

Report:

PROCEEDINGS OF THE 1978 CONFERENCE ON THE SEATTLE AND DENVER INCOME MAINTENANCE EXPERIMENTS

OCTOBER 1979

State of
Washington
Department
of Social & Health
Services



**Analysis and
Information Services
Division**

PROCEEDINGS OF THE 1978 CONFERENCE
ON THE SEATTLE AND DENVER INCOME MAINTENANCE EXPERIMENTS

edited by

Joseph G. Bell

Patricia M. Lines

Michael Linn

Office of Research

Department of Social and Health Services
Analysis and Information Services Division
Olympia, Washington 98504

October 1979

STATE OF WASHINGTON
Dixy Lee Ray, Governor

DEPARTMENT OF SOCIAL AND HEALTH SERVICES
Gerald J. Thompson, Secretary

ANALYSIS AND INFORMATION SERVICES DIVISION
Hans Carstensen, Director

OFFICE OF RESEARCH
Timothy R. Brown, Ph.D., Chief

This is the report of proceedings of the 1978 conference on the Seattle and Denver Income Maintenance Experiments held in Washington State. The studies were conducted under contract to the Department of Health, Education, and Welfare. The contents do not necessarily reflect the official view of DHEW.

Library of Congress
Catalog Card Number
79-620060

ACKNOWLEDGEMENTS

This book is the work of many hands. It first took form as a conference agenda, developed by Charles Froland, Michael Linn, and David Parker. Preparations for the conference and arrangements with the authors were made through the State of Washington SIME Office; Curt Funk and Frank Waynewood were responsible for much of the success of the three-day meeting. Patricia Lines was the principal editor of these proceedings. She provided a complete transcript of the conference, and worked closely with the authors to shape the original papers and discussions into their present form. Bonnie Bayes assisted in the production activities, and, along with Michael Linn, edited many of the papers. Pam Busick typed most of the edited drafts and served as proofreader. Linda Lynd and Joan Dickhaut helped with the typing and the final assembly of the volume.

In addition to the people who had a direct role in producing these proceedings, credit should be given to the numerous other individuals who helped the authors gather, analyze, and report on the data. The management and administrative staffs of SRI and MPR offered continuous assistance on the project. Financial support was provided by the Department of Health, Education, and Welfare.

Without all of these contributions, this publication would not have been possible.

Joseph Bell
SIME Project Director

CONTRIBUTORS

Miriam Aiken, Jodie Allen, Marcy Elkind Avrin,
Harold Beebout, David M. Betson, Matthew Black,
Gary Christopherson, Constance F. Citro,
Virgil Davis, Barbara Devaney, David Edson,
Henry Felder, Linda Drazga, Margaret Grady,
David H. Greenberg, Lyle P. Groeneveld, Arden Hall,
John Hall, Harlan Halsey, Michael T. Hannan, Alan M. Hershey,
Terry R. Johnson, Richard A. Kasten, Michael C. Keeley,
Barbara H. Kehrer, Kenneth Kehrer, David N. Kershaw,
Larry Manheim, Cheri Marshall, Myles Maxfield, Mary E. Minchella,
Robert A. Moffitt, William Morrill, Arnold H. Packer,
John Pencavel, Kristen Puckett, Philip K. Robins,
Robert G. Spiegelman, Ricardo Springs, Michael Stern,
Manika Sukhatme, Peggy Thoits, Cynthia Thomas,
Billy J. Tidwell, Nancy Brandon Tuma, Samuel Weiner,
Richard West, Charles Wolin

TABLE OF CONTENTS

ACKNOWLEDGEMENTS		<i>iii</i>
PREFACE		xv
I. INTRODUCTION AND PERSPECTIVE		
An Introduction to the Seattle/Denver Income Maintenance Experiment: Origins, Limitations and Policy Relevance	Jodie Allen	3
The View From Congress	Michael Stern	23
The Administration's Approach: The Premise of the Better Jobs and Income Proposal	Arnold Packer	29
What We Have Learned About Research and Policy-Making	William Morrill	39
What We Have Learned From the Experiments	Charles E. Metcalf and David N. Kershaw	45
Research Issues in SIME/DIME: An Overview	Richard W. West	57
Summary of the Research Findings	Robert Spiegelman	67
Designing Income Maintenance Programs: Evidence, Problems and Suggestions for Further Research	Kenneth C. Kehrer	75
II. SIME/DIME DESIGN AND DATA HANDLING		
The Structuring of SIME/DIME: An Overview	Robert Spiegelman	91
Data Collection	Cheri Marshall and Gary Christopherson	113
Data Collection From Collateral Sources: SIME/DIME Experience	John Hall and Kenneth Kehrer	127
SIME/DIME Data Processing	Virgil Davis	143

Validating SIME/DIME Income Data	Harlan I. Halsey	161
SIME/DIME Data Files Available to the Public	Constance Citro, Miriam Aiken and Kristen Puckett	181
III. IMPACT ON WORK AND LABOR SUPPLY		
Work Effort Response by Race, Site and Experimental Duration	Philip K. Robins and Richard K. West	205
Methodological Issues in Labor Supply Analysis	Michael C. Keeley	229
Employment Issues for Low-Income Women: Values and Choices	Linda Drazga, Barbara Devaney, Margaret Grady and Manika Sukhatme	247
Unemployment Insurance and the Duration of Unemployment	Henry E. Felder	267
The SIME/DIME Manpower Program	Arden Hall	283
IV. COST IMPLICATIONS OF THE LABOR SUPPLY RESPONSE		
Generalizing Experimental Results With Microsimulation	Harold Beebout and Myles Maxfield, Jr.	311
National Labor Supply Effects and Costs of a Negative Income Tax	Myles Maxfield, Jr. and Philip K. Robins	319
Support Levels and Tax Rates: Implications for a Negative Income Tax Program Design	David Edson and Myles Maxfield, Jr.	353
The Macroeconomic Impact of a Negative Income Tax	Robert Moffitt	377
A Simulation of the Program for Better Jobs and Income	David Greenberg, David Betson and Richard Kasten	391
V. MAJOR IMPACTS ON THE FAMILY		
Income Maintenance and Marriage	Michael T. Hannan, Nancy Brandon Tuma, and Lyle P. Groeneveld	413

Methodological Issues in Analysis of Marital Stability	Michael T. Hannan	443
Progress Report on Studies of the Effects of Income Maintenance on Infant Health Status	Barbara H. Kehrer and Charles M. Wolin	455
Impacts of the Seattle and Denver Income Maintenance Experiments Upon Psychological Distress	Peggy Thoits	471
SIME/DIME Housing Study: Policy Questions and Progress Summary	Cynthia Thomas	489
The Effect of Income Maintenance on the Utilization of Subsidized Housing	Marcy Avrin	499
The Effects of Alternative Negative Income Tax Programs on Migration	Michael C. Keeley	511
Effects of Income Maintenance on School Performance of Children	Larry M. Manheim and Mary E. Minchella	531
SIME/DIME Child Care and Public Policy	Samuel Weiner	551
Impact of Future Income Uncertainty on Saving Decisions	Matthew Black	569
VI. ADMINISTRATIVE IMPLICATIONS		
Effects of Income Accounting and Reporting Periods in Transfer Programs	Ricardo Springs	597
The Colorado Monthly Reporting Project	Alan M. Hershey	613
Participant Responses	John Hall and Billy Tidwell	625
Alternative Methods of Scaling Welfare Payments to Family Composition	Terry Johnson and John H. Pencavel	653
Welfare Reform and State Fiscal Flows	Myles Maxfield, Jr. and David Edson	677

APPENDIX A: Monthly Income Report Form	699
APPENDIX B: Authors and Editors, Biographical Notes	707
APPENDIX C: Conference Participants	715

PREFACE

This volume presents a preliminary analysis of the data accumulated thus far in the largest controlled social experiment ever undertaken, the Seattle, and Denver Income Maintenance Experiments (SIME/DIME). The concept of social experimentation is itself not new; modes of organizing for public welfare and distributing benefits and services within communities are virtually always first attempted on a small scale prior to their broader implementation. Yet the use of rigorous statistical methods to test the effects of policy changes is a product of only recent years, and the income maintenance experiments are perhaps its foremost example. This research was designed to assess the social and economic impacts of a welfare system based on the principle of a negative income tax. Earlier and more limited studies were conducted in New Jersey and Pennsylvania, in Gary, Indiana, and in two rural settings. SIME/DIME was the most carefully designed of the series; it contains the largest sample of families and the greatest number of experimental variations. Thus, it allows the examination of a wide range of issues, and permits a highly refined analysis of the probable effects of various reforms to the welfare system.

The papers were first presented by researchers working under government contract on SIME/DIME, at a conference held in May 1978 at Orcas Island, Washington. The conference was a first step in sharing this work with other researchers and with selected representatives from administrative and

legislative branches of government, members of community organizations, and other interested persons. The setting allowed all those present to review the findings to that date, and to discuss their implications for policy. The primary goal was to promote communication between the researcher and the policy maker: it is too often the case that these groups simply speak past each other. Researchers need to be reminded of the importance and the difficulties of communicating technical findings to non-researchers, particularly to those who may base their actions on this knowledge. Policy makers can benefit from an understanding of research findings, and also of the assumptions that are made and the methods followed, so that results may be seen in their proper light. Excerpts of the discussions that took place after each presentation at the conference, included here, and reflect the value of this interaction.

The papers and discussion contained in this book are aimed toward several purposes. First and most importantly, they provide a rare opportunity to examine the nexus between research and policy-making. A number of excellent papers discuss this interchange from different perspectives, including those of policy makers and their staffs. The discussions demonstrate that productive exchange between government decision-makers and the research community is in fact possible, and under appropriate conditions, may even be expected to occur. This collection will provide instructive reading for the student of government, and should prove useful as a resource for legislators, administrators, and researchers. It would also serve well as a supplementary text for courses in policy analysis and social welfare.

Second, the papers represent a first attempt by the SIME/DIME research team to provide a comprehensive and up-to-date report to policy makers and the public. Those who are responsible for welfare policy can consult this work, to assess the research design and methodology, and ultimately to assist them in determining how the experimental findings and other types of knowledge may contribute to effective reform of the welfare system. The material gathered here on issues surrounding the question of a negative income tax, and the implications of this research for welfare program operations, raise the likelihood of informed decision making in this area.

Third, the papers serve as an introduction to the overall research plan, and will be useful to independent researchers who are interested in studying the data. One set of papers describes how the data was collected, the form it is in, the ways in which it has been processed, and how sets may be obtained for additional analysis. Other papers summarize the studies performed to date, enabling the researcher to chart new or extended territory, rather than repeat work already done. Analysis under the primary contract with the government--by Stanford Research Institute (SRI) International and Mathematica Policy Research (MPR)--is expected to continue for a few more years, but research under other auspices is highly encouraged. SIME/DIME has produced one of the richest data bases that is currently available on lower-income households, and the largest developed through a controlled experiment. As has been the case with previous surveys and panel studies, the research potential of this data will undoubtedly attract many independent scientists in the years to come.

The papers presented at the conference and edited for this book are, in many cases, summaries of longer and more detailed research studies. Citations in each paper inform the reader as to the existence and availability of these original documents. Bibliographies have also been prepared which provide a comprehensive listing of the work done on the individual topics. Editing and abridging of the conference papers was carried out with the active assistance of the authors. It was not possible to contact all of the discussants for approval of deletions and word changes, so the discussion sections were prepared directly from the conference transcript. This transcript, along with the original papers (a combined total of nearly 3,000 pages), is on file at the State of Washington SIME Office.

These papers do not attempt to provide complete or definitive answers to the policy questions underlying the experiment. Some serve chiefly to explain the theoretical framework and the methods for analysis. A few focus on the behavior of the control group (matched for comparison to the SIME/DIME sample) for non-experimental research--to determine, for example, differences in welfare program participation between workers covered by unemployment insurance and those who are not. The largest number of the papers report on the interim findings of the experiment, and discuss the directions indicated for public policy. Others use the results obtained as a basis for projecting the effects of NIT programs on national populations, or for estimating the costs of other kinds of programs. As the first of a number of volumes to be published on the experiment, it sets a background

for interpretation and future analysis of the data.

This collection is diverse, in terms of both the level of detail in the papers and the scope of their subject matter. It is not intended to be read from cover to cover, but can most usefully be examined selectively. General-interest readers will find it helpful to cover Part I, which provides an introduction and perspective on the research, and proceed through other sections. Active researchers may prefer to focus first on Part II, covering design and data management issues, then turn to specific papers in subject areas of interest. The more detailed research memoranda can also be referred to, as necessary. For all readers, some acquaintance with each of the major subject areas is recommended, because of the close interrelationship among the individual research topics. For example, to consider only the effects of reduced work effort in estimating the number of people likely to participate in a negative income tax program would be incomplete; changes in marriage patterns and an increase in the number of single parents with children may also be expected to occur, and will significantly affect the structure of the eligible population. Similar relations among other areas are revealed as well, reflecting the breadth of household response to the negative income tax. One of our chief objectives for this volume is to convey a sense of the wide range and interconnectedness of program impacts. If this is achieved, we are confident that readers' time spent with these papers will be well rewarded.

The Editors
Seattle, Washington

I

INTRODUCTION AND PERSPECTIVE

AN INTRODUCTION TO THE
SEATTLE/DENVER INCOME MAINTENANCE EXPERIMENT:
ORIGINS, LIMITATIONS AND POLICY RELEVANCE

by

Jodie T. Allen
Special Assistant to the Secretary
Department of Labor

The papers presented in these proceedings contain a great deal of discussion of the design, implementation and outcomes of the Seattle and Denver income maintenance experiments (SIME/DIME). As a prelude to these presentations and discussions, I will try to place SIME/DIME in three frames of reference, by relating it first to its own history and that of its companion projects; second to its uses and limitations as the largest and most complex social experiment completed to date;¹ and third, to its impact on policymaking in the income maintenance area.

ORIGINS

The concept of income maintenance experimentation originated in 1967 as an outgrowth of mounting interest in the defects of our current welfare system, and in the need to design better methods of income supplementation. Analysis of the various alternatives most widely promoted--ranging from universal children allowances to negative income taxes (NIT) and wage subsidies--quickly stalled for lack of data. There was little, if any, hard data on behavioral responses to the various incentives, both favorable and perverse, implicit in any of these schemes. Since many of these so-called "induced" effects of income maintenance policy have potentially huge fiscal

or societal effects, it appeared sensible to base any long term program of income maintenance on a sound program of research. It also became clear that, given the subtlety of the anticipated individual behavioral responses, it would be impossible to assess accurately many of the most important potential consequences of income maintenance reform using traditional social science research techniques. That is, the required information could not be obtained either by extrapolation from static cross-sectional data or from information gathered from the type of relatively uncontrolled demonstration projects previously attempted. The problem gave birth to a new idea--that of initiating carefully controlled experimental projects designed to yield statistically reliable data on stated hypotheses.

Among the most important questions in the income maintenance area which appeared to require exploration through the experimental method were the following:

- How will proposed programs affect the incentive to work?

Theory suggests two major effects--an "income effect" associated with the size of the basic guarantee offered to a person with no other income, and a "substitution effect" associated with a reduced return on work effort caused by the reduction of welfare benefits as earnings increase. We needed to know, if standards were raised to a point where some persons on welfare might be almost as well off as persons in low paying jobs, whether individuals would stop or cut back work effort in order to qualify for increased public assistance benefits (the income

effect). We also needed to know whether more liberal provisions for the retention of earned income while receiving assistance payments could encourage individuals on public assistance to enter the labor force or increase their work effort.

- What interactions will occur with manpower and work-related programs and services, such as jobs creation and training, day care and transportation services? Will skill training and other human investment programs serve to offset the long run impact of any reductions in work effort induced by the income maintenance system?
- Will proposed programs affect mobility--in particular, will they tend to accelerate, decelerate or reverse the current rural-urban migration pattern?
- Will welfare policy changes affect family stability and, if so, how?
- Will certain types of programs--particularly child-oriented allowance systems--produce an effect on family size?
- Will a new system affect demand for social services, both public and private? Will the injection of additional money into a community of itself promote development of private medical, legal and educational services? Or will the burden continue on public sector services?
- Will consumption patterns change among low-income families? How high must payments be before families will budget significant amounts of money for long-run investments in health, education and general well-being?
- What will be the general effects on the social and economic life of

a community? Will prices of goods and services change? Will community cohesiveness be enhanced? Will the location of businesses shift?

A parallel set of questions focuses on how new policies might affect different population groups and different areas. Will there be differences for urban or rural populations, whites or non-whites, female-headed or male-headed families; or for aged or non-aged individuals, or persons in families compared to unrelated individuals and childless couples?

It is clearly not possible to obtain reliable answers to all these questions in a single experiment and some of them probably can't be answered at all. At the same time, experiments are costly and difficult to design and implement. Consequently, income maintenance planners in both the Department of Health, Education and Welfare (DHEW) and the Office of Economic Opportunity (OEO) were forced to keep the number of experiments as small as possible, attempting to limit them to a series of well-controlled and carefully designed projects, each of which was intended to be a necessary and integral part of an overall effort. An overall strategy for income maintenance experimentation and related research was developed for DHEW by Larry Orr, Robinson Hollister and others² then at the Institute for Research on Poverty at the University of Wisconsin. While the full strategy has never been implemented, four major projects were launched: the New Jersey, North Carolina/Iowa (rural), Seattle/Denver and Gary experiments. The New Jersey and rural experiments were sponsored by OEO and the Seattle/Denver and Gary experiments were sponsored by DHEW.

All of these experiments focused on the controversial problem of work incentives in an income maintenance system. The first two OEO experiments

dealt almost exclusively with this crucial question and were designed to determine the effects of financial treatments on the work response of male-headed families in both rural and urban areas. Male-headed families were of particular interest since they represent a large portion of the working poor and generally are not covered under existing welfare programs. Disincentives to work in an income maintenance system should be most discernible in the work effort of those working already.

The New Jersey project concentrated on the urban poor in five communities in New Jersey and Pennsylvania and completed its three-year operational phase in 1972. Original analysis of the results was completed in 1974 although additional analysis of the data has subsequently been undertaken by others including a series of studies sponsored by the Brookings Institution in 1976³ and a recent study by the Rand Corporation.⁴ The rural income maintenance experiment in Iowa and North Carolina tested the work incentive effects of an NIT on predominantly rural populations. The experimental phase was completed at the end of 1972 and analysis emerged in December 1976.⁵

The more complex DHEW experiments focused on a variety of issues of major policy concern in addition to the effects of income transfers on work incentives. The first and most important DHEW experiment was the SIME/DIME, and it remains the most comprehensive of all the experiments. The population studied included white, black and Chicano families, having either one or two parents present. The Seattle portion of the experiment began in November 1970. The Denver portion of the experiment began in late 1971. The two site samples included over 5,000 families. The experiment was intended to

test the combined effects of an NIT with a manpower program in areas of high and low unemployment. A new hypothesis to be tested was that manpower programs, including job training, counselling and vocational guidance services and day care services, could serve to offset the presumed work disincentive effects associated with cash transfers.

The last of the experiments, the Gary, Indiana experiment, also sponsored by DHEW, paralleled the OEO experiments in that it focused on the impact of alternative sets of income maintenance structures on the work-leisure decision. Approximately 1,800 families, 60% of them receiving experimental treatments, were enrolled in a three-year experiment. Experimentation began in 1971 and the analysis was ready in 1977.

In addition to measuring the impact of direct cash transfers on various aspects of family life, Gary, like the Seattle/Denver experiment, had a secondary treatment--in this case, social services. The experiment was originally intended to measure the demand for, and the impact of, a variety of social services (such as day care, homemaker services, housing and medical information services) in the presence and absence of income maintenance transfers although in this respect the experiment had little success.

DISTINCTIVE CHARACTERISTICS OF A SOCIAL EXPERIMENT

The papers presented in these proceedings discuss the lessons learned thus far from the SIME/DIME experiments. Before examining them, it is useful to delineate the legitimate role of experiments in the social policy area and to establish reasonable expectations for such work. In addition to the substantive issues to be explored, issues appear relating to the intent and

form of a major social experiment that are quite different from issues associated with a demonstration project. In its simplest form all that a demonstration attempts is to show that a particular "treatment" can be administered to a given population and that, when it is, the status of this particular population will be altered in some discernible fashion. No attempt is made to control for the effect of non-treatment factors affecting the treated population. Thus, no rigorous generalization of the results to other populations, other times, or slightly altered treatment variables, is possible. One might, for example, decide to demonstrate a family allowance by giving such an allowance to residents of a particular town and observing changes, if any, afterwards. This would demonstrate that such a plan can be administered but would not allow us to infer which of the changes were due to the allowance and which to other factors. Further, we would not have quantitative measures of the relationship between the level a type of allowance and the observed responses.

An income maintenance experiment has a more ambitious objective. It seeks to provide information on the effects of a given "treatment" which can be generalized to populations other than the one treated. Further, an income maintenance experiment seeks not only to discover whether the work effort of recipients will increase or decrease given a certain support level and tax rate, but also to develop a statistical description of this relation from which one can infer what the labor force response will be to variations in the particular standard and tax rate chosen. Similarly, an experiment might focus on the birth rate response to changes in income maintenance

policies affecting the cost of children. The analysis can reveal not only that the birth rate changed, but also the predicted amount of change at various levels of payment.

In short, an income maintenance experiment seeks to provide information to the policymaker who can then rationally choose among numerous options, knowing the likely social consequences of any particular choice.

To achieve this objective in an experiment an attempt is made to simulate, as much as possible, the laboratory conditions of the physical sciences. Care is taken to gather adequate information on a non-treatment group which is similar enough to the treatment group to permit comparisons. Other characteristics of the subjects and their environment are continuously measured so that non-experimental changes in variables can be taken into account. The experimenters try to isolate differences in the subjects' responses attributable to differences in personal characteristics. Further, the number of subjects in both the treatment and non-treatment groups, taking into consideration the range of variation in policy variables to be tested, are sufficiently large to allow the application of the principles of statistical inference to determine the statistical significance of any observed differences in response. Without these controls, we cannot arrive at conclusions with a high degree of confidence.

To sum it up, a demonstration is relatively easy to conduct, and can be useful in terms of working out the "bugs" in the administration of a particular program or in generating public awareness or acceptance of such a program. And this can be very important. An experiment, while conceptually far more difficult and often more costly, yields more powerful information for the policymaker.

LIMITATIONS

Before undertaking a major social experiment, it is important to fully understand the costs and difficulties of the experimental method. Several major problems must be addressed.

For example, an experiment applies varying treatments including a "null" (or control) treatment to estimate functional relations. Because we are using people, questions of inequity will arise. Inequity also occurs in any demonstration which covers only a small geographical area or sample but it grows more visible when similarly situated persons in the same area receive different treatments. The appearance of inequity may be even more troublesome when that treatment is money. Dealing with people as experimental subjects also raises innumerable ethical and practical problems. Harold Watts has best dramatized this problem by imagining the difficulties of the chemist if he had to secure the individual cooperation of each molecule he wished to observe in an experiment.

Experiments are also very expensive because they must cover a sufficiently large number of people to produce statistically significant results. One should carefully question if field experimentation is necessary, if it will produce sufficient solutions for social problems, and whether cheaper alternatives are available. Experimental exploration of some problems, particularly those involving large scale community responses or the interaction of complex economic and social institutions may not be practical. In other cases experimentation is not necessary.

The existing world offers many instances of natural experimentation-- where different program variants are tried in similar settings, or similar programs are tried in different settings. A rigorous evaluation effort, employing many of the same techniques as social experimentation, may suffice.

Since most social and economic processes are rather slow to adjust, the experimental period must be rather long. Allowance must also be made for distortions in the behavior of the experimental population produced during the start-up phase of the experiment by the newness of it, and during the last period by the anticipation of its termination. How long is long enough to allow researchers to begin to see patterns of permanent response? Observed consumption patterns may provide some indication of the extent to which the income maintenance payments are perceived as transitory--but one cannot infer with certainty that labor supply and consumption responses are identical in this respect. Control for the time factor was built into the Seattle/Denver experiment by extending the life of 25% of the sample to a five year period. Preliminary analysis indicates that even a two year time difference in program duration may produce a very substantial, statistically significant increase in disincentive effects for male heads of households.

Social experimentation also shares many of the problems of more traditional research on human subjects. In both demonstrations and experiments the problem of Hawthorne effects arises, that is, the participants' reaction to being experimental subjects can be confounded

with the effects of the treatments; but in this respect the experimental mode has the advantage of a control group which, since it is also exposed to the experimental observation techniques, can presumably provide a reference for isolating the effects of the experimental treatments themselves. Demonstrations and evaluations also share common problems with regard to duration bias, attrition bias and learning effects (i.e., lagged, imperfect or evolving understanding of treatments by recipients). Other shared problems include adjusting for participation interaction different in scope or duration than that which might be expected under a full scale national program, and preserving the confidentiality of the data supplied by program participants for research purposes.

Of course demonstrations, evaluations and experiments alike must design questionnaires, administer them to the studied population and process and analyze the data thus produced.

Finally there is the problem of environmental change. Two major disruptive occurrences have rocked the income maintenance experiments. The first was the introduction by the State of New Jersey, shortly after the New Jersey experiment began, of a welfare program for unemployed fathers which offered eligibility and superior benefits to many of the treatment and control families. As a result, many of the experimental cells were effectively decimated and comparisons with the control group, many of whom were receiving more generous transfer benefits and facing stronger work disincentives than experimental families, lost most

of their meaning. The second calamity was a precipitous rise in the unemployment rate in Seattle from less than 3% at the time the design phase of the experiment was initiated to a high of 13% shortly after the experiment got under way. This was obviously disastrous for an experiment designed to measure the pure effects of income maintenance on labor supply. Here again the experiments have the advantage of the presence of a control group subject to the same exogenous shock, as well as the opportunity for some flexibility. In the case of New Jersey, a higher guarantee treatment was introduced to improve the competitiveness of the experiment versus the welfare program. In the case of Seattle it was possible, given the cooperation of the State of Washington, to shift half of the experiment to a low unemployment area--Denver--and thus, in the process, introduce an additional dimension to the experimental design by allowing measurement of interactive effects with the unemployment rate. (Here again, preliminary analysis indicates that such effects may be important in measuring the work effort responses of persons normally in the labor market.)

Clearly, there are many pitfalls in social experimentation. Throughout the papers, there are countless references to how the SIME/DIME researchers coped with them. But many lessons have been learned already, and the lessons have been translated into improved experimental design and operations in the later experiments.

A first major lesson is that a successful program of social experimentation must rely heavily on the cooperation of state and local

officials. Numerous impending disasters in program design and operations were averted by the sympathetic and knowledgeable cooperation of officials in the states of Washington and Colorado.

Second, experimenters in their desire to produce efficient sample designs should be careful not to define the policy space too narrowly. External changes or shifting policy interest can overtake them. In this respect serious mistakes were made in all the income maintenance experiments. For example, on the seemingly obvious assumption that no one would consider imposing confiscatory taxes on the poor, no one of the experiments tested the impact of cumulative marginal taxes (including payroll and income taxes as well as benefit reduction rates) in excess of 70%. Subsequent experience with welfare reform has indicated considerable Congressional interest in cumulative taxes approaching or even exceeding 100%. In New Jersey the focus on male-headed families permitted a small sample size but at the cost of criticism from Congress which wanted to know more about current welfare recipients, predominantly female-headed families. Indeed, I think it is generally conceded now that in matters of sample size it is very easy to be "penny wise and pound foolish," with the rural experiment as a sort of extreme case in which it is difficult to conclude anything except that farmers know how to manipulate negative taxes just as well as positive ones.

We experimenters probably outsmarted ourselves again in the case of sample design. The relatively modest income guarantees first offered in New Jersey and the design of the sample assignment model--which, in the

interest of sample efficiency tended to minimize project costs by, in simplest terms, assigning better off people to the richer transfer plans and poor people to the poorer--not only aggravated considerably the problem of defections to welfare and general attrition, but introduced a serious complication into the subsequent analysis of labor supply effects. As you will hear in later presentations, "assignment model bias" is also a complicating factor in the SIME/DIME analysis although in this case the far larger size of the sample has allowed researchers to compensate for the bias in the analysis of most effects. Similarly in Seattle and Gary, the initial guarantees also proved too low to attract substantial numbers of families away from the cumulative cash and in-kind welfare benefits already available.

Further, great care must be taken in devising rules of operation for the experiment. The design of such rules is a complex task. They must provide as complete a description of program operations as a normal program authorization statute, and also specify the peculiar obligations and features of the experiment. The rules may provide incentives for behavior neither wanted or intended by the experimenters. To cite one of many examples, a rule which "taxes" benefits from other welfare programs at 100% prompted a family in Seattle to abandon an aging grandmother and a blind daughter, both receiving regular welfare benefits, in order that the residual unit could qualify for the experimental transfer. Not all such unintended side effects can be avoided but the importance of such occurrences should be carefully

considered by experimental rulemakers.

In addition, rulemaking has all the problems of statute writing--to some extent the more explicit the letter of the law the more easily is the spirit violated. If as in the Seattle experiment, subsidies are offered for vocational training and related expenses at the going price, some enterprising participant will demand 3000 flying hours to become a commercial pilot. Another will wish to purchase \$2000 of photographic equipment as part of a photography course. The rules must thus incorporate some catch-all generalities to give the administrators discretion in handling such cases.

When making rules, it is important to remember that experimental participants are just as "human" as anyone else and a great deal more clever than is commonly supposed. Loopholes in the experimental laws are quickly detected and exploited. Allowing persons to deduct medical and other expenses from income, when net income determines entitlements, will produce sizeable deductions. Requiring recipients to submit receipts will effect prompt and sizeable reductions in such claims. Taxing "unearned income" at 100% may seem like a smart way to minimize program costs but in no time at all, this unearned income will "dry up." The problem of misreporting or rule-bending is particularly serious in an experiment. At the very least it is disillusioning and, further, the experimenters may, and have, found themselves involved in disputes with tax and transfer program enforcement authorities. But misreporting can also challenge the very purpose of an experiment. Dishonest reporting means that the data collected will be biased to an unknown extent. Experimenters have come, perhaps too

slowly, to realize that in primarily self-administered systems, it is essential that opportunities for abuse be minimized and that participants be made aware that auditing and compliance enforcement are rigorous.

Finally, adequate data processing can be the Achilles heel of social experimentation. The data gathered by experiments are voluminous in size and complex in design. All of the experiments to date have suffered from the failure to insure the adequate data processing capability was in place before the start of the experiments to permit ongoing analysis of experimental data as it is collected. SIME/DIME and similar experiments have encountered such problems and it has delayed the availability of experimental findings and further strained the patience of the policymakers who have, since the inception of the experiments, been pressing for results. Perhaps even worse, unprocessed data minimizes the possibility of "feedback" from early field results to later design modifications. This last point with regard to sequential program design is important. Although there is a presumption against altering the rules of the game in midstream, given the high cost and lengthy duration of social experiments, a Maginot Line philosophy does not become the social experimenter. External circumstances change, policy interests shift and some things just don't work. And in these cases some flexibility is in order.

All of this should in no way be interpreted as a counsel of despair. In fact you will find that, without exception, everyone associated with the income maintenance experiemnts remains tremendously enthusiastic about them. Indeed the contagion has spread to other fields including training, health,

and most recently, employment. At the least we can say that social experimentation is feasible.

If there was one common failing among all of us early experimenters, I think it was a failure to appreciate not just the theoretical, but the practical difficulties of operating social experiments. Each of the experiments is in itself a mini-welfare system with all the well known problems of such systems, aggravated by the need to collect analyzable data. The realization that running a transfer program is easier said than done has been a sobering one for all the income maintenance experimenters, and they are uniformly less bright-eyed than they once were. But as byproduct, a great deal of administrative know-how has been developed throughout the various experiments, and the application of this combination of theoretical insight and practical technique to the problems of the actual world of welfare may, indeed, be one of the major social outputs of the income maintenance experimental program.

IMPLICATIONS FOR NATIONAL POLICY

I will conclude with a very brief comment on the relevance of the experiments for policy purposes. We might ask whether the experiments were worth the \$100 million or so they cost the taxpayer. I think that the answer to this is so clear that I will cite only two examples in its support.

Take for example the major finding of the SIME/DIME experiment on monthly reporting of income by welfare recipients. We now know this is administratively feasible, can improve the responsiveness of the welfare system to the changing need of recipients and can considerably reduce

the cost of the system to taxpayers by reducing the number and size of incorrect payments. It is estimated that such a system, which the Administration is now testing in several states, could in itself save some \$800 million a year in welfare costs while improving the level of service offered to the system's clientele.

In the area of program design the impacts are even more dramatic. The SIME/DIME findings have made it possible to improve substantially our direct and indirect estimates of the cost of various welfare alternatives. While earlier experiments could estimate gross program effects on labor supply, SIME/DIME for the first time permits isolation of guarantee and tax effects which, in turn, could be incorporated into the models used by DHEW for estimating the costs and impacts of new programs. We now know that the costs of labor supply responses can be very substantial unless care is taken in program design to minimize or offset such effects. These and other findings have already exerted considerable influence on the design of the Administration's Better Jobs and Income Program for welfare reform with its emphasis on jobs, rather than cash transfers as the major vehicle for assisting the employable poor. Unless subsequent research or experimentation contravene or modify the SIME/DIME findings, it seems very likely that they will continue to affect the income maintenance policy debate for years to come.

NOTES

1. Dollar expenditures on the various housing allowance projects far exceeded the SIME/DIME budget, but relatively little of the total expenditures under those projects was for controlled experimentation.
2. Larry L. Orr, Robinson G. Hollister and Myron J. Lefcowitz, Income Maintenance, Interdisciplinary Approaches to Research (Chicago: Markham Publishing Co., 1971).
3. Joseph A. Pechman and P. Michael Timpane, Work Incentives and Income Guarantees: The New Jersey Negative Income Tax Experiment (Washington, D.C.: Brookings Institution, 1975).
4. John F. Cogan, "Negative Income Taxation and Labor Supply: New Evidence from the New Jersey-Pennsylvania Experiment" (Rand Corporation: Santa Monica, February, 1978).
5. John L. Palmer and Joseph A. Pechman, Welfare in Rural Areas: The North Carolina-Iowa Income Maintenance Experiment (Washington, D.C.: Brookings Institution, 1978).

THE VIEW FROM CONGRESS

by

Michael Stern
Staff Director,
Senate Finance Committee

In 1970, testifying before the Finance Committee on the proposed Family Assistance Plan, then Secretary of Department of Health, Education, and Welfare (DHEW) Finch said, "The first objective of the program is to provide strong work incentives in the welfare system, both for those on welfare and for those working people who have a high risk of entering the welfare population. The central reason for the particular reform structure embodied in the family assistance plan is its importance to work motivation in our society." And he subsequently said: "The preservation of the system which provides a prima facie incentive, a clear financial reward for family breakup, seems vicious and irrational."

Now I'd like to quote from testimony of Secretary Califano earlier this year before the Senate Finance Committee on the Administration's welfare reform program: "The existing welfare system contains substantial family splitting incentives." He also said, "Under the President's proposal low-income fathers would no longer have an incentive to leave their wives and children in order to make families eligible for cash support." And he said, "For those expected to work the tax rate should not under our plan exceed 52%. This will insure an adequate work incentive."

Has anything changed in the two years between those two statements?

What has changed is we now have information from demonstration projects, particularly the ones in Seattle and Denver. Unfortunately, the information that is now available didn't get worked into Secretary Califano's statement.

During the course of the hearing earlier this year, Senator Moynihan, the Chairman of the Finance Committee's Public Assistance Subcommittee, asked Henry Aaron, the Assistant Secretary of DHEW for Planning and Evaluation, "What studies or other evidence exists to back up Secretary Califano's assertion that the President's welfare reform proposal would provide an adequate work incentive, and that the existing welfare system contains substantial family-splitting incentives that the President's welfare proposal would correct?"

This was Mr. Aaron's answer: "The real point that I think should be expressed is that one has to ask whether a program that drastically alters the incentives the people face--like for the first time providing aid to two-parent families and not requiring as a condition of aid that two-parent families split; a program that for the first time assures that a fully employed worker can support his family above the poverty threshold deserves the presumption that it, on its face, improves these incentives. I would suggest that really the burden of proof ought to rest on those who deny that such a patently sensible change in the incentives that American families face. The burden of proof should rest on those who deny that it will improve, over the long haul, the behavior and the prospects for family stability."

Subsequently, Senator Moynihan responded by saying, "Dr. Aaron, that

is not a burden which will weigh heavily on the members of this committee. They will not think they must go out and disprove your data. They are not-- God bless them and let the republic rejoice--they are not social scientists."

In this regard Senator Moynihan quoted from what he terms Forrester's Law, named after Jay Forrester who invented the memory-core computers: "With respect to complex social situations, intuitive solutions are almost invariably wrong."

Despite these assertions by DHEW officials, in the absence of any actual information, welfare reform efforts in 1970 and 1972 failed in the Senate. Let me suggest why this happened. Senators have very strong feelings on the subject of welfare and welfare reform, but their views are polarized on what welfare reform means. At one pole, there are senators who believe that welfare reform means providing a federal guaranteed minimum income and extending benefits to people that are not now covered: single people, childless couples, and couples with families or families with male heads. At the other pole, there are those Senators who believe that welfare reform means reducing benefits, increasing work requirements, and eliminating fraud and abuse. There is no compromise between the two views.

In 1970 and 1972 no majority could be found for any one proposal; in fact, there was a majority against every proposal. So, although the senators may have had very different reasons for voting the way they did, in fact, a majority combined to oppose any particular proposal. Take for example, a proposal to provide a guaranteed income at a modest level with some kind of work requirement. A majority of senators would vote against it--part of

them on the grounds that they didn't believe in a guaranteed minimum income at all (since it would put millions more people on welfare), and part on the grounds that benefits were too low and for some recipients would be cut back. Some might also feel it punitive to have a work requirement. I believe the proponents of such an amendment mustered 41 votes in the first vote, and in the roll call on every subsequent version of the proposal there were less votes. Finally the sponsors gave up, because anytime they modified their amendment to have a higher guarantee level they lost votes on one side, and everytime somebody moved to have a lower guarantee level he lost votes on the other side.

That is what happened in 1972. In my opinion, attitudes on welfare reform are no less polarized today. However, one thing could change the situation and enhance the chances of passing welfare reform legislation-- more information, new information of the type that has been developed in the SIME/DIME experiments. That is the reason why these experiments are so important and why members of the Senate Finance Committee are following them so closely.

Also, the kind of results that are reported here really show the importance of testing alternative approaches. I suspect that the way you give people money is as important as the fact that you give it at all. The public clearly believes this; that becomes apparent when you look at the differences in public acceptance of programs that are employment-related, like social security or unemployment compensation, compared to the public attitude toward welfare programs, such as Aid to Families with Dependent Children (AFDC).

DISCUSSION

RIKKI BAUM: Do you feel that there is any chance that Senator Moynihan's Subcommittee is going to report out this bill in committee before they get tax legislation later on?

STERN: The House has taken the view that virtually anything in the Senate Finance Committee's jurisdiction is revenue legislation and therefore, the Committee is constitutionally restricted against acting on legislation before it passes the House. I think the Committee would wait until something passes the House, especially since the House has moved as far as it has. As far as prospects go this year, however, I don't expect separate welfare reform legislation. I base that on the fact that the Speaker of the House, after consulting with the President and other officials at the White House, produced a list of top-priority legislation which did not include welfare reform. So I doubt that the House as a whole will act on welfare reform legislation this year; there may be some committee action, but that would be all.

UNIDENTIFIED PARTICIPANT: What about H.R. 7200?

STERN: H.R. 7200 is pending on the Senate calendar. It was reported by the Finance Committee at the end of last year, but hasn't been acted on, basically because the Senate was working on the Panama Canal Treaty. H.R. 7200 doesn't deal with welfare reform, but with foster care, subsidized adoptions, child welfare services, and supplemental security income. It also has some administrative provisions affecting AFDC. Since that bill went on the calendar Senators Baker, Danforth, and others have co-sponsored a middle-range welfare reform bill. I'd call it incremental, but that term seems to invoke considerable emotional response. In any case, instead of being an overall plan it makes a whole series of changes in the different parts of the welfare program, and it does involve several billion dollars. There has been some talk--I guess Senator Danforth himself in the hearing raised the possibility--that some or all of that bill might be offered as an amendment on the Senate floor when H.R. 7200 is considered. Even so, it will be difficult to get the House conferees to be willing to consider significant parts of welfare reform if the House hasn't acted already. Generally, the House doesn't like to go into conference on a bill and discuss a new Senate provision, especially where something very significant is involved.

Much would also depend on the Administration's position on the proposal. They might view it as harming the chances of passing their own bill. Without some strong Administration support I don't think amendments of that magnitude could pass, even in the Senate.

JAMES SHORT: Could you elaborate on your interpretation of some of the things that you have heard coming from the research reported here. I am

a little curious about how you think the Congress is going to interpret these findings.

STERN: The two major results, I think, are that when you offer a guaranteed minimum income program you can expect a significant reduction of work effort-- significant in the sense of adding significantly to the cost of the program. That will no doubt make a big impression on the Congress. Second, instead of a guaranteed minimum income program contributing to family stability, it seems to have the opposite effect. For the time being I think that Congress will value having families continue to live together.

SHORT: Do you think it is possible for the Congress to consider/make more than traditional assumptions? I refer to the assumption that family stability is the overriding value, regardless of what a marriage means, or what a family means?

STERN: Knowing that enactment of a particular program might increase family instability by 50% or 100% might well lead many members of Congress not to probe further, in my opinion.

KAY THODE: I am concerned, because my understanding from the morning workshop was that, although there had been an aggregate reduction in the neighborhood of 100 hours a year on work effort as a result of the guarantee, that no analysis has been done on whether that was a reduction in overtime work or part-time withdrawal from a part-time job. If, in fact, it stems from a reduction in overtime work a second person will fill the job, and you won't have reduce productivity. I think you are possibly basing decisions on preliminary analyses. As for family stability, the results seem to indicate that if you give families a substantially higher income guarantee, you reduce family breakup. I hope Congress is not going to assume that the guarantee necessarily produces a family breakup and reduction in work effort on regular work hours.

STERN: You make a good point. But I wouldn't have tried to intepret a day's worth of sessions if I hadn't been asked.

BETH HARRIS: Is it your general assumption then that women who are at home caring for their children aren't doing any real work?

STERN: No.

THE ADMINISTRATION'S APPROACH:
THE PREMISE OF THE BETTER JOBS AND INCOME PROPOSAL

by

Arnold Packer
Assistant Secretary for Policy, Evaluation and Research
U.S Department of Labor

First, I want to answer Michael Stern: It is not quite accurate to suggest that the Carter administration's proposal is Nixon's Family Assistance Program (FAP) warmed over. And, it's not a negative income tax (NIT), although in some ways it resembles one. This administration has a very different approach.

The biggest difference is in the treatment of jobs, this is the key, not only to providing a means for getting congressional support, but to providing income.

I think a number of important things have happened in the last decade. One is the findings of the income maintenance experiment. Another, as Senator Moynihan has said, "Is just a very different feeling among women about what work means." As Senator Moynihan pointed out, ten years ago, "If you said you were going to provide work for AFDC women, it was considered a cruel and inhuman way to approach the problem." Today, if you decided to provide a job guarantee for two-parent families and did not make that available to one-parent families, you would very soon run into a political buzz saw--and deserve it. That is a big change in a short period of time.

I think also there is a feeling that an NIT is something designed by economists and that it is based on too narrow a view of human motivation. This view assumes everyone makes all their decisions on the basis of money and money only and, furthermore, that cash from a job is the same as cash from welfare. The Nixon administration accepted this view. President Carter, who has taken a very personal role in the design of the administration's Better Jobs and Income Program (BJIP), has a different understanding of human behavior.

Under the BJIP there would be a very low guarantee for those expected to work but who did not want to do so. This decision reflects a judgement that it's not the government's responsibility to provide a guaranteed income to everybody, regardless of whether he/she discharges his/her responsibilities to his/her family and to society. The BJIP reflects a judgement that if there is someone in a family who can take care of young children and someone else who can work, one of the parents is expected to work.

A willingness to work is only the family's half of the dual responsibility shared with the government. The responsibility on the government's part is to make sure that there are jobs available. So we want to guarantee people income substantially over the poverty line--not at 65% or 75%, but at 117% of poverty, in general by providing jobs and if the jobs do not pay a sufficient amount, we will add a wage subsidy to it.

The wage guarantee itself is set low enough so that we don't think we will attract very many people from existing jobs. Now, there are lots of people who work at the minimum wage, but very few of those are family

heads. Ten percent of low-wage persons, providing 20% of all low-wage hours of work, are family heads. The remaining hours are provided by single persons, teenagers, or secondary family earners.

So I think we are looking at a proposal that was proposed by a president who wants his programs designed with a moral basis. The moral basis in the BJIP recognizes the special needs of single parents and their young children. These families are assisted by providing a guarantee that's two-thirds of the poverty line to all those who are not expected to work. This is the same group that is not expected to work right now--women with children under the age of six. It reflects a premise that has been accepted by this country for quite some time.

Heads of two-parent families and one-parent families with children over 14 are supposed to work if they want an income; and we as a government have a responsibility to provide them with jobs. This is not going to be easy: The most difficult challenge we have in this country is to find jobs for those without skills. At the moment we leave this extremely difficult task of finding such jobs to those least prepared with our complicated society.

I think that most Americans and most politicians, feel a certain sympathy for the man or woman who wants a job, and most believe we should give them one at the minimum wage or slightly above, and even to supplement that income to bring it 10 to 15% above the poverty line.

Certainly we will face problems: How do we make sure those jobs are not life-time careers? How do we ensure that people move through them--that the wage is not too high or not too low? I think we can

solve these problems.

The Baker-Bellman approach, the Ullman approach, as well as the President's approach starts out with a fundamental assumption that the society owes a job to the heads of families with children and that women with young children are owed a choice. They can either stay home and accept the welfare guarantee or if they choose they are guaranteed access to a job.

Finally, this proposal is much less expensive than an NIT and provides a much higher income for those who work. The estimates on cost savings are quite dramatic where the program has incentives to work rather than not to work.

Interestingly, when this program was designed we did not have the results we now have from SIME/DIME, but I think our strategy is supported by the research findings reported here. The minimum income should be high enough to avoid an affect on family stability. What we aspire to is to neutralize the financial incentive for family structure. A woman who is living with a man whom she finds impossible should have access to a job. On the other hand, for the great majority of poor women who wish their husbands had a decent wage, this program will provide it. And in the frequent case of a pregnant young lady who has not yet married the father of the child, the BJIP will extend a job guarantee to her, or to her spouse if they do get married.

To conclude, it's incorrect to suggest that what we are talking about here is FAP, updated. It's a completely different approach. I think it's politically saleable. I think in a short period of time, we have shifted the debate from how much cash to how many jobs, how long a waiting period would it be, and what should the wage be.

DISCUSSION

HENRY FELDER: Your remarks seem to be based on the assumption that there are jobs available for people who want to work; but with the rate of unemployment currently close to 7%, it's not clear, where all of those jobs are? There also seems to be an underlying assumption that people must be forced to work, otherwise they will sit around and draw benefits. So, according to the program that you are outlining, the work component must be a major part of any welfare reform act.

If the government institutes a program of job guarantees similar to that in the Humphrey-Hawkins proposal, what will be the approximate cost of such a program; and why should the government create jobs that the private sector doesn't seem to want?

PACKER: There is a large job component in the BJIP. As proposed, there would be 8.8 billion dollars for the creation of 1.4 million jobs. Now, the number of unemployed at last count was 6 million, of which less than half were household heads, and only a fraction of which were household heads with children. We think that 2½ million people will want these jobs during a year if the unemployment is ½% less than what it is now--5½%. It is unlikely that all of those 2½ million people would hold these jobs continuously through a year, so we think a stock of 1.4 million jobs could, in effect, provide a guarantee.

Second, why the private sector would rather not hire those people is a long and complicated story. But they will hire people if it maximizes profit--which means if they would rather hire a young person without any family responsibilities who is well trained, than an untrained person with family responsibilities. But because of the greater social costs of a lack of income for those with children and with family responsibilities, we want to provide work for them in the public sector--that is, 1.4 million jobs that will, in essence, create a job guarantee.

FELDER: What kind of jobs will these be?

PACKER: Well, we have prepared a list based on estimates of what is needed. There are jobs working in the health area; geriatric nurses; day care; security jobs; transportation; recreation; and construction, in some places.

The Labor Department has just entered into some agreements with the Department of Energy to provide solar heating for low-income families. We have also done weatherization projects, mosquito abatement, and many similar projects. We think if we combine those jobs with training they will be temporary and people will move through them to private sector jobs.

CHRISTINE SWIFT: It is my understanding that during the eight-week required job service period, benefits are reduced and there are no job-related services. I am interested to know the logic behind that.

PACKER: I think you have got to remember that there is nothing in this program that takes anything away from anybody. It has a hold-harmless principle. Two-parent families (except for 140,000 families in the AFDC-unemployed fathers program) are not covered today. If those people are out of work for 5 to 8 weeks, or 18 months, they won't get anything ever except food stamps. The BJIP may not be ideal, but in no case does it really withdraw benefits that are presently available. Further, one-parent families would start at a higher guaranteed payment, and should have no problems. For two-parent families, however, there is a lower minimum welfare benefit for a period of 5 weeks. On the other hand, I think it takes about this long to begin payments even in the present welfare program, including those few states that extend benefits to unemployed fathers.

KAY THODE: You indicated that it was impossible for low-income families to solve the problem of job availability in the private sector. But you have also given us a list of jobs to be created, but some 90% of people on assistance are women, and the jobs are not the kind that are likely to lead to private-sector jobs.

Second, why, if your major goal is to emphasize the value of earned income, do you propose eliminating the earned income tax credit? And why are you setting a time limit on these new public-sector jobs? Why don't you recognize the absence of jobs in the private sector?

PACKER: We do have a difficult balancing job. There are jobs in the private sector, they come and they go. At any point in time there are not enough jobs, but there are jobs. We have tried to create a system of incentives so that people will prefer jobs--subsidized jobs if necessary--over not working, and so that people will prefer private or unsubsidized jobs to subsidized jobs. If you establish incentives on the way up, it looks like punishment on the way down. This is unfortunate, but it is true.

Second, we think women can handle these jobs, and in fact they are doing so. Further, many of these areas are considered growth industries, for example, care of the elderly. There should be about half men and half women filling the new jobs.

ROBERT SPIEGELMAN: Have the income maintenance experiment results suggested any modifications in the proposed BJIP?

PACKER: Frankly, Bob, we have had some inkling of the results before.

SPIEGELMAN: There have been leaks?

PACKER: It has not really taken us completely by surprise. We are pleased that some of our guess-work was correct--that the jobs program is crucial, and that it's terribly important to have jobs in place prior to expanding the welfare roles. Even the Bellman-Baker proposal mandates aid to unemployed fathers, as well as jobs. Their proposal first contemplated

about 350,000 jobs, but only for two-parent families. But equity as well as politics suggests that you can't restrict these jobs this way. So they may have to double it--and create 700,000 jobs, that's still substantially less than the administration's proposal of 1.4 million.

CAROL MAHONEY: Won't these new jobs have an impact on the public-sector jobs?

PACKER: We think we have differentiated the jobs. We have at least an agreement on this in the Corman subcommittee--on that part of the bill that the subcommittee passed--and we think the AFL-CIO has acquiesced, if not given their explicit approval. We also think there are additional jobs that can be created in the public sector. The salary structure that we have today in Comprehensive Employment Training Act (CETA) is similar.

This is a tough issue--providing enough jobs and still avoiding a serious impact on wage structures. You do it by making the new jobs relatively unattractive, compared to private sector jobs, and you do it by making these new jobs temporary. We think we have worked out an acceptable compromise.

THEODORE LANE: I am concerned that the administration seems to think that it needs to keep providing incentives to push people back into the private labor market without recognizing that there aren't any jobs there for these people.

PACKER: That's not the idea. The idea is to have countercyclical jobs phased down. The house and senate subcommittees want this. People will be retrained--and there would be additional funds for training.

PHILIP ROBINS: Mike Stern appeared somewhat pessimistic about passage of welfare reform legislation for this year. I was wondering what you think the future will hold?

PACKER: I think that Mike referred to a meeting with Senator Long, Senator Moynihan, Mr. Korman, and Mr. Ullman. There was an indication of flexibility on the President's part and a willingness to have less than the whole load pass on the first effort. Whether the first increment will pass this year or next year is hard to say. But I am optimistic that something can be done if not this year, then next year. And if not the entire program proposed by the president, then something that starts the ball rolling.

BETH HARRIS: Is it the Administration's view that there is going to have to be, for a long period of time, public subsidized jobs in order to maintain employment, and that the private sector really can't maintain full employment?

PACKER: Well, yes, there will be permanent expenditures for structural-unemployment programs. But the vast majority of the jobs will be in the private sector.

SWIFT: During the first two months in BJIP a family of four would be expected to live on approximately \$190 a month. Does the Administration consider this possible?

PACKER: Usually the person who is working and is laid off will have unemployment insurance. The one-parent family with children under 14 will have more, and as I said before, the two-parent family today has only food stamps.

VEE BURKE: Could you tell us how and when you are planning to test the jobs component of the BJIP?

PACKER: Let me turn the podium over to Jodie Allen who can answer that.

JODIE ALLEN: We are currently requesting authority under the CETA Re-authorization to undertake a series of demonstration projects. Because of the magnitude of the projects, we are seeking special authorization from Congress, and such authorization does appear in the versions of the bill reported by the committees in both houses. Of course, we will have to wait and see what the appropriation committees do in financing these requests.

We would like to test the jobs component in a geographically representative set of sites scattered across the country--at least one in every region. We are at the moment trying to identify a geographically representative set of sites.

We think that the major issues to be resolved are two: First, we need to know if we have correctly estimated the demand for these jobs. It is important to remember that we are talking about a structural-unemployment program requiring both jobs and training, and a combination of both. We have already used two different simulations to estimate the number of jobs, and both come out about the same. But the real world may be different.

Second, we want to know what the characteristics of the beneficiaries will be. Again, the simulation models give us some idea, but we want to find if these characteristics differ from area to area, and how we can best match up applicant characteristics with jobs.

HENRY LAJEWSHI: Jodie, how long is this research going to take? You are just now asking for funding for it.

ALLEN: First, we are talking about a demonstration project. We will not have control groups. We will not be trying to measure very carefully responses over some surface, as in SIME/DIME. We think that we can get feedback on our two most important issues relatively quickly. We hope to select the sites some time this summer, give out planning grants to those

areas for 6- or 8-month planning period, and begin enrolling people in February or March, 1979. We would also have a period of outreach, so people know that the jobs are available. We think that by the end of the year we can get good readings on who is going to participate. During this 8-month planning period we will try to identify many more job types than we think we need--so that we might see what seems to work best.

Of course, we have all assumed an incremental approach. We can't phase in 1.4 million jobs overnight. So we want a demonstration program in 1979, more permanent, nation-wide reform, with a job component in 1980, and full implementation in 1982.

WHAT WE HAVE LEARNED
ABOUT RESEARCH AND POLICY-MAKING

by

William A. Morrill
Senior Fellow, MPR

I would like to focus on how the income maintenance experiments and the knowledge gained from them affect the real world of policy-making, specifically welfare policy. There is a wide range of opinion and only a limited amount of careful research on how, if at all, research findings enter into the process of making policy decisions. One major issue, as I see it, arises out of the expectations of sponsors and designers of research. There may still be a lingering notion, carried over from physical science research, that everyone should be persuaded by the results. It is becoming quite clear, however, that everyone is not going to be persuaded by the results, nor should they be persuaded without careful questioning first.

A second expectation is that the research will tell the policy-maker precisely what to do. This rarely happens. Some questions are and always will be beyond the reach of answers through social science research.

Another expectation--held by researchers as well as by policy-makers--has been that no relevant issues should remain unanswered or ambiguous. Clearly our struggle in the income maintenance area over a decade has demonstrated that there are always going to be results that must be interpreted and that not all questions can be answered. To ask this is to place an unreasonable load on any experiment.

Another belief that is held by some members of the research community is that an experiment is a scientific, nonpolitical event. My own observations suggest quite the contrary. The health insurance experiment, for example, was the subject of intense ideological controversy before it was even begun, and, indeed, the very commitment of resources of the magnitude required in many experiments has political repercussions.

Perhaps another false expectation to which we fell prey is that experimental results can be translated quickly into adjustments to policies. I think we need remind ourselves that most national issues of great social importance are generally in vigorous public discussion for at least a decade before something substantial occurs in the legislative process. I cite the Medicare and Medicaid legislation as interesting cases in point. When you have an issue such as welfare--which goes to the very heart of the value structure of a great many people in the society--you cannot expect overnight acceptance of a new reform. You must accept the American political process, which seems to prefer incremental action and compromises, reflecting our pluralistic nature. This requires patience. People take their time assimilating information and changing their opinions, especially when research results are either counter-intuitive or counter to the prevailing values of society. When research results sound "wrong"--that is, disconcerting or disturbing--the first step by a reluctant audience is to attack the methodology of the study. If no flaws can be found in the

methodology, the skeptic focuses on the competence of the researcher. Only after these avenues have been exhausted (and sometimes the process properly stops here) does one finally get around to considering the substance of the research findings. This lengthy process is perhaps most appropriate when particularly major programs are being considered, so that a significant portion of the population will be satisfied that the decision to change is a correct one.

This is how social science research finds its way into public policy. There is rarely a case where a single project, even one so large as this experiment, provides all the evidence one wants for a major public policy decision. My own experience suggests that the results from the experiments in the income maintenance field have affected the policy-making process within the executive branch in very substantial ways. The results from the Seattle and Denver income maintenance experiments have entered directly into the design considerations that resulted in the Carter Administration's welfare reform proposal. The experiments also affect the focus of public debate on welfare reform. By introducing solid evidence about a subject, the experiments combat arguments based solely on rhetoric and emotion.

These experiments also produce indirect effects. Congress is quite interested in sponsoring monthly reporting demonstrations--an interest that developed directly out of SIME/DIME. Another indirect and subtle product of these experiments is the development of an expert community for clearer focus--a group of people who are conversant with the research, who are appointed to jobs at the federal and state level, and who advocate whatever additional research is necessary.

This raises the issue of redundancy. Some might argue that we conducted the New Jersey and the rural income experiments and that was enough. But with an issue such as welfare reform, with its powerful emotional overtones, several experiments, with improved methodologies and attending to varying socioeconomic groups, help to accumulate the persuasive evidence needed in the political process to create change.

If the impact of experimentation and large-scale research depends on slow and incremental assimilation into the policy-making process, are the costs, particularly for the more expensive experiments, justified? To answer this question, consider the alternative avenue to reform. One option is an "instant" national experiment--Congressional enactment of a new program without any preliminary testing. Our experience with that approach has been sufficiently mixed that most would find some other, more careful, design for new or revised government programs to be the preferred course. Second, the cost of research should not be compared to the cost of less rigorous research, but with the cost of ignorance. Expenditures of \$100 million or even \$300 million for experiments in a universe where we spend around \$140 billion a year on transfer programs at the federal level alone represents a modest investment in knowledge.

In closing, the future of experimentation as a policy-making tool seems to be secure. The effort has been highly focused thus far on issues arising in programs where the government is affirmatively spending money, such as income maintenance, housing, health, or education. Perhaps ahead of us are some exciting new possibilities, particularly in the field of regulation. Experimental techniques might be applied in those areas, although that clearly will raise new ethical questions.

One of the keys to successful experimentation in the future will be a continuing effort on the methodological frontier. As we increase our understanding of what experimentation can do, what its limitations are, and how it can be improved, we will be able to produce a more effective policy-making tool.

WHAT WE HAVE LEARNED FROM THE EXPERIMENTS

by

Charles E. Metcalf
Director of Research
Mathematica Policy Research

and

David N. Kershaw
President
Mathematica Policy Research

Vast intellectual and financial resources have been applied to the conduct of the social experiments of the last ten years. In addition to social scientists already established as researchers, a whole new generation was attracted to policy research because of its interesting and important policy issues and its challenging technical questions. The lessons we have learned from doing the experiments will help us to improve both the conduct of large-scale experiments and demonstrations in the future and the day-to-day evaluations of ongoing and new programs. In addition, our sensitivity to the issues involved in effective program evaluation has been increased.

Perhaps the most important thing we have learned is that social experimentation on a large scale can be done--and that it can be done in a way consistent with the operation of a program. Policy research, which is often highly technical in its internal structure, can help answer relevant policy questions. It can be slow, and there are difficulties of language that hamper communication, but the information learned can be--and has been--conveyed to many different

decision-making and policy-making audiences. We have learned important substantive things about research, about technical and methodological issues, and about some operational issues, all of which will improve our capacity to help policy makers in the future. It is not possible here to discuss these issues in depth, but I would like to review with you briefly a few general design and implementation matters that are important to keep in mind.

SPECIFY THE RESEARCH FOCUS

We have learned how to decide what questions have to be asked. In the early days there was a tendency to use a scatter-gun approach and often to try to ask too many broad questions. What we tended to get, therefore, was a lot of information that was potentially rich but often not analyzed in detail. I think we have learned to be sensitive to that problem and have tried to focus on a few critical questions in designing our experiments.

I think we also have learned to make certain that the questions on which we focus our experiments (especially the long-term ones) are policy relevant but not overly tied to the specifics of the particular program. We must maintain a sufficiently broad policy space because the policy options when the results become available may be very different from the program options with which we started several years before. Rather than focus on the wrinkles of the specific program, therefore, especially if we are planning a large-scale effort, it is important to do a lot of thinking about what is fundamental, what is likely to remain important even if the nature of the policy discussion should change.

BE ABLE TO COMMUNICATE THE RESULTS

However complicated our analysis may be and however intricately designed the experiment is, we have learned that ultimately it must be capable of providing information simple and direct enough to be interpreted by people who are not technicians. They have to understand enough about the experiment to be able to believe the results. They are not going to take the technician's word on blind faith. And if they did, they would probably be making a mistake. So, technicians must be concerned with how they can convey their results. This obligation includes pointing out to policy makers that the issues may, in fact, be more complicated than they had earlier thought.

When to report research results is, in my view, an inherently insoluble problem. There was some feeling that there was pressure from the sponsoring agency to report the New Jersey results too early, but that the rural results came too late. There are probably some feelings now that we should be getting the SIME/DIME results sooner. What I have learned is that when you have committed yourself to a large experiment testing a controversial program and you are using a scientific, hypothesis-testing method, you pretty much have to let the chips fall where they may; and, in particular, you cannot keep a result from getting out.

MAKE THE STIMULUS STRONG ENOUGH

The experiments have taught us much regarding the specification of the treatment variations that should be built into an experiment. This is closely related to the first lesson--the importance of

restricting the number of research questions--especially if you have small samples. How strong to make the stimulus is a crucial question in social experimentation. There was some fear that the New Jersey results, for instance, could not really be trusted because the treatment was too weak to find out what we wanted to know. This fear has been allayed to some extent with the SIME/DIME experiments because they used more extensive samples over somewhat longer periods of time.

With respect to income maintenance experimentation, the particular stimulus that is the biggest unresolved issue is, in my judgment, how much the tax rate matters. Other sources of information and types of results suggest that the tax rate of a welfare program does make a difference. In New Jersey we restricted our range of tax rates to the small band that we thought to be directly policy-relevant. The unintended result was that we didn't get a strong enough variation in outcomes to get a definitive test. Thus, I think we learned from New Jersey not to have too complex or too subtly varying a set of treatments because, again, we may be drawn away from the fundamentals.

The demonstration/experiment distinction is important here. The basic distinction we usually make is that experiments have an experimental group and a control group, and assignment to each is random--decided by flipping a coin. The distinction I want to stress here, however, is somewhat different. In a demonstration you want to see how a particular program works. You set it up, you operate it, and you infer what you can from its operation--about not only its feasibility but to some extent its effects.

In an experimental framework, in contrast, the programs you implement are what we have tended to call treatments, which are viewed as vehicles for information and, in some cases, may not be the particular programs that are directly applicable in a policy framework. What ultimately comes out as welfare reform, for example, will not correspond specifically to any treatment that has appeared in any income maintenance program to date. Now, this does work against the simplicity objective because you may be experimenting with things that are not exactly what you would like to see. But in experiments there is no choice.

THERE ARE MORE PITFALLS TO SAMPLE DESIGN THAN WE ONCE THOUGHT

We have learned a lot about sample design issues since the New Jersey experiment. We have learned, for instance, that it is very important how you specify the sample frame for your experiment. A number of issues come to mind where, indeed, the experimenters made mistakes. A major one in the New Jersey experiment was sample truncation. By defining the sample to be in that range of income where people could actually receive negative income tax payments, we excluded from the sample most families with multiple earners who, should the spouse or some other earner withdraw from the labor force, would then move on to the transfer payment rolls-- a substantial portion of the population. It seemed reasonable at the time not to include those with family income high above the benefit cutoff level, but that decision made us unable to observe the labor supply responses of those who would have become eligible for payments as the result of drastic reductions in family income.

Another issue is whether you should have a dispersed national sample or concentrate your experiments in particular locations. There has continued to be a fair amount of disagreement about this. But I think that most of the people who have been involved with the experiments still believe that we were correct, given the resources at our disposal, to concentrate experiments in particular locations and then try to deal with the admittedly difficult question of how to translate this into what would happen if the program were implemented in the entire country.

Another issue of sample design has only recently come to the fore and been taken as seriously as it should be--the issue of sample size. Do you have, given the nature and complexity of the questions you're going to ask, a sample size adequate for statistical precision in measuring effects of the magnitude you expect to occur? For example, the New Jersey sample was barely large enough to measure effects for a fairly homogeneous population. As it turned out, the results tended to be quite different among the black, white, and Spanish-speaking populations. But the subsample sizes for black and Spanish-speaking populations were too small to get a full understanding of what those results were.

In the rural experiment, the sample was cut entirely too many ways. The full sample was only 800 families, split among whites and blacks, farm owners and non-farm owners. It turns out that the entrepreneurial nature of farm ownership made a big difference between farmers' response and the response of people in nonfarm jobs. The experiment was just not big enough to measure with any reliability any of the information one might have hoped to get out of that experiment.

Another crucial facet of sample design is sample allocation. Sample allocation models are, in principle, extremely important. These are expensive experiments. It is therefore very important to have a cost-efficient design and an experiment focused on the answers we would like to address. The simple lesson is that when you are doing the analysis, remember what you have designed. This is critical. A complicated (and probably in the early experiments an overly complicated) design structure produces data that are not general purpose data. They are focused in particular ways, and one must account for that in the subsequent analysis.

DON'T UNDERESTIMATE THE FIELD PROBLEMS OF EXPERIMENTATION

The experiments have taught us much about survey research among low-income populations. We have learned about constructing and administering survey instruments in situations where we have to obtain extensive and detailed information about and from relatively uneducated respondents. We have learned about the right kind of interviews to use for this kind of population. We have learned about methods of verification and quality control, and we have also learned that we had better learn something about the populations with whom we are experimenting before we jump into the field. In the New Jersey experiment, in particular, we had absolutely no conception of what was going to happen. We made the mistake of believing census data. Incomes turned out to be much higher than we had anticipated and, given our selection process, the racial composition was also very different from what we had anticipated, causing problems later in data reduction and processing.

In the data processing area, in fact, we have to learn two lessons: first, avoid an over-designed system--but second, design early. In most of the experiments we have too often done neither. We decide on research questions; we get together to develop the survey instruments; and then almost as an afterthought we realize that we have to process the data. We are finally learning that the design of the data processing and the data management system has to be done before the rest of the design of the project.

ENSURE LOCAL GOODWILL

It is important that there be good relations among the experiment or evaluation people and the program operators (particularly where program operators are run independently of the organization in charge of the evaluation), the local communities, and the program participants. Unless we are able to achieve minimal acceptance at the grass roots level of what we are trying to do and some appreciation of its importance, our experiment is simply not going to work.

RECOGNIZE THE INEVITABLE LEGAL AND ETHICAL PROBLEMS

Clearly, there are some very complicated issues here and, although we have not solved all of them, we have at least learned to grapple with them openly. One important issue centers on any experiment where some of our treatments make people worse off than they currently are. In such a case a voluntary program will not work because no one can be expected to join a program that will make him worse off, and a mandatory program cannot be instituted because it is unethical to impose an option on anyone that will make that person worse off.

Let me offer an example of a hypothetical experiment that many people would reject on these ethical grounds. For a period in the mid-1970s the unemployment compensation system in the United States was such that, with various federal supplements, in most states you could collect as many as 65 weeks of unemployment compensation of various forms. There was considerable feeling in both the research and policy communities that the availability of unemployment compensation for this long could be expected to pose severe work disincentives. As it turned out, the policy response was that the perceived need for long-term benefits tended to wither as unemployment rates fell; so the program, which was temporary in nature, was allowed to lapse. But let us suppose that it was a permanent program and the real policy question was whether too much unemployment compensation was on the books. Should we, as ex-Governor Burns of the Federal Reserve Board suggested, cut unemployment insurance back to 13 weeks? Any experimental option for answering that question would have involved making some people worse off; and unless you took a rather harsh ethical view of the matter, that likelihood would have precluded this particular aspect of unemployment compensation as an area of experimentation.

Closely related to the no-worse-off principle is the question of random denial of access--whether it is ethical to preclude someone's access to a program solely by a coin flip. Let us suppose that only one-half of the available population can be covered with the available funding. The obvious decision for an experimenter who wanted to evaluate such a program is to say: Well, if only half of them can get it anyway, why not do it randomly and flip a coin? Some people,

however, have very strong feelings that randomization is no better than any other justification for a situation where half the group gets something and the other half does not.

Finally, we have become very sensitive to the rights of privacy of participants and to protecting the confidentiality of the information they provide. This protection is sometimes difficult to maintain because some of the activities we supposedly have knowledge of may well be illegal. This represents an ethical and legal conflict: because of our research objectives, we of course are interested in that information, and certainly the courts--if they know that we have that information--are going to be interested in getting it.

ACKNOWLEDGE THE LIMITATIONS OF EXPERIMENTATION AS A RESEARCH TOOL

Now for the final lesson. We have learned that large-scale, controlled experiments are not the answer to all of our questions. We began social experimentation because the defects of nonexperimental methods of analysis were such that we could not be provided with definitive answers. In fact, common textbook criticism of all social science research was our inability to do controlled experiments, and the attempt to do something about that was in large measure responsible for the social experiments. Unfortunately, they have not been a research panacea. As a research method, controlled experiments are extremely expensive, and not all policy questions are worth the cost. Second, every time we run one of these experiments we see that the results are not definitive.

The important point here is that experiments are not imperfect just because of mistakes, the fallibility of researchers, and the lack of money. They are imperfect because they possess inherent inadequacies.

We are never measuring precisely what we would like to measure. People leave the experiment, and they leave in varying ways. And, as to what is called the control group, we don't really have good control over what they do or what environment they face. Also, an experiment of limited duration does not necessarily allow us to see what would happen in a continuing program. And an experiment in a particular location, such as in Seattle, does not ensure that we will know what it would be like if generalized to the rest of the United States. Finally, the program we have experimented with may not be the one we turn out to be interested in. And so, what we have, on the one hand, is a desire for a simple experiment whose results can be directly observed and applied to policy but, on the other, inevitable complications creeping in because we must adjust for the reality that the proper test has not been performed and, quite often, could not have been. It's not quite the right program; it's not quite the right environment.

Ideally, what we need is a large random sample of U.S. citizens with which to conduct full-scale saturated programs. Clearly we cannot do that. As a result, we have results that must be interpreted and then communicated. And, in order to interpret our results we must use the very same imperfect, nonexperimental techniques that controlled experimentation was set up to avoid. So, rather than a clear choice between conjectural, nonexperimental analysis versus controlled experiments that will answer all of our questions, the controlled experiments have turned out to be simply one technique that provides some answers. But from there one must still proceed with standard

imperfect analysis. Thus, the experiments, while a major innovation in social research, are not the definitive vehicles that our public relations people, namely ourselves, have made them out to be. And they must be used with the same restraint and with the same humility about their imperfections as are other research methods.

RESEARCH ISSUES
IN SIME/DIME: AN OVERVIEW

by

Richard W. West
Economist
SRI International

A basic goal of many social science research efforts is to determine how public policy affects human behavior. The Seattle and Denver Income Maintenance Experiments (SIME/DIME) share this goal. SIME/DIME is an ambitious project that has produced an enormous amount of data on income maintenance policies and a wide variety of human behavior, including not only work effort and marital instability--the focus of much investigation--but also migration, human capital formation, fertility, expenditure patterns, occupation changes, wage rate changes, job search behavior, child care usage, the performance of school children, and so on.

Some of the papers that follow will analyze the effects of existing income maintenance policies, such as Aid to Families With Dependent Children (AFDC) and unemployment insurance, using only the SIME/DIME control group.

There are problems with this method. Those receiving income maintenance and those who do not will undoubtedly differ in important respects. For example, people receiving unemployment insurance are likely to have different educational levels, job histories, and so forth, compared to those who do not. Point-in-time differences between the two groups could be attributed to these other variables rather than to the presence of the program, and this prevents conclusions as to cause and effect. This problem

can be partially overcome by the skillful use of regression analysis, where the researcher compares strata of people who are as alike as possible in all the characteristics the researcher feels will be important. Regression analysis only permits conclusions as to relationships, however; and the researcher still cannot make conclusions about cause and effect. Examining the same groups at another point in time, possible with the SIME/DIME control group, as longitudinal analytical files become ready, permits analysis of causal factors, but there remains a possibility that unidentified socio-economic factors continue to confound the analysis.

Research using the experimental data avoids these problems, because the groups to be compared are both selected randomly. One portion of the sample receives no special treatment--the controls. The remainder of the sample--the experimentals--receive treatment. In SIME/DIME, experimentals were divided into subgroups and received different treatments. Where assignment to treatment or control status is completely random, the different groups can be assumed to be identical, on the average. Thus, differences between groups can be presumed to be due to the experimental treatment.

To illustrate the importance of a control group, take as an example the behavior of young males who were not heads of the families, as observed in the SIME/DIME samples. Those receiving treatment worked an average of 6.22 hours before the experiment. During the third year of the experiment they worked 14.78 hours per week. The increase does not tell us much about the effect of the experiment, because the elapsed time could allow other variables, such as the business cycle, to affect the result. Here, however, attribution

of the change to the experiment is wrong for a much more obvious reason. These young men are three years older in the third year of the experiment and young persons are likely to work more as they get older. The extent of the error might be detected by looking at the control group, who increase their work effort by 12.33 hours per week over the same period. Thus, the controls enable the researcher to correct for environmental or other changes that have an effect, independent of the experiment, as long as these changes have the same impact on both controls and experimentals.

The example also illustrates the way in which the researcher can develop an estimator of the effect of the experiment. The simplest estimator, or measure of predicted effect, is the differences between experimentals and controls. In the example, this difference was -4.9 hours of work in the third year of the experiment. This becomes the basis of estimates of the effect of financial treatment on hours of work.

Generally the researcher develops a model to justify the use of a particular estimator. The model used in SIME/DIME assumes that, within each strata, hours of work for a control equals mean hours of work for controls in the absence of treatment plus a random error term (a measurement of all the idiosyncratic factors that cause hours of work to vary among individuals, which should average zero). For an experimental, hours of work equals the mean hours of work in the absence of treatment, plus the error term, plus the mean effect of the experiment.

This estimator relies on an implicit assumption that, apart from the effects of the experiment, there is no reason to expect any difference between the experimentals and the controls. Of course, one could be unlucky

and generate a random assignment with large differences between experimentals and controls, but such an occurrence is very unlikely with large samples.

SIME/DIME does not present quite so simple a situation, however, as families were not assigned to treatments in a purely random fashion. As described in the Marshall/Christopherson paper, the probability of being assigned to any particular treatment depended on: (1) an estimate of the family's normal income level (the income expected in the absence of an experiment); (2) the race of the family head; and (3) whether the family was headed by a husband and wife or a single person. While assignment to treatment was random within each income/race/structure category, or cells, different cells had different assignment probabilities.

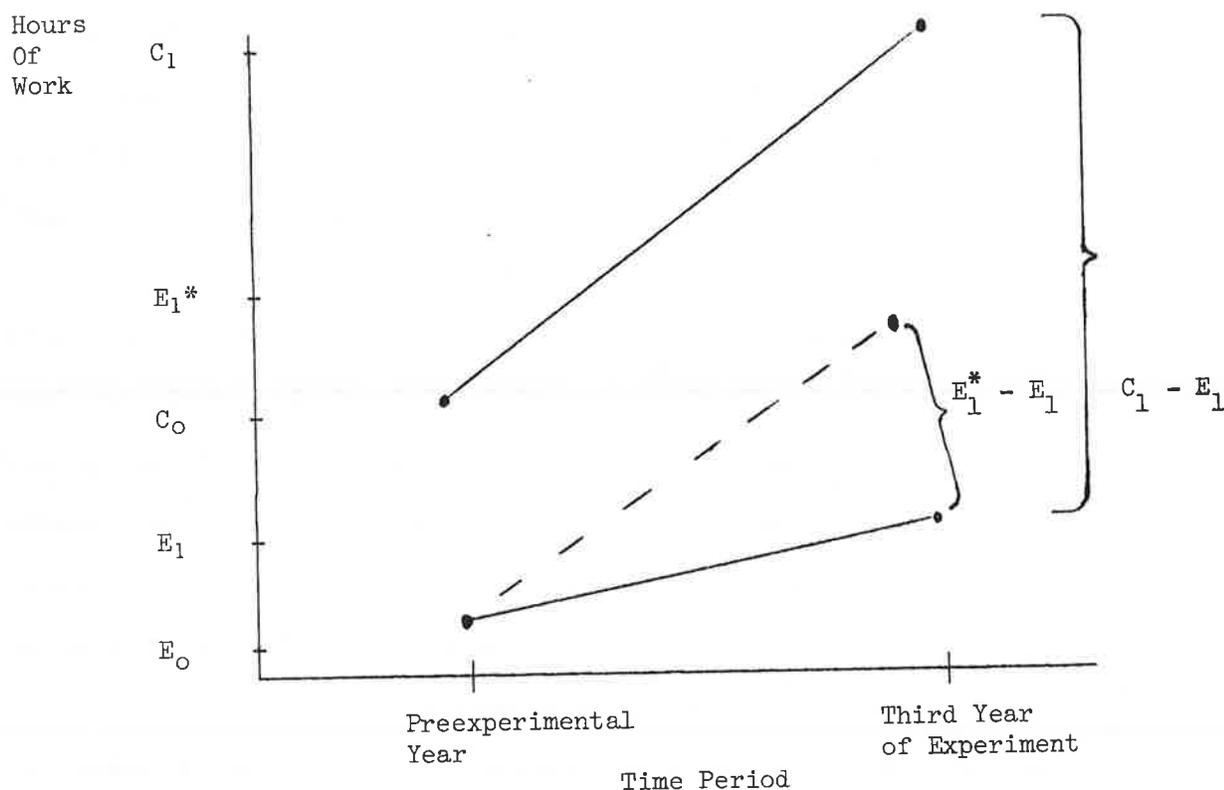
Because we used this stratified random assignment, we cannot expect controls and experimentals to be the same on average. Returning to our example of young male nonheads, controls worked an average of 7.35 hours per week in the pre-experimental period. This is 1.23 hours more than the mean for experimentals. If this difference was due to factors other than the experimental variables, the estimator based on a simple difference in means overstates the disincentive effect of the financial treatments by 1.23 hours per week.

A more precise estimator would be based on the difference in mean changes exhibited by the experimentals and the controls. This estimator accounts for a constant assignment-induced difference in means. Figure 1 demonstrates how we calculate a new estimator of -3.76 hours per week.

However, it is unlikely that the non-experimental difference between experimentals and controls would remain constant over time.

FIGURE 1

ILLUSTRATION OF THE DIFFERENCE IN MEAN CHANGES ESTIMATES



Explanation: In the pre-experimental period, controls work C_0 hours and experimentals work E_0 hours. The difference is $C_0 - E_0$. In the experimental period controls work C_1 hours. Under the assumption that the control-experimental difference would have remained constant, in the absence of the experiment, experimentals would work E_1^* hours during the experimental period if they were not subject to the experimental treatments. E_1^* is equal to $E_0 + C_1 - C_0$. In fact, they work E_1 hours per week and consequently the effect of the experiment is given by $E_1 - E_1^*$. This estimate is precisely equal to the difference in the changes of hours of work or $E_1 - E_0 - C_1 + C_0$. This differs from the difference-in-means-estimator, which equals $E_1 - C_1$, by $C_0 - E_0$. In the textual example of young male nonheads, where mean change is +8.56 hours of work for experimentals and +12.33 for controls, the difference in change is -3.76 hours. This estimator differs from our original estimator of -4.90 hours by exactly the preexperimental difference of 1.23 hours per week.

Each person's hours of work can be presumed to have a permanent component and a temporary, or transitory, component. Suppose the control group happens to have both high permanent hours and high transitory hours in the preexperimental period. The difference in the permanent components between controls and experimentals would persist over time, but the difference between the transitory components will dissipate. Consequently, the difference in mean hours worked between the two will tend to narrow over time. The difference will remain constant only if the preexperimental difference is due solely to differences in permanent labor supply.

Because assignment is random within each assignment stratum, we assume the mean effect due to the temporary component is equal to both experimentals and controls, within each assignment stratum. We assume that each person's labor supply in the experimental period is equal to the mean labor supply for the person's stratum, plus the mean treatment effect (which is zero for controls), plus an error term. The experimental effect in this model can then be estimated using regression analysis.

In our example of young male nonheads, this regression estimator is -4.10 hours. This is somewhat larger than the estimator based on difference in the mean changes (-3.76). This is what we would expect if the experimental-control difference induced by the assignment process narrows over time. As that factor grows smaller, a larger share of the actual difference (4.9) can be attributed to the experimental effect.

Most of the analyses of SIME/DIME experimental data have been based on modifications of this estimator. The modifications will be discussed in

detail in some of the other papers in this collection, but I would like to give a brief overview of some of the reasons for modification.

First, there is no reason to expect that a person with one normal income level will have the same response to a financial treatment as a person with another normal income level. Indeed, it can be expected that higher income persons will have a lower response to any given NIT treatment, simply because they would have to give up higher income sources to become eligible.

However, the estimators described above yield estimates of the average effect of the treatment on the experimental sample, and the sample does not resemble any general population. The estimator cannot be applied to the whole world. In principle this problem can be dealt with by using separate estimators for each strata. Unfortunately, the experimental samples are not large enough to permit precise estimates using this procedure. Consequently, the researcher is forced to narrow the number of factors to be examined, that is, to constrain the way the responses vary over assignment strata. These constraints may be based on a behavioral theory, or be ad hoc. Of course, if these constraints are wrong, the researcher may make some erroneous conclusions.

Second, we must analyze several experimental treatments, each of which can have a different effect on the outcome variable. In principle, this problem can be dealt with by including a dummy variable for each treatment in the equation being estimated (comparing each group with the aggregate data for every other group in the sample). However, in SIME/DIME there are eleven different financial treatments, three different manpower treatments, and two

different experimental durations, making sixty-six possible groupings for study (not counting controls). The sample within each unique treatment group is not large enough to obtain precise estimates of the main effects of each of these treatments, let alone account for all the possible interactions among them. Thus, the researcher must impose constraints.

The need to constrain the response naturally leads the researcher to make the response a function of the parameters of the treatments. In the case of the financial treatments these parameters are the support level, the initial tax rate and the rate of decline of the initial tax rate. Having constrained the response to be functions of these parameters the researcher is able to combine subgroups experiencing a particular parameter for analysis.

In our research on the effects of the experiment on the labor supply of husbands, women sharing head of household status with a husband, and single female heads of households, we have devised a model that we believe deals with these problems in a reasonable way. In this model we have represented the response for persons initially below the tax breakeven level as a linear function of the changes in disposable income and the after-tax wage rate induced by the financial treatment. These variables are correlated with normal income because they depend on preexperimental income and after-tax wage rates. It is hypothesized that the variation in response over normal income level is explained by the variations over normal income level of the changes in disposable income and the after-tax wage rate. The response also varies by financial treatment because the experimental support level affects the change in disposable income and the experimental tax rate affects both

the change in disposable income and the change in the net wage rate.

Many of the papers below use this model in analyzing the effects of treatment in SIME/DIME. Further, as described in greater detail in the paper by Beebout and Maxfield, it is also used to generate nationwide estimates of labor supply effects and costs of alternative NIT programs.

SUMMARY OF THE RESEARCH FINDINGS

by

Robert G. Spiegelman
SIME/DIME Project Director
SRI International

This paper summarizes the research results from the Seattle and Denver income maintenance experiments (SIME/DIME). This is not a final summing up of SIME/DIME, because these conference papers do not represent an end point, but a milestone. Although the operational phase of the experiments is drawing to a close, all the data have not been assimilated and the evaluations are far from complete. Nonetheless, we have learned enough to make it worth reporting now.

These interim papers are particularly successful in providing answers to the crucial question of how welfare programs with different support levels and tax rates will affect work effort. We can now tell Congress, with some reasonable degree of reliability, how the costs of a \$3000, a \$4000, or a \$5000 per year support program will differ, or what will happen to real national income if the tax rate is raised from 50% to 70%.

For example, the paper by Maxfield and Robins describes the characteristics of a program at 75% of the poverty level and a 50% grant reduction rate--which is quite similar to the Carter Administration's tier 1 cash grant program. This program (without state supplementation) would provide benefits to 10.6 million families at a total cost, net of the Aid to Families with Dependent Children (AFDC) and Food Stamp programs, of \$8 billion

in 1974 prices. We can also estimate that 30% of these costs are the result of reduction in hours of work due to the program. If the tax rate is raised to 70%, the total cost of the program drops to \$2.2 billion in 1974 prices, 60% of which is due to reduction in hours of work, but only 5.3 million families benefit from the program. Thus we can present to decision-makers the various trade-offs among program cost, number of beneficiaries, and work effort effects. In general, our results show that a moderate-size program will not cause massive reductions in work effort, but that work effort effects are a substantial portion of total program cost.

Looking at the more detailed results, the research to date indicates that in a two-parent family, husbands and wives reduce hours of work by about the same amount. However, as a proportion of total hours worked, the reduction is approximately three times larger for wives. Our results further show that reduction of work effort does not usually take the form of reducing hours of work to zero.

In the paper by Robins and West, we trace the adjustment process over time and find that work effort response took from five to ten quarters after the start of the experiment to achieve equilibrium. This paper has presented evidence that the three year experiment may have been too short to capture the effects of a permanent program. For husbands, the reduction in hours of work on the five year program was considerably larger than on the three-year program, and the differences were statistically significant. The reduction for wives was also larger for the five year program, but the differences were not statistically significant. There were no duration effects for female heads. If further modeling substantiates these

preliminary findings that the five year effects are significantly larger than the three year effects, then the work effort results used in simulations, at least for husbands, will have to be revised upward in future presentations. We are now attempting to refine our estimates of work effort response in order to distinguish transitory from permanent response, and also to determine if there are response differences by race and site.

As you know, in addition to the NIT (negative income tax) component, the experiment has a substantial manpower program which provides intensive job and career counseling and education or training subsidies. This component was designed in an effort to offset the expected reduction of work effort due to an NIT. We do not as yet have evidence as to the long-range impacts of this program on labor market behavior. To date, we have investigated only the immediate impact of the program in terms of its utilization. Use of the manpower program by members of families eligible for the program is optional. The experiment operates essentially like a voucher system in that an individual was free to select a course of study, restricted only by the boundaries of the experimental objectives; that is, a subsidy of 50% or 100%, depending on experimental assignment, was available provided that the training or education course enhances the individual's capabilities to perform in the labor market, and provided that it be the least expensive training available, given the individual's goals.

Hall's paper does indicate that the manpower program was used by a substantial fraction of the eligible population: 6350 persons were eligible for counseling, and 4367 of those were also eligible for training subsidies.

Of those eligible for counseling, 30% took advantage of the option when first offered, and an additional 15% took counseling later in the experiment. Most of those who took the option later in the program had been ineligible at the beginning; e.g., under the minimum age limit of 16. Of the 1798 individuals who took counseling when initially offered, 1284 wrote a plan of action to carry out some labor market goal, three-quarters of these plans were to take a training or education program. About one-third of those who initially started counseling entered a training or education program, including for many, pursuit of a general education degree as an option.

There is little doubt that the training subsidies increased substantially the likelihood that an individual would take training. Moreover, the 100% subsidy treatment was more effective than the 50% subsidy, indicating high sensitivity to the cost of training on the part of the client population. The models reported by Arden Hall show that for heads of families above the age of 25, a full subsidy increased quarters of schooling attended during the first two years of the program by 50 to 300% for heads of families and spouses. These large proportionate increases are particularly striking when it is realized that the amount of the direct subsidy was not large. Those on full subsidy received benefits averaging \$956 over three years and those on half subsidy received an average of \$382. In research now under way, we will measure the impact of the additional education or training on such labor market variables as wage rate, proportion of time unemployed, and job satisfaction.

In addition to the experiment's focus on the effects on work habits, there was a major search for "side effects" of an NIT. In fact, the most

publicized and controversial results from SIME/DIME have been those on marital status. Concern has been expressed by members of the Senate Finance Committee considering welfare reform that our findings on marital stability refute the Administration's contention that the Program for Better Jobs and Income (PBJI) helps to stabilize families. As described in the paper by Groeneveld, it is difficult to reconcile our findings with the view that an NIT-like program will increase the stability of two-parent families. The experimental results clearly show that the immediate impact of such a program would be to increase substantially the rate of marital dissolution. In the experiment, these findings were mitigated only at very high support levels (about \$8500 in current dollars).

These findings ran counter to our expectations, which were guided by intuition and accepted dogma, i.e., that the existing AFDC system broke up families. The research findings have dispelled the notion that an NIT will help stabilize families and have helped dispel the notion that the existing AFDC system necessarily breaks up families. The research findings suggest a new set of hypotheses that indicate that an NIT simultaneously acts to stabilize families by providing income support to the existing family, and destabilizes them by providing support to the wife and her children outside of the marriage. Furthermore, the experimental findings suggest that the NIT is considered a better and more reliable source of support to a wife considering divorce than the existing welfare system. In essence, a dollar of potential support from the existing system is discounted relative to a dollar from the NIT.

Our results in this regard cannot be automatically transferred to the Administration's welfare program. First, we do not know as yet whether the marital dissolution rates will persist in the long run if such a program is inaugurated. However, even if it does not persist, the short term effects are so strong as to result in an increase in the proportion of families with a single parent for some time, and this will have major, prolonged cost implications for a new program. Secondly, there are some important differences between the NIT under SIME/DIME, and the Administration's PBJI. Some of these differences are likely to mitigate the impact on marriage while others exacerbate it. To conjecture informally, the jobs component for male heads of two-parent families will probably help stabilize marriages. On the other hand, a female spouse generally must leave the marriage in order to acquire one of the public service jobs and this is likely to increase separation. The existence of the lower tier support for job-eligible families coupled with the work requirement, implies a weaker form of financial support for the two-parent family than is the case under SIME/DIME, and this is likely to be destabilizing. On balance, however, we cannot say with certainty whether the PBJI will have more or less effect on marriage than SIME/DIME.

I believe too much has been made of this issue anyway. The following newspaper editorial provides a sensible perspective on this matter.

"What the preliminary results clearly indicate, however, is that while the existing AFDC program has discriminated against families, the use of welfare to encourage family stability is considerably more difficult than the administration first imagined. They are a reminder that federal intervention into families--either to encourage or discourage

conventional domestic arrangements--is a risky and uncertain exercise of federal authority. Welfare programs should be used precisely for the purposes they were designed for--helping to support people who have no other means of adequate support."

Sacramento Bee, 3/8/78

There are many other behavioral responses to the experiment of interest besides work effort, training decisions, and marital status. Only a few have been analyzed to date. I will take this opportunity to capsule some of these: Thoits has studied the impact of the experiment on psychological distress. She has found a somewhat erratic pattern of responses that generally indicates a tendency for the experiment to increase psychological distress. Scores on an index of distress for white husbands in Denver, black husbands in Seattle, wives in Denver, and single black female heads in Denver and Seattle were about 10% higher for treatment than for control families. These effects did not tend to dissipate in the second year of the experiment, which led Thoits to conclude that the experiment increased distress primarily by increasing the number of life events to which the families on the experiment were exposed.

A response of potential importance in designing optimal welfare programs has been the decrease on the holding (or reported value) of assets, such as homes, which are subject to tax under the SIME/DIME payment system. SIME/DIME taxed an imputed flow of income from assets. Another experimental effect is the apparent tendency for experimental families to increase use of private sector housing relative to public sector housing. In a study by Avrin, it was found that the financial treatments reduced the likelihood that a family would remain in subsidized housing

(if initially in such housing), or would enter subsidized housing. She found that this effect is concentrated in small families, probably reflecting the tendency for subsidy values of public housing to increase with family size. This effect might translate into a permanent reduction in the demand for subsidized housing if an NIT is inaugurated.

In a study by Keeley on the impacts of the experiment on geographic mobility, it was found that on the average, being on a financial treatment caused two-parent white families in both Seattle and Denver to increase by about 50% their rate of out-migration.

Overall, SIME/DIME has presented the research community and those charged with the responsibility for setting welfare policy with a perhaps never-to-be-repeated opportunity to examine, to record, and to take into account in setting policy, the needs and reactions of the people at risk in the welfare programs. We who have been responsible for the design, implementation and evaluation of this experiment hope that we have been worthy of the trust placed in us in the conduct of this study.

DESIGNING INCOME MAINTENANCE PROGRAMS:
EVIDENCE, PROBLEMS, AND SUGGESTIONS
FOR FURTHER RESEARCH

by

Kenneth C. Kehrer
Director, Survey Division
Mathematica Policy Research

The current debate over welfare reform proposals recalls earlier attempts to reform welfare. However, one aspect of the debate that is substantially different from the welfare reform discussions during the Nixon Administration is that significant findings are now available from four experiments that tested simple negative income tax (NIT) programs with different population groups in various parts of the country. But, across all experiments, there was considerable variation in the generosity of the plans tested. Thus reviewing the findings from all the experiments might provide some insight into the consequences of introducing alternative income maintenance programs. The purpose of this paper is to summarize the findings from these income maintenance experiments, noting whether the Seattle-Denver income maintenance experiments (SIME/DIME) results support or contradict these findings.

THE DESIGN AND ADMINISTRATION OF
THE INCOME MAINTENANCE EXPERIMENTS

The initial experiment, the New Jersey graduated work incentive experiment, was conducted from 1968 to 1972 in four urban areas (Trenton, Paterson-Passaic, Jersey City, and Scranton). The sample, consisting of white,

black, and Spanish-surname families, was composed principally of intact families containing prime-age males. The rural income maintenance experiment was conducted in selected counties in Iowa and North Carolina between 1969 and 1973. The North Carolina sample of the rural experiment contained both blacks and whites, and the Iowa sample was all white. While it focused on intact families, the rural experiment also enrolled some female-headed and aged families. The Gary experiment, which began about the same time as the Seattle experiment in 1970, complemented these earlier experiments by testing the impact of selected NIT plans on a black urban sample containing a high concentration of families with female heads.

All of these experiments tested a simple form of NIT, with payments depending on only family income and family size. However, each experiment was designed to test several plans of varying generosity. The NIT plans tested in the four experiments are presented in Table 1. The average tax or benefit-reduction rate in the experiments was about 50%, and the typical guarantee or support level was slightly less than 100% of the poverty line. The New Jersey and rural experiments were relatively less generous than the other experiments, while SIME/DIME provided the most generous plans. Taken together, the experiments provide a range of generosity that encompasses the policy options for designing income maintenance programs.

On the other hand, no one experiment encompasses the entire range of plans tested. Unfortunately, SIME/DIME, although the largest and perhaps the most useful of the experiments, does not include any plans that have support levels even somewhat similar to the plans currently being considered

TABLE 1

BENEFIT-REDUCTION RATES AND SUPPORT LEVELS
TESTED IN THE INCOME MAINTENANCE EXPERIMENTS

Support Levels (Percent of Poverty Line)	Benefit-Reduction Rates (Percent)					
	30	40	50	60	70	80
50	N.J.		N.J. Rural			
75	N.J. Rural	Gary	N.J. Rural	Gary	N.J. Rural	
90			SIME/DIME		SIME/DIME*	SIME/DIME**
100		Gary	N.J. Rural	Gary	N.J.	
120			SIME/DIME		SIME/DIME*	SIME/DIME**
125			N.J.			
135			SIME/DIME		SIME/DIME	SIME/DIME**

* The benefit-reduction rate declines by 2.5 percent for each \$1,000 of income received by those families assigned to a declining tax plan.

** All of these plans had a declining tax.

in Congress. The least generous support level is approximately a third greater than the support level in the Administration's plan. If the response to alternative benefit levels is discontinuous, as many suspect, responses to generous plans may provide misleading information about responses to less generous ones. Thus, considerable caution should be exercised in interpreting the results to estimates of the potential response to the Administration's welfare reform plan based on SIME/DIME data, as is done in some of the papers below.

The choice of experimental benefit-reduction rates and support levels partially reflected the generosity of the Aid to Families With Dependent Children (AFDC) program in each of the states where the experiments were conducted. In designing the experiments, practical considerations required that the experimental payments dominate the existing AFDC payments, to assure that families would be better off on the experiment, and therefore willing to participate.

There were other important differences between SIME/DIME and the earlier experiments. Families in the New Jersey, rural and Gary experiments were enrolled for only three years, while SIME/DIME enrolled some families for five years. The New Jersey and rural experiments provided cash benefits to the families in experimental treatments, but provided no special social services. On the other hand, the Gary experiment included an experimental child care subsidy program and an experimental component that provided access to social services. Manpower services were unique to SIME/DIME. More families participated in the Seattle-Denver experiment than all of the other experiments combined.

In some other ways the administration of the experiments was quite similar. Eligible families were randomly assigned to experimental or control groups, and interviewed periodically--quarterly in the New Jersey and rural experiments and three times a year in Gary and Seattle-Denver. All of the experiments utilized a system of monthly retrospective self-reporting of income, in contrast with the existing AFDC program in most states, where caseworkers ask families to project their income over the next six months. The experiments each had an accounting period longer than a month to take into account high earnings in the recent past in computing benefits. There was considerable variation across experiments in the details of the length of the accountable period and how past income was carried over into the payment calculation.

Indeed, there was considerable variation in the detailed rules of operation across all of the experiments, and even in the operation of an experiment over time. Because these differences in rules may have created inconsistent incentives, comparing the findings from the experiments should proceed with caution. With this caveat, some observations on the main findings of the experiments are presented.

WORK DISINCENTIVE EFFECTS

The total cost of an NIT depends largely on its disincentive effect on work effort. Any income support program is expected to have some disincentive effect on work effort because giving an individual income support payments takes away part of the reason to work. The challenge in designing an income support plan is to develop a program that provides adequate benefits

with a minimum of work disincentive.

The findings from the income maintenance experiments indicate that the experiments did have a disincentive effect on the work effort of household heads, but work effort declined by less than was generally expected. (See Table 2.) Husbands in intact families reduced their total hours worked by an average of about 6%, and their wives reduced their hours of work by around 20%. Female heads of households without an adult male present (who were generally on AFDC prior to the experiments) reduced their hours of work between 5 and 12%. However, because both female heads receiving AFDC and

TABLE 2
EFFECTS ON HOURS WORKED OF THE AVERAGE NIT PLAN TESTED
IN THE FOUR INCOME MAINTENANCE EXPERIMENTS
(Experimental Response as Percent of Hours Worked by Controls)

Experiment	Sample		
	Husbands	Wives	Female Heads
New Jersey	-6	-31	
Rural	-1	-27	
Gary	-7	-17	-5
Seattle-Denver	-6	-17	-12

wives worked few hours prior to the experiment--about six hours a week, averaged over all women including those that were not employed--their reductions in work effort had only a small impact on total family work effort, earnings, benefits, and program cost.

The income maintenance experiments did not have a jobs component or a work requirement, so the effects on work effort could be substantially different under the Administration's proposed Better Jobs and Income Program (BJIP). Provision of a public service job could ameliorate disincentive effects among those who withdraw from the labor force after looking for a job for a considerable time. Indeed, the work effort reductions in Gary were concentrated among those who might be expected to benefit the most from a jobs program--young and aged males and males who have been looking for a job unsuccessfully for some time.

Somewhat surprisingly, the range of the work effort response estimates from different populated groups in different parts of the country (with somewhat differing tax and transfer programs) is quite narrow, between 1 and 7% of total hours worked for husbands, 17 to 31% for their wives, and 5 to 12% for female heads of families. The experiments thus have been able to considerably narrow the range of uncertainty about the work effort effects of NIT plans with an average tax rate of 50% and an average guarantee level of about 100% of the poverty level, relative to the range of non-experimental estimates (0 to 60% for males, 13 to 100% for females).

However, the similarity of estimates of the work effort response across experiments might be illusory, for the estimates reported in Table 2 were obtained with somewhat different methodologies. For example, the existence of alternative welfare programs such as AFDC and Food Stamps was not always controlled for in the same way, nor were differences in labor market conditions. During the next year, research at Mathematica Policy Research

(MPR) and Stanford Research Institute (SRI) will provide estimates of the work effort responses across the experiments which are computed on the same basis.

FEASIBILITY AND COSTS OF MONTHLY REPORTING

In all the income maintenance experiments, each family was sent an Income Report Form along with its income support check on the first day of the month. The family reported income for the previous month and any changes in family composition. The completed forms were supposed to be returned within a week. Over 90% of the families filed their reports on time. About 2 to 3% of the families tended to file late reports, while between 2 to 7% of the families would fail to submit any report at all.

About 90% of the forms received were able to be processed without need to contact the family to resolve discrepancies or questions.

The payments information was processed in time to be used for the income support payments for the following month. Thus, July payments were based on income earned in May and reported and processed during June. The evidence from the Colorado monthly income reporting experiment now in progress indicates that the payments cycle can be shorter and hence even more responsive to changing family circumstances. The Colorado experiment, which is discussed in one of the papers below, has provided strong evidence that this self-reporting system is much more accurate than the current AFDC system, and promotes considerable costs savings.

This self-reporting system did not appear to burden recipients. In fact, participants in the income maintenance experiments seemed to like the

experimental program better than AFDC. In Gary, families were asked about their perceptions of the program. Over 60% described their participation as having been "very worthwhile" and 81% indicated that the program had helped them. The findings from Gary also suggest that those families that had been AFDC recipients viewed the experimental program more favorably than AFDC. For example, almost 48% of the families thought that AFDC rules were too difficult to understand, while only 16% held the same view about the Gary program. Similarly, 85% believed that AFDC rules were too intrusive; 48% thought so about the Gary program. Finally, almost 75% felt that AFDC rules were not enforced equitably, while just 28% felt this way about the Gary program.

The administrative system not only pleased participants more, but by simplifying program rules and streamlining administrative procedures, the income maintenance experiments were able to effect substantial savings in administrative costs. Administrative costs in the New Jersey experiment were about one-third of the cost of payments administration under the AFDC program. Benefits were administered at a cost of about \$100 a year per case. Costs of payments administration in the Gary experiment were quite similar. The highly computerized system being tested in Colorado is actually somewhat more expensive than AFDC to administer, but these additional costs are more than offset by the large savings derived from the more accurate reporting, which greatly reduces income levels reported, on average.

HOW THE FAMILIES USED THEIR PAYMENTS

The evidence from the income maintenance experiments suggests that families would improve their well-being and use the payments in ways that

will make them less dependent in the long run. For example, families used the payments to reduce debt and increase their savings; they also bought home appliances and additional food, clothing, medicine, and automobile repairs and insurance, but did not buy automobiles. The experimental payments enabled families in Gary to reduce their use of social services, and move out of public housing. Families in the New Jersey experiment used the payments to move to better neighborhoods. Young families in Gary used the payments (which were not restricted to a particular state as in AFDC) to move to areas with better job opportunities.

The experiments also appear to have had important effects on the well-being and long-run earnings capacity of children and teenagers. Teenagers in Gary and the rural experiments tended to quit jobs and stay in school. Evidence from the rural and Gary experiments suggests that young children performed better in school. In the rural experiment, the only experiment that measured nutrition, there was a marked improvement in family diet in North Carolina. The Gary experiment investigated the effect of the payments on the health of children at birth, and found significant gains among women who were at risk of giving birth to underweight babies. As reported in this conference, this study is now being extended to the Seattle-Denver experiment.

These findings are not always consistent across experiments, nor do they occur in every population group studied. Nonetheless, taken altogether, the cumulative findings from the experiments are suggestive of the broad benefits that this kind of welfare reform would bring to poor families and to society at large.

On the other hand, the findings on marital stability have provoked considerable concern. On the one hand, the SIME/DIME analysis reports large effects on marital breakup and the New Jersey experiment revealed some possibility that the experimental payments encouraged breakup. These should be contrasted, on the other hand, with the lack of effect in the rural and Gary experiments. As in the case of the estimates on the work effort response, the estimates of the experimental effect on marital breakup across the experiments have not been calculated on the same basis. The Department of Health, Education and Welfare (DHEW) is funding additional analysis at MPR and SRI in the second year of the final SIME/DIME research contract to re-estimate these responses using common methodologies and similar data sets.

The differences in rules among the experiments are especially important in understanding and interpreting differences in the marital breakup response across the experiments. For example, in SIME/DIME, if the husband and wife split up, both individuals (and their dependents) were eligible for continued payments. In Gary, until mid-experiment, a husband who left his wife and children was not eligible for payments. Moreover, in Seattle-Denver a husband who split was eligible for the same support level (\$1,000 a year), no matter what the generosity of the NIT plan his family had been assigned. These anomalies in the rules of the two experiments may have been the cause, or part of the reason, for the inconsistencies in the findings on marital breakup. The rules in Gary provide a disincentive for marital breakup relative to the Seattle-Denver rules. Similarly, the rules in Seattle-Denver provide a relatively greater incentive to break up for the

least generous plans--as borne out by the empirical results. These considerations suggest that differences in rules must be carefully integrated into any cross-experimental analyses.

SUGGESTIONS FOR FURTHER RESEARCH

The income maintenance experiments provide considerable evidence that an income support program based on a system of retrospective monthly self-reporting of income is not only feasible, and possibly less expensive than AFDC programs with similar support levels, but also responsive to the needs of participants. The evidence also suggests that families will use the support payments to improve their lives in ways that will reduce their dependence on income support in the long run.

However, some of the evidence is inconsistent across experiments: in other cases, the consistency might be partly illusory. What is needed to resolve these questions is careful research which examines the data from all of the experiments using a common methodology. DHEW is planning to address some of the major open questions as part of its research program on the experiments in the next two years. During the next year, the Policy Studies Division of MPR will create a data file from the New Jersey, Gary, and Seattle-Denver experiments. During the same period, researchers at MPR and SRI will re-estimate the average work effort response estimates reported in Table 2 for these three experiments, using a common methodology and similar data sets and definitions. During the second year, more sophisticated approaches will be used to pool the data from all of the experiments to investigate differences in the work effort and family breakup responses.

My review of the research on the experiments to date suggests that the analysis of the pooled data should carefully account for differences among the experiments in labor market conditions and the availability of other welfare programs. This analysis should also endeavor to integrate the differences in the rules of the experiments in the analysis. Finally, it should exploit the lower support levels in the New Jersey and Gary experiments in simulating the experimental findings of welfare reform plans that are less generous than those tested in Seattle-Denver.

Another area that remains to be investigated is the linkage among the work effort responses, family stability effects, and effects on family well-being. Are the beneficial responses concentrated in the same families? For example, are the beneficial effects on children concentrated in the same families as the other beneficial effects on housing, consumption, and debt? Do the relatively few teenagers who drop out of school live in the families where the household heads reduced their work effort? Or are there offsetting experimental effects? For example, are the beneficial effects on school performance concentrated among those families where the mother reduced her work effort, or where the father and mother stayed together? These kinds of questions provide a large and important agenda for further research.

II

SIME/DIME DESIGN AND
DATA HANDLING

THE STRUCTURING OF
SIME/DIME: AN OVERVIEW

by

Robert G. Spiegelman
Director, Center for the Study of Welfare Policy
SRI International

A negative income tax (NIT), in theory, is a particular type of income maintenance program differing from present programs like Aid to Families with Dependent Children (AFDC) in that: (1) its components are more rationally selected and designed to permit a particular rate of reduction in payments as the recipient's earnings increase (the tax), (2) it would be tied to the regular income tax system, and (3) income maintenance payments would be considered a "right", available to any family whose reported income falls below a designated amount.

An NIT resembles other income-conditioned transfer programs in that it has: (1) a support level (the amount available to a family with no other sources of income); (2) a tax on the income available to the family from other sources; and (3) a time period over which the support is guaranteed (the accounting period).

EXPERIMENTAL COMPONENTS

Support Level

The support level in the Seattle and Denver Income Maintenance experiments (SIME/DIME) is the amount of income guaranteed to the family over a period of one year. This annual guarantee should not be confused

with a monthly system, such as used by AFDC; a \$3,600 annual guarantee does not guarantee \$300 each month. Annual accounting eliminates payments to people who have seasonal periods of low-income, such as construction workers, and provides equity across families with different annual income profiles. On the other hand, it creates adjustment problems for persons suffering sudden reduction in income.

The SIME/DIME support levels were set at \$3,800, \$4,800 and \$5,600 per year for a primary family of four persons, adjusted to account for actual family size and status in the household.¹ Support levels were increased or reduced according to family size (see Table 1).²

The lowest support level of \$3,800 was selected because it was just sufficient to bring all family incomes up to the government's poverty level, and matched (in most cases) the support available from AFDC and the Food Stamp Program. This was absolutely necessary; if the experimental treatment was not superior to that available from these programs, many families would undoubtedly refuse to participate in the treatment group. Financial superiority is especially necessary in states like Washington and Colorado that have both AFDC and AFDC-unemployment (AFDC-U) programs. Differences in work incentive or other effects between the \$3,800 support level selected for SIME/DIME and AFDC will be mainly caused by the difference in system operation and recipient attitude.

The middle program of support at \$4,800 reflects our estimate of the minimum differential (\$1,000) needed to yield measurable differences in effect. In addition, it equals the highest level available from the

TABLE 1

SUPPORT GUARANTEE BY PROGRAM AND FAMILY SIZE

Number of Family Members	Primary Family Index	\$3,800 Program		\$4,800 Program		\$5,600 Program	
		Primary Family	Secondary Family	Primary Family	Secondary Family	Primary Family	Secondary Family
2	.62	\$2,356	\$1,984	\$2,976	\$2,604	\$3,472	\$3,100
3	.83	3,154	2,656	3,984	3,486	4,648	4,150
4	1.00	3,800	3,200	4,800	4,200	5,600	5,000
5	1.12	4,256	3,584	5,376	4,704	6,272	5,600
6	1.23	4,674	3,936	5,904	5,166	6,888	6,150
7	1.32	5,016	4,224	6,336	5,544	7,392	6,600
8	1.38	5,244	4,416	6,624	5,796	7,728	6,900

New Jersey experiment, and thus provides an important point of comparison between SIME/DIME and New Jersey.

A third support level of \$5,600 was introduced into the experiment to widen the range of support payments, and improve our ability to statistically distinguish the marginal effects of higher support levels. In the real world it is highly unlikely that Congress would consider support levels this high (equal to \$8,500 in 1978 dollars).

These support levels have been approximately maintained over time in real dollars by the application of an automatic cost-of-living escalator, which is adjusted every quarter. The following shows the status of the nominal support levels at the end of the three- and five-year program in Seattle.

COST-OF-LIVING INCREASE
IN SUPPORT LEVEL

<u>Initial Support</u>	<u>April 1974</u>	<u>April 1976</u>
\$3,800	\$4,320	\$5,200
4,800	5,460	6,560
5,600	6,360	7,650

The NIT Tax

As noted above, the NIT system requires reduced benefits to the family as its other earnings increase. SIME/DIME chose both constant and declining tax rate systems. Each has particular advantages.

A constant rate is simple to administer, to analyze, and to explain to families. On the other hand, a constant tax rate provides limited data on tax rate effects and is probably inefficient in terms of work effort goals.

In a declining rate tax system, the rate starts high and declines as income increases. It can be shown that, compared to a constant tax, a declining tax will tend to push some people into corner solutions. That is, if they tend to work very little, they are very apt to work not at all. If, however, they tend to be work-oriented, then the declining rate provides a greater work incentive than a constant rate. Furthermore, if government policy establishes both a support level and a breakeven point (the income point at which the grant will be reduced to zero), then a declining rate system yields a lower cost program. A declining rate system could also be integrated into the positive tax system without discontinuity in the marginal tax rates.

SIME/DIME chose four different tax systems for experimentation. Two of these used a constant rate (at 50% and 70%) and two used a declining rate (with an initial tax at 70% and 80%, both declining at an average rate of 2.5% per \$1,000 of income, or at a marginal rate of 5% per \$1,000 of income). No experiments were conducted with declining rates starting at 50%, or with a support level of \$5,600 combined with a 70% declining rate, because these would fail to provide for full recovery of the grant before a zero tax is reached. Also, SIME/DIME did not experiment with a constant rate system at 80% because it was believed to be too confiscatory to be of any practical interest to policymakers.

The experimental tax rate was kept at the assigned level by reimbursing all or some of the income taxes paid to federal and state governments. In addition, all supports from other income conditional programs were taxed at 100% to eliminate their effects.

Composite Picture of an NIT

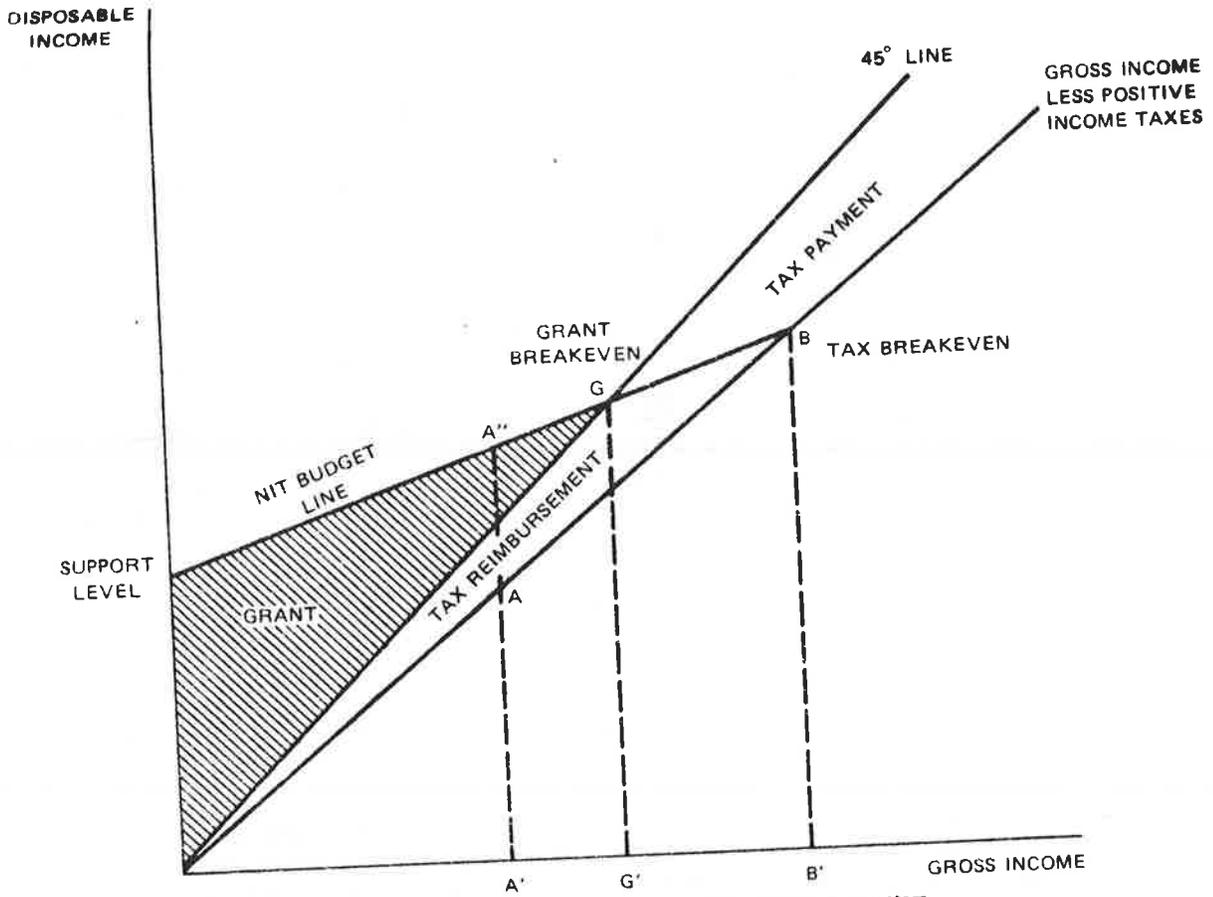
The total income a family enjoys depends on the support level and the tax rate, as shown in Figure 1. The horizontal axis shows gross income (income before taxes or transfer payments such as welfare grants); the vertical axis shows disposable income (income after taxes and with transfer payments). If gross income is zero, the NIT grant is equal to the support level. Without the NIT, a family with gross income A' would have disposable income A after paying positive income taxes; with the NIT payments that family would have disposable income A'' .

Notice that the NIT payment has two components: a grant and a reimbursement of positive income taxes. At gross income level G' , the NIT grant has declined to zero, but the family still benefits from the program by receiving reimbursement for its positive income taxes. Between the gross income levels G' and B' , the family receives partial reimbursement of its positive income taxes. Families with incomes above the tax breakeven level, B' , do not receive any benefits from the NIT program.

Table 2 displays all of the NIT tax plans subject to experimentation in SIME/DIME. There were eleven such plans, labeled F1, F2, F3 and so forth. Four of these tax plans were assigned a support level of \$3,800; four, \$4,800; and three, \$5,600. Table 2 reveals some interesting aspects of SIME/DIME. Some of the experimental plans provide payments to families with rather high incomes. For example, families on the plan designated as F7 do not reach the tax breakeven point until earned income equals \$19,700.00.

FIGURE 1

A NEGATIVE INCOME TAX PROGRAM
WITH POSITIVE TAX REIMBURSEMENT



NOTE: Figure assumes no income outside of earnings and a linear positive income tax system.

TABLE 2

SIME/DIME BREAKEVEN LEVELS
(1971 Dollars)

Plan	TAX		Supp't Level	Grant Breakeven Level	Tax Breakeven Level
	Initial Rate (%)	Rate of Decline per \$1000 Income			
F1	50	0)		\$ 7,600	\$10,250
F2	70	0)		5,429	6,350
F3	70	.025)	\$3800	7,367	10,850
F4	80	.025)		5,802	7,800
F5	50	0)		9,600	13,150
F6	70	0)		6,867	8,520
F7	70	.025)	\$4800	12,000	19,700
F8	80	.025)		8,000	11,510
F9	50	0)		11,200	15,700
F10	70	0)	\$5600	8,000	9,780
F11	80	.025)		10,360	16,230

Note: These figures are for a family of four with only one earner and no income outside of earnings. Positive tax reimbursements include the federal income tax and Social Security taxes. The calculation of federal income tax assumes a standard deduction. Colorado income taxes (there is no state income tax in Washington) are ignored in calculating the tax breakeven levels. The actual tax breakeven levels for the Denver families would be slightly higher than shown.

SIME/DIME is the only income maintenance experiment to date to include a manpower program. This component was included in an effort to offset some of the negative work effort effects of the NIT. The NIT discourages work effort by lowering the value of market work. The manpower program is intended to improve job search and increase wage rates, thereby increasing the value of market work.

The manpower program is conducted at three levels, called M1, M2 and M3. Briefly, M1 comprises only counseling services, M2 includes counseling services and a subsidy of 50% of the direct costs of any training taken over the life of the experiment, and M3 includes counseling plus subsidy of 100% of the direct costs of training.

Counseling is provided by members of a special staff associated with community colleges in Seattle and Denver. Any member of an enrolled family, 16 years of age or over and mentally and physically capable of gainful employment is eligible. Since people are not always aware of the value of counseling, a contact was made to each eligible individual, encouraging the use of the service. In fact, by the end of three years, 45% of all eligible persons had used the service.

SIME/DIME had several experimental goals for this counseling component. First, a relevant informational package must be determined for each individual based on his or her characteristics, aptitudes and interests. Second, the staff must deliver this information in a clear way, so that the recipient understands it and is able to use it effectively. Finally, every individual with the same characteristics, aptitudes and interests must receive the same treatment. Since the essence of the counseling program is to allow indivi-

duals to make efficient decisions in terms of work, then it follows that the counselors should not interfere in the decision-making process. The individual voluntarily decides to seek counseling, and after, must make individual decision, influenced only by the information acquired through counseling.

Training subsidies of 50 and 100% of direct costs were provided to the sample assigned to the M2 and M3 groups. These subsidies were not conditioned on the selection of any particular training program. The only time limit was that dictated by the total experiment--three years for most of the sample, and five years for a subset of the sample. This formulation of the training subsidy component has "textbook" simplicity, in that it serves essentially to reduce the "price" of training, without altering the set of training options provided by the market. The treatment performs the same function as a voucher for education or training. Although the range of subsidized training was wide, some limits were imposed. Any education or training course was eligible for subsidy as long as it represented preparation for an occupation or career. Some schooling that could be considered primarily as consumption (such as flying lessons) was disallowed. The subsidies were intended to cover all reasonable direct costs of instruction. Reimbursement for tuition was limited to the lowest cost alternative that provided the desired course if instruction. Thus, SIME/DIME would not pay Harvard tuition, if the same program was available at the University of Washington. For a variety of reasons, the largest single provider was the community college in each city. Of those who entered counseling and had a subsidy option, about half entered a training or education course.

SAMPLE DESIGN

Control Over the Environment

The very essence of the experimental approach is the exercise of control over the environment in which the experiment takes place. This environmental control is accomplished in SIME/DIME in two ways. First, part of the sample is assigned a control status. The control group is the same as the treatment group, except that it receives no treatments. It is monitored by the interview process in exactly the same way as the treatment group. The control group does not represent a true null treatment. It represents the status quo, receiving supports and paying taxes just as the rest of the population would. Comparing the behavior of experimentals and controls is the most straightforward method of determining the impact of the experiment. The second means of control relies on comparison among experimental treatments.

Sample Stratification

Once the budget for each year of the experiment was known, the sample could be developed. The sample was stratified by location, race, and family composition. The stratification by location was, of course, Seattle and Denver. This raises some important issues. Why did we conduct site specific experiments when the purpose of these experiments is to provide guidance for national policy? Why didn't we select a random sample of the U.S. population for the experiment? If there are significant response differences by site, site-specific experiments give us limited information for national policy, while a random sample of the entire population would probably require a prohibitively large sample. If on the other hand, the site differences are

minimal, then site-specific experiments are satisfactory, and certainly cheaper. Preliminary evidence from Seattle and Denver and crude comparisons across experiments (New Jersey, Gary and SIME/DIME) suggest that mean work effort responses are of the same orders of magnitude at each site. Our findings not only give us confidence in extrapolating findings to the nation, but suggest that it is reasonable to merge site-specific results for some analysis.

If future analysis shows that the effect of location is strong and interacts with a large number of other variables, two different equation systems may be needed for Seattle and Denver. The data permits separate or combined analyses.

Location plays an important role in our study, mainly because of the great differences in the rate of unemployment in the two study sites. As discussed above by Ms. Allen, at the start of the experiment, Seattle had an unemployment rate of 12%, or twice the national average. This situation posed serious problems for an experiment designed to measure the labor supply response to income maintenance programs. In a situation of low and declining job opportunities, it is hard to distinguish the effects of changes in labor supply from changes in demand, unless some adequate control is provided through comparable information gained in a more favorable labor market. Thus, circumstances forced us to select a second site to provide information on effects in a more normal labor market.

Denver was the city selected as the second site, because among all the cities it best met the following criteria: First, it had an unemployment rate below the national average, and was experiencing growth in total employment; therefore, supply of job openings was guaranteed to be sufficient to

measure labor supply response. Second, Denver had a diversified economy which served as protection against future idiosyncratic problems that might arise in an area dominated by a particular industry (such as aerospace in Seattle). Third, Denver was similar to Seattle in terms of (1) demographic characteristics of its population and (2) its sociocultural background. Denver also had the incidental advantage of permitting the experiment to be expanded to study Mexican-Americans as a major segment of the poverty population.

The sample was also deliberately structured so that if racial characteristics significantly affect response, we could measure separate functions for each of three different racial or ethnic groups: black, Mexican-American (Denver only), and "other white" (all whites other than those identified as Mexican-American). We eliminated all other minority groups to achieve socially homogeneous groupings for analysis.

Finally, the sample was stratified by family composition. This characteristic has many dimensions, including size, number of children, and the ages, sexes and relationships of the family members. These attributes may all be involved in the response to an NIT. Therefore, an effort was made to achieve some degree of homogeneity in the population mix by confining the experiment to families having certain structural characteristics. First, the family, rather than the household, was selected as the basic experimental unit, on the assumptions that the family is more stable and more likely to be a resource-pooling unit. The families that were allowed in the experiment were those containing either (1) an adult with a dependent child (and could also

an unmarried, cohabitating partner); or (2) a married couple. An eligible family must have at least two members; once a family met the basic eligibility criteria, it may have attached to it other children or other relatives. The sample was stratified according to whether the unit was headed by a single adult or a couple, with 60% of the budget allocated to couples. Labor supply response from these two types of families was expected to be significantly different.

Another restriction on the eligibility of families to participate in the program is that a single head, or the male of a family headed by a couple must be between the ages of 18 and 58 at the time of enrollment and physically capable of gainful employment. A child who turned 18 during the experiment could also become a head of a family.

ASSIGNMENT TO TREATMENT GROUPS

Within each of the twenty strata represented by two cities, two races (plus one more for DIME), two family types, and two experimental lengths, families were assigned to treatments by a mathematical model that sought to maximize the amount of information obtainable at a given budget.³

This introduced a new, nonprogrammatic variable affecting the assignment model: a family's "normal income" (the income predicted for the first year of the experiment, excluding the NIT and other transfer payments). The estimates of normal income assume normal circumstances for the family and the regional economy, and are based on pre-experiment data that eliminates transitory components (such as absence from work to participate in a training program, or pregnancy) and exceptional and temporary fluctuations in the

economy (such as those facing Seattle because of the decline in Boeing's fortunes).

For assignment purposes, every family income was converted into the equivalent income for a family of four (using the family size index shown in Table 1). Families were then grouped into seven income classes, as shown in Table 3.

TABLE 3
INCOME GROUPINGS FOR
ASSIGNMENT PURPOSES

<u>Class</u> <u>(E Level)</u>	<u>Family of 4</u> <u>Equivalent Income</u> <u>1970-71 Dollars</u>
1	less than \$1,000
2	\$ 1,000 - 2,999
3	3,000 - 4,999
4	5,000 - 6,999
5	7,000 - 8,999
6	9,000 - 10,999
7	11,000 - 12,999

Assignment to treatment was limited to one-worker families with incomes below \$9,000 (or its equivalent) (classes 1 through 5), and two-worker families with incomes below \$11,000 (classes 1 through 6). These income limits were imposed because families at higher incomes were judged to be so far above all support levels and most cut-off points, that they would not reduce their work effort in response to the program. The higher limit on two-worker families was chosen because of a need to include more couples with working

women for observation purposes,

The Assignment Requirements Table

The assignment model was applied in 1970 with a total budget constraint of \$6.3 million per year for program operations,⁴ exclusive of the supplemental budget for a Mexican-American subsample in Denver. This budget was first divided equally between Seattle and Denver, and within each city, equally between the black and "other white" population. Each racial group in each city had a budget of \$1,575,000, of which 75% was allocated to a three-year sample and 25% to a five-year sample. Later, the assignment methods used for black families were applied to Mexican-American families in Denver. Additional families were assigned to control and manpower-only status, to help assure that effects due to city differences could be measured.

The final assignment requirement totalled 5,202 families in both Seattle and Denver, of which 4,800 were enrolled. The distribution of the enrolled sample by city and race is shown in Table 4.

TABLE 4

	Seattle		Denver		Total	
Black	901	44%	961	35%	1862	39%
Other White	1141	56	930	34	2071	43
Chicano	<u>0</u>	<u>0</u>	<u>867</u>	<u>31</u>	<u>867</u>	<u>18</u>
Total	2042	100%	2758	100%	4800	100%

The higher white sample reflected the expectation that black families would evidence a somewhat higher labor market response and family separation rate than white families. The smaller Mexican-American sample in Denver is due to the absence of the special control assignment that was allocated only to the families that would provide a comparison between Seattle and Denver.

Table 5 shows that 58% of the total sample was assigned to some financial program, while about half of the remainder are on a manpower program only, and the remaining 22% receive no treatment. The sample design permits measurement of the separate and combined effects of the NIT and manpower components.

TABLE 5

	ENROLLED FAMILIES BY TREATMENT AND SITE					
	Seattle		Denver		Total	
No treatment	518	25%	523	19%	1041	22%
NIT-only	369	18	577	21	946	20
Manpower-only	417	20	595	22	1012	21
Manpower-NIT	<u>738</u>	<u>36</u>	<u>1063</u>	<u>39</u>	<u>1801</u>	<u>38</u>
Total	2042	99%	2758	101%	4800	101%

Note: Total percentages are not 100% because of rounding.

Sample Selection

To acquire the 4,800 families enrolled in SIME/DIME, an elaborate and time-consuming process of finding, culling and selecting took place at the two sites. Table 6 shows the major steps in the process. Sample selection started with the listing of almost 100,000 households in the two cities.

TABLE 6

SIME/DIME SAMPLE CREATION

<u>SIME/DIME Sample Creation Viewed as Percentage of Original Housing Unit Listing</u>	<u>SIME</u>	<u>DIME</u>
Sample Housing Unit Listing	36,024 100%	57,827 100%
Sample Selected (DIME only)		53,581 92.66%*
Vacant	3,512 9.75%	5,192 8.98%
Not Found/Not Complete	3,542 9.83%	4,642 8.03%
Refused Screening Interview	5,352 14.86%	4,079 7.05%
Terminated (SIME only)	13,432 37.29%	
Miscellaneous	60 0.17%	
Completed Screening Interview	10,126 28.11%	39,668 68.60%
Coded Screening Interviews on file at SRI	N.A.	39,668 68.60%
Ineligible (SIME only)	3,428 9.52%	
Eligible for Pre-experimental Interview	6,760 18.77%	N.A.†
Added from ineligible list (SIME only)	221 0.61%	
Total Selected for Pre-experimental Interview	6,981 19.38%	7,350 [†] 12.71%
Moved	647 1.80%	486 0.84%
Refused	443 1.23%	1,258 2.18%
Ineligible	789 2.19%	591 1.02%
Miscellaneous (duplication, incomplete, false screening, excess 1-parent, employees, unaccounted)	413 1.15%	332 0.57%
Completed Pre-experimental Interviews	4,689 13.02%	4,683 8.10%
Added from new families who moved into houses of previously "ineligible" families who had moved out	126 0.35%	
Total Completed Pre-experimental Interviews	4,815 13.37%	4,693 8.10%
Assigned to Treatment	2,542 7.06%	3,361 5.81%
Refusals	126 0.35%	246 0.43%
Terminations	84 0.23%	83 0.14%
Moved out of area	106 0.29%	111 0.19%
Miscellaneous	183 0.51%	163 0.28%
Total Enrollment	2,043 5.67%	2,758 4.77%

* Percentages are based on the housing unit listing (i.e., for sample selected, 92.66% = $\frac{53581}{57827}$). All percentages may not add up due to rounding.

** The extra 62 coded screening interviews at SRI included some terminated interviews. Note also that 1655 of these completed screening interviews also represent completed pre-experimental interviews due to the fact that these two followed one another during the second set of interviews.

† At least 13,202 of the completed interviews were for families eligible for the pre-experiment.

These listings were drawn from census tracts with incomes below the national mean. About one-third of all the housing units in the cities were contained in these tracts.

After listing, the next step was the administration of a five-minute screening interview to determine if the household contained a nuclear family with characteristics that might make it eligible for enrollment. The screening process resulted in 50,000 interviews. (An additional 13,000 interviews were terminated in Seattle on the basis of clear ineligibility, e.g., over-aged head or no nuclear family.) Through screening, we acquired a sample of approximately 20,000 potentially eligible families. These formed the list of families eligible for an extensive pre-experimental interview. Of these, 1400 families were selected for a detailed 1½-hour interview that provided the information necessary to estimate family income and otherwise determine the family's eligibility. Some 6,000 potential interviewees were not selected for a pre-experimental interview in Denver, as they fell into categories where we had more than a sufficient number of observations. Due to refusals, moves, and other reasons, 5,000 of the attempted interviews were not completed. Of the 9,000 that were fully interviewed, 6,000 families were assigned to treatments, and, finally, after additional refusals and moves, etc., 4,800 families enrolled in the experiment.

In the end, only 1 of 20 screened households were enrolled in SIME/DIME. Most of the remaining households contained ineligible household units. However, 11,000 were refusals at either the screening, pre-experimental interview, or enrollment stages, and more than 9,000 represented household

units that had moved away or were not found at home after repeated tries, These omissions could have biasing effects on the analyses, and are now being investigated,

CONTROLLING FOR DURATION OF THE EXPERIMENT

One of the major issues in income maintenance experimentation is the need to measure long-run responses on the basis of the results obtained in an experiment of limited duration. This need will be fulfilled if the experimental family makes the same kind of an adjustment to the experimental treatment that it would make to a national program of indefinite duration. However, it is likely that the adjustment made to the experiment will be incomplete in this regard,

To directly measure duration effects, some families were placed on three-year programs, while others were on identical programs for five years; a small sample was subsequently assigned to treatments for twenty years, About 25% of the annual program budget was assigned to families on the five-year plan and 75% to those on a three-year plan, At the beginning of the experiment, support staff told each family the length of the guaranteed payment period. We are now attempting to estimate the difference in response as a function of the difference in program length. On the basis of such an estimate, we hope to be able to extrapolate the results to provide an estimate of the long-term response.

CONCLUSION

Overall, SIME/DIME had eleven NIT plans, a control group, three manpower treatments and a manpower control group, yielding 48 possible combinations.

Added to that, about 75% of the families were on the program three years and 25%, five years; and a small group will remain for 20 years. SIME/DIME is a large, complicated and important experiment, which is now drawing to a close after 8 years of operation. The papers which follow present our preliminary findings and discuss the procedures developed to handle the complexity of the design, which has been discussed above.

NOTES

1. A primary family is one that either owns the dwelling unit or is responsible for paying the rent. A secondary unit is a family that resides with a primary family as a subsidiary unit and may or may not contribute to the rent or housing cost. Its support level was \$600 less than a primary family's.
2. The table was developed according to then-current indices of actual differences in cost for larger and smaller families, but current research suggests that the ratios chosen underestimated (overestimated?) the additional expense of a larger family.
3. See pp. 234-38 below for a description of the assignment model.
4. This budget did not include any set-up costs or annual costs of operations that would not vary with the size of the sample, e.g., system design, evaluation expenditures, SRI administration.

DATA COLLECTION

by

Cheri Marshall
Vice President, MPR

and

Gary Christophersen
Seattle Site Director, MPR

Most of the data collected for the Seattle and Denver Income Maintenance Experiments (SIME/DIME) came from two sources: (1) initial and periodic interviews of both treatment and control groups conducted by carefully trained staff; and (2) monthly reports prepared and submitted by the families receiving payments. These data were used for both analytical purposes and to determine eligibility and size of the payment.

The personal interviews generated data for analysis of differences between experimental and control groups and is the most important data set for research analysis. The monthly reports produced data only on families receiving a financial treatment, but nonetheless has yielded valuable information on family response and differences between treatment groups.

This paper will describe this data in general terms: the creation of the instruments; the interviewing and reporting techniques; efforts towards error control; and efforts to perfect the collection of the data. (SIME/DIME also obtained data from files of other agencies, as discussed below in the paper by Hall and Kehrer.)

A RESPONDENT'S EYE VIEW

From the participant's perspective, the data collection process began when an interviewer arrived at the resident's doorstep explaining that she/he was from "Urban Opinion Surveys," and was helping to conduct a study for the Department of Health, Education, and Welfare (DHEW) on the economy of the area. A ten minute interview followed, made up of questions about the household composition and income.

This was the SIME/DIME screening interview. From it the research team identified families who were probably eligible for the experiment. The screening interviews yielded more subjects than were needed in some categories. Higher-income, two-adult families, lower-income, single-parent families, higher-income whites and lower-income minorities were over represented. The size of these groups was reduced randomly before the next round of interviewing.

Shortly after the screening interview, another interviewer appeared, explaining the Urban Opinion Surveys wanted more information. This time the interview lasted one and one-half to two hours and covered topics such as wages, working hours, fringe benefits, flexibility in setting hours of work, overtime patterns, and the costs of travel to work. These questions were asked about every job in the last year for every member of the family. Periods of unemployment were queried in similar detail. Finally, the interviewer turned to the family, with questions about education and job training, opinions about various work-related issues and family relationships. Some questions were asked only of the husband, or only the wife. The interview ended with a \$5.00 payment to partially compensate the family for its time.

This second interview was the baseline or "pre-enrollment" interview, and was used for several purposes. First, it provided a more refined and accurate measurement of income for a final eligibility determination. Second, it allowed for testing the randomness of the random assignment process since there should be no identifiable differences between groups based on pre-enrollment data. Finally, it provided baseline data against which changes during the experiment could be measured. After this interview, the sample was further reduced in size, eliminating any remaining ineligibles and randomly culling subgroups which were still larger than necessary.

A few weeks after the pre-enrollment interview, yet another person contacted the family, explaining that the research group would like to enroll the family in an experiment for three (or five) years, to be interviewed three times per year. The family was again interviewed and paid a modest fee for this effort -- the "enrollment" interview -- which updated the family's employment and income data. If the family was in the more fortunate group, they would have their financial and/or manpower plan explained to them at that time. Enrollment took place over about a year, beginning in Seattle in November 1970, and in Denver in late 1971. Families agreed to abide by the rules of the experiment, to participate in the periodic interviews, to report changes in family composition each month, and, for those selected for the financial treatment, to provide monthly written reports on income and certain expenses.

In return, SIME agreed to keep all identifying data entirely confidential, and to compensate the respondents for time spent on interviews and

reports. Those not selected for financial treatment--the controls-- received a stipend of \$8.00 monthly for returning a postcard reporting their current address and any changes in family composition. (This helped combat sample attrition.) Around 95% of those offered an assignment to either treatment or control accepted the offer.

After that, an interviewer appeared every four months and asked a standard set of questions about jobs and income along with a very eclectic assortment of other questions. Some questions were asked only once during the entire experiment while others repeated in specified cycles. Some were addressed to the husband, some to the wife. The interviewers followed each piece of the family if it split up, and followed them anywhere in the country if they moved. (However, if they moved a distance, the interviews would occur only once a year.)

DESIGN OF THE QUESTIONNAIRES

From the respondent's point of view, the periodic interviews probably seemed simple (if endless). This was due to heroic efforts on the part of the staff. In the survey design area, there were three major problems to solve. First, we had to select and schedule into manageable formats an enormous number of questions. Each question was justified, formulated, pretested, and polished. Second, almost every topic was profoundly influenced by the nature of the lives led by low-income families. Concessions were made, ideas changed and measurements improved to deal with uncertainty, instability and rapid change as well as a lower educational level prevalent among the working poor.

Third, frequent changes in the composition of the families themselves dictated a more complicated survey instrument than was originally anticipated. Yesterday's youth became today's household head; families gained and lost children; children gained and lost parents. Worse, people in one experimental plan met and married people in another experimental plan. These changes required responses in both the regular periodic survey design and in the form of special interviews.

Topic Selection and Scheduling

The scope of research reports is a good indicator of the breadth and depth of data collected by the experiment. For every hypothesis there was at least one, and sometimes several complex models. Each model contained many variables, each of which required careful measurement. Because of the state-of-the-art nature of the designs, we frequently tried several formulations of a variable or methods of measuring it since no well-tested measurement was available. It is obvious that the experiment generated an enormous demand for data arising from these requirements.

At times, families must have wondered at our originality in devising methods (some of which failed) at creating new measures for difficult and abstract concepts. Once when grappling with measurements of subjective discount rates, we asked families whether they would prefer one free trip to New York this year, or two free trips next year. Most families told us that they didn't want any free trips to New York at any time.

All in all, in addition to the standard "core" questions asked each time about family composition, labor force participation and income, some 56 special topics or "modules" were scheduled across the regular periodic interviews.

A researcher's preference for response by one or another particular member of the family added another factor competing for space on the surveys. Some questions were directed at certain individuals, e.g., each worker was interviewed about his or her own earnings and job experiences. Others called for contrasting answers from different family members. Thus, many topics were measured not only over time, but also between respondents. For example, SIME/DIME suspected that a woman's decision to work was influenced by her own and her husbands' attitudes towards her working so both were interviewed about this. Moreover, such differences of opinion might help explain marital instability--another target for the research.

These kinds of issues generated some scheduling problems. Often a choice had to be made between having the husband and wife answer a question at the same time each year, in which case a difference of opinion could be pinpointed, or having them answer at alternating interviews, but inter-respondent influences could not be controlled. Differences in responses, moreover, could reflect either actual differences in opinion, or changes in opinion since the last interview.

Another important scheduling problem concerned balance in the interview. Each interview had to make concessions to the fact that the respondents were humans, not computers. Each interview ideally balanced hard and easy topics, boring and interesting topics, questions for the husband and questions for the wife, and sensitive and non-sensitive questions.

Finally, different topics were cycled at different intervals; some appeared only once. Choosing the time interval or actual time involved a number of factors. How likely were responses to change, and how quickly

would they change? Would the accuracy of responses be damaged by length of time since the change, because of recall difficulty? Would the responses be affected by seasonality? This would be true, for instance, of some health status questions. Should responses be obtained for Seattle and Denver at the same length of time since enrollment or the same real time?

Effects of the Experimental Population on Survey Design

It is easy to see that the demands for both precision and breadth imposed considerable demands on the survey design. Add to these demands the ones endemic to the sample itself. Many sample members were poorly educated, limiting the level of concept and vocabulary which could be used in the surveys. This ruled out duplication of some survey questions or groups of questions which had proven useful with more middle-class samples.

A second (and in many ways more challenging) factor in designing measurements for the sample was the unusual complexity of the economic and sociologic lives of the low-income population. One outstanding example is the measurement of work effort and income. Unlike middle- and upper-income families, low-income families do not, by and large, experience a steady, albeit low, income stream during the year. A typical pattern is that of rapid fluctuation--feast and famine--although even the "feast" is modest. Many of the working poor experience rapid job turnover and changes of hours or even wages within jobs. Families made many stopgap adjustments to these uncertainties, with one person's job loss triggering another person's entry to the job market. It was obviously important for research reasons to capture these fluctuations.

This complexity in economic conditions dictated development of non-traditional measurements. Respondents would have trouble with the usual questions such as "How many hours did you work last month altogether, on all jobs?" However, they could easily recall each individual job, its wage rates, and when they worked on it, as well as other details. For these reasons we moved to a question pattern which was very detailed and simple to answer.

With such detailed questions, however, the possibility of contradiction on the part of the respondent and clerical error on the part of the interviewer was aggravated. Thus, many cross-checks needed to be built into the survey instruments. For instance, the interviewer checked responses on wages and hours against the respondent's estimate of total monthly earnings and available pay stubs.

As another hedge against reporting error, certain questions were tied to the previous interview. ("The last time we spoke, in January, you were working at Triple A Plumbing as a laborer. You were making \$3.75 per hour at that time. Did your wage rate on that job change after that?")

A final adjustment to the questionnaires was made on wording. The general vocabulary and concept levels had to be very clear. Questions which, in a standard survey, might have asked about marital history, were changed to ask about "marriages and other relationships you considered permanent at the time." The people interviewed were apt to be very literal in their interpretations of questions. (A pretest question on hours of television watched each day yielded suspiciously low numbers until it was changed to specifically include nights as well.)

Mobility and the Need for Special Interviews

Many special situations required the development of survey instruments designed to collect supplementary data. These included baseline data for new family members, catch-up interviews for people who had missed one or more interviews, and interviews for sample members who moved away from Seattle and Denver. (Movers were interviewed once per year, with key data paralleling the regular periodic interviews throughout the year, but other questions on reasons for mobility and local labor market conditions had to be devised.) Finally, a special survey had to be constructed to allow us to measure work effort for the self-employed, as a surprisingly large segment of our sample had entrepreneurial aspirations from time to time.

SURVEY MANAGEMENT

The demands placed upon survey design by the research plan and by the respondent group carried over into the survey management area where there were several major challenges.

The first management challenge was selection of interviewers. Obviously, their tasks were difficult, but attention had to be paid not just to brilliance but to the image they would present to families and to their willingness to work in low-income areas. It was important that they present an informal yet professional image. They took care to avoid resembling welfare department caseworkers, as SIME/DIME wanted to clearly separate interviewing from both welfare and the disbursement of SIME/DIME payments.

Consequently, interviewers whose characteristics were closer to the experimental population were preferred over traditional interviewers (middle-class housewives picking up extra money working daytimes). The interviewing

staff had many minorities, and many were males.

After selection, interviewers underwent extensive training, which involved up to one week of intensive lectures, role playing and exercises.

A second management challenge was respondent mobility. The respondent group moved frequently, and to maintain sample integrity, they had to be found. The moves were sometimes easily traced, but many were difficult--especially where the family or a part of it was fleeing from the landlord or the bill collector or from other members of the family. This turned our interviewers into detectives--studying the family's file and doing intensive street work. Their work was made more difficult by our promises of confidentiality, which prevented interviewers from disclosing the reasons for their search to potentially helpful outsiders. Despite the odds, the field staffs were extraordinarily successful, obtaining overall response rates of 90% to 95% of the original sample on each wave.

The final management problem was the quality control effort taken back at the office after the interview was completed. Very detailed procedures were developed and enforced to insure the highest possible data quality. Each interview was given a thorough reading by two different members of the highly trained quality control staff. These readings focussed on interview completeness and internal consistency and insured that the vast volumes of policy decisions were applied consistently across all interviews. These two readings often stimulated questions directed to the interviewer who in many cases needed to re-contact the respondent to clear up the issues.

THE MONTHLY REPORTS

As noted above, after a family was enrolled and assigned to a treatment, they were required to submit monthly reports as a condition of receiving a grant. These reports were called the Income Report Forms (IRFs) (reproduced in appendix A).

On the IRF the family reported income from all sources for the month just ended. The report included earnings, transfer payments from other programs, such as Aid to Families with Dependent Children (AFDC) and Unemployment Insurance, business income, the sale of capital assets, child support payments, certain deductible expenses (particularly medical costs and work-related day care costs) and any changes in their family composition that may have occurred during the month.

The challenge in developing the IRF was to make it as thorough as necessary in collecting information, yet simple and short enough that participating families would be both willing and able to regularly and accurately complete it.

The form itself was mailed to each experimental family along with their payment check as calculated from the previous month's IRF.¹ The family was obligated to complete and return it to the payments office within two weeks in order to receive the next check "on time." If a form was late but still within two weeks of the deadline, the check would come two weeks late. If the IRF was not submitted by the end of four weeks, the payment was forfeited.

These filing deadlines were readily met by most families. During the average month about 89% filed their IRF's on time, 8% filed late, and only 3% failed to file and forfeited their payment.

When the IRFs were received in the office payment analysts submitted them to a rigorous series of quality control tests in preparation for data processing and calculation of the payment. They examined the IRFs for completeness, internal consistency, and consistency over time. Each payments analyst had a caseload of about 275 participating families and usually became familiar with a family's idiosyncracies.

The analyst also checked signatures, pay stubs and receipts documenting medical, day care, and business expenses. In examining for consistency, the analyst compared previous IRFs with current ones to check items such as continuity of reporting Social Security, AFDC, unemployment benefits, and so on. The analyst also examined the "year to date" figures on the pay stubs submitted in order to determine whether all earlier earnings had been reported on previous IRFs. At the same time, the analyst coded the IRF in preparation for data processing.

Next, after key punching, we ran a computer audit of the data, including another check for completeness and data validity. The corrected data was then merged with the files containing the prior IRFs for each family, and the payments calculation was computed, based on the merged data; the computer produced a check and a page outlining the check calculation for each family.² These were returned to the payments analyst who compared the checks with prior ones for each family to insure either that it was consistent with past payments or that the IRF data explain any difference. The check, grant explanation, and another blank IRF were then mailed to the family.

In addition, the payments supervisor at each site performed a rigorous audit of a 2% random sample of each month's IRFs. Periodically, random audits are also performed by a professional auditing staff. These steps combined to guarantee a low rate of error and also safeguard against possible fraud.

OTHER CONTACTS WITH THE FIELD OFFICE

The field offices were responsible for handling any family contacts necessary to check on irregularities in the payment process, to answer questions, and to explain the program to families who contacted the field offices for advice or with problems regarding the program.³ The field offices also undertook a special reexplanation project about eighteen months after each family's initial enrollment. This consisted of two stages: an interview to determine existing knowledge levels concerning the rules governing payments and eligibility; and a new explanation to those families whose understanding was found inadequate in any area.

Families generally understood the most basic concept of the payments calculation (that if their income goes up their payment goes down and vice versa), and had a fairly solid understanding of other basic rules. However, more subtle aspects of the experiment required further explanation for many of the families. For example, there was a relatively low comprehension of concepts such as the "breakeven" level of income, how the declining tax rates actually worked and to what degree different types of transfers affected their monthly grants. For the 25% of our sample enrolled for five years, we followed up our first "reexplanation" with another, coming at the end of the third year of participation. A record of all contacts with the field offices was maintained and is another modest source of data on the experiment.

NOTES

1. Roughly, a check for a month is the grant level minus the tax rate times net income (income less allowable expenses). The actual calculation is much more sophisticated than can be appropriately described here. It involves income and expenses for both the current month and the eleven preceding months. This data base is assembled and stored on computer tape from which it is accessed monthly to calculate payments. For a detailed description see Spiegelman, Kurz, and Brewster, "The SIME/DIME Payments System," (SRI Research Memorandum 17, 1973).

2. The page which outlines the calculation of the grant is a computer printout which shows step by step how the month's grant was calculated.

3. To make this latter service convenient for families with questions, we initially had three field offices in each city. In Seattle these were soon consolidated into one office because the number of family-initiated contacts at the outlying offices was very low, and Denver followed suit later on.

DATA COLLECTION FROM COLLATERAL SOURCES

by

John Hall
Senior Survey Researcher, MPR

and

Kenneth C. Kehrer
Director, Survey Division, MPR

As with most large-scale social science research projects, the Seattle and Denver income maintenance experiments (SIME/DIME) depended primarily on data provided directly by the respondents in personal interviews of self-administered questionnaires. One could also use observational techniques but this could be cumbersome, costly, and intrusive on subjects. Another choice, used in SIME/DIME, takes advantage of existing public records that contain data previously reported by the respondents or others. This paper assesses the use of this type of data--called "collateral source" data--based on our experience in SIME/DIME and related studies.

COLLATERAL SOURCES

SIME/DIME

SIME/DIME initially used collateral source data only for the following purposes. Public records such as drivers' licenses and automobile registrations were used to locate respondents who did not report that they had moved. Copies of federal tax returns, obtained from both respondents and the Internal Revenue Service (IRS), were used to reconcile tax reimbursements made to families receiving financial treatment. A monthly audit of Aid to Families with Dependent Children (AFDC) records determined whether SIME/DIME families were

receiving unreported welfare support, These data were later used in the validation study described below by Halsey. However, starting in 1975, collateral source data were also used in analyses of the impact of the experiments on (1) the educational performance of children, (2) the health status of children born during the experiments, and (3) juvenile delinquency.

The most extensive use of collateral source data to date in SIME/DIME has been in the school performance study. (Early results of that study are presented below.) In that study, we obtained individual student records that included elementary academic records, secondary transcripts, compensatory education program files, health records, and aggregate data about the peers and school environment of the sampled youth. Some of these data were provided on machine-readable cards or tape; others were coded. Several thousand person hours were spent coding data for nearly 5,000 students who attended over 250 schools (in about ten school districts) over a six year period. Collecting and preparing the data for analysis was a major and problematic task, given the mix of school districts, schools, types of records, and forms of data (machine readable, tape, card, and manually coded).

Other Income Maintenance Experiments

Collateral source data have also been used in the other income maintenance experiments, but not as extensively as in SIME/DIME. All the income maintenance experiments, for example, used IRS data to reconcile tax reimbursements to families and used official records to validate reports of income. The New Jersey and Rural experiments relied on data from the Social Security Administration (SSA) for validation purposes. The Gary experiment used data from numerous federal, state, and local agencies for an extensive validation effort for the entire Gary sample.¹ The Rural and Gary studies also used collateral source data for analysis purposes.²

THE DATA COLLECTION PROCESS

In SIME/DIME the steps taken in designing and executing the collection of data from collateral sources were not always straightforward. The records from which data were collected were designed for administrative purposes; as administrative needs changed, so did the records. Explanations of the records were obtained from the keepers of the records, sample forms were examined, and even pretests were conducted, but they still did not reveal all of the problems we faced in converting the data to research use. However, we were able to define certain tasks necessary for any effort of this type.

Identification of Data

In converting the data to research use, one must ascertain, first, what data are desirable and, second, which of the data are available and in what form. Where data are sought by a researcher outside the record-keeping agency (as has been the case in SIME/DIME), interaction between the researcher and the agency is necessary to discover the availability of the desired data and how the records are kept: Are the data found in one or many records? Are they found in automated files? Have the forms for manually kept records changed over time? How often are the data entered into the records? Where are the records kept? This investigation may show that a substantial portion of the desired data are either nonexistent or irretrievable, in which case the effort must be reevaluated. In SIME/DIME we have not had to abandon any collateral source data collection, but, after a preliminary investigation we have often revised the list of data to be collected.

Arranging Access

Once the researcher has determined what data are available, a request for access must be made. There are, of course, degrees of accessibility, depending on whether the records are privately or publicly held and whether there are legal, ethical, or customary barriers to releasing the records or the data they contain. When SIME/DIME staff found that access was legally restricted, we either obtained the consent of the research subject or met other (usually narrowly defined) criteria.

To arrange access we first made informal contacts, after which we submitted a formal request that typically included the following:

1. A precise description of the data
2. Justification for the proposed use of the data
3. Legal qualifications for access (e.g., consent forms signed by research subjects)
4. The proposed schedule for collecting the data
5. Methods for collecting the data

Sometimes the records-keeper would simply deny access to the records. At other times, the first agency staff person contacted would grant access to the records, only to be denied by a superior (or vice-versa). We had some success in obtaining records after they were initially denied, by submitting a reapplication (sometimes after waiting a few months or a year) or by approaching a superior.

Designing the Data Collection

This task requires a knowledge of what data are to be collected, in what form they will be found (hard-copy, tapes, etc.), where they are kept, when they will be collected, how up-to-date they will be at the time of

collection, and who will conduct the data collection activities. We have found that to collect all the desired data, it may be necessary to go to more than one type of record. In the SIME/DIME School Performance Study, we used automated records and four types of hard-copy records from only one school district. (The hard-copy records were kept in one of two offices in the last school attended or in one of three central locations, depending on the age and enrollment status of the pupil.)

The design of the data collection effort must allow for (1) data from potentially different sources to be merged into one analysis file, and (2) checks to be made on data quality. It must also attempt to minimize costs and the disruption of the agency's ongoing activities. These goals, however, are not always compatible. For instance, when manually coding hard-copy, an approach that ensures high quality may be too disruptive and costly; or, for example, the degree of supervision available may lead one to opt for a coding form that is easy to use in the field, but that requires somewhat more cost and effort to convert the data to computer-readable files. In SIME/DIME we used such a coding form in one school district because of schedule restrictions and because of the district's sensitivity to disrupting school activities.

If data are to be obtained in machine-readable form, the collector must determine the nature of the data, whether it is compatible with the computer system and planned analysis file, and, if not compatible or immediately usable, how much work will be involved in preparing these data for research use.

Much of the data needed in SIME/DIME was contained in hard-copy administrative files. To prepare these data for research use, we had to

design specific collection forms. First we obtained samples of records--preferably records or copies of records that were already filled out, although we found that blank forms were quite helpful. We then prepared a draft from that was reviewed by the records-keeper as well as by our own staff. After revisions, we conducted a pretest with a sample of records that was large enough to ensure that the more common variations between records and record-keeping practices were discovered. The pretest also helped to determine what skills the coders needed. After the pretest, the data collection plan and schedule were finalized and agreed to by both the record-keeping agency and the research staff at Mathematica Policy Research.

Collecting the Data

As noted, most of the collateral source data were obtained from hard-copy files. The manual collection of this data was comparable to the collection of survey data; however, we used coders rather than interviewers, and respondents became office managers, secretaries, and school principals who did not answer questions, but guided the coders to the records from which they extracted data. Occasionally refusals occurred,³ and administrative records, like respondents, were often impossible to locate or did not provide meaningful information.

Thus, in collateral source data collection, coders need more clerical skills, while in survey data collection, interviewers need more interpersonal skills. However, both must be able to understand the purpose and content of the data collection forms they use, fill them in clearly and neatly, provide adequate marginal notes in cases where doubts or new situations arise, and work independently. We found that recruiting

and training could often be simplified if current or former employees of the record-keeping agency were used.

Once coding began, supervisors assisted the coders by resolving problems with officials who kept the records and by answering questions about problematic situations that were not covered in training. Thus, a supervisor might have been faced with the following questions:

"The office secretary said she never heard of us, and the manager isn't here. What should I do?"

"This school doesn't give A's and B's; they give stars and checkmarks. How do I code them?"

"How can I ever find all these people's records? This office only records the first initial, not the whole first name."

Finally, when manually coding collateral source data, we conducted some manual checking of coders' work as concurrently as possible to the coding. This was accomplished by a combination of (1) having coders cross-check each other's work, (2) having a supervisor recode a few samples of each coder's work, and (3) conducting editing for completeness and internal consistency.

DATA QUALITY

Administrative records can be a source of high-quality data. They are often more objective than survey data, which are affected by respondent recall, intentionally misleading responses, and bias in the interviewing process. Some administrative records, such as those for check disbursements, are presumably complete and are usually audited. Obtaining access to the data does not depend on the ability to find the individual

being studied, nor does it depend in all cases on the consent of the study subject.

However, collateral source data are not devoid of data quality problems. For instance, records may be incomplete or missing, or they may be filled out inconsistently by different clerks or in different offices. (The paper on validation by Halsey cited below suggests that some public records were incomplete.) Even worse, data may be recorded inaccurately, perhaps in some unknown, systematic manner, or the data may be recorded on different forms in different offices or over time. Only careful data collection procedures can overcome such problems.

Our experience suggests that high-quality data can be obtained if the data source is located within one administrative jurisdiction, or if the data are an essential part of the operation of the program that maintains the records. Data quality problems are likely to be severe when data for the same study are being collected from employment insurance offices in several states, or from schools in more than one school district, or in situations where record-keeping is fairly decentralized. For example, in the School Performance data collection in Seattle, we discovered that elementary schools had the freedom to design their own academic-marking system, which rendered much data unusable. On the other hand, we have found striking similarities between birth records across states,

Where data are being used to determine eligibility, make payments, or operate an established program, we have found the data to be fairly reliable. Payments data, which are usually audited appear to be the most reliable.

LEGAL BARRIERS

Legal barriers are perhaps the greatest major obstacle facing the researcher in collecting collateral source data. There are several federal laws that provide privacy for persons who are subjects of public records,⁴ as well as regulations by each agency affected, numerous state and local laws,⁵ and judicial interpretations that must be considered. Legal restrictions may also require that the research and the data collection forms undergo prior review and approval, that the informed consent of research subjects be obtained, that limitations be placed on access to certain government bodies and their agents, and that the use and or release of the data be constrained.⁶ The researcher may also be contractually restricted by the agency that is funding the data collection or by agreements made with research subjects. It is quite possible that legal action may be taken by an individual research subject under one of these protective statutes or regulations, or under a more general theory of invasion of privacy.

Such restrictions are rarely insurmountable barriers to collateral source data collection, although they may cause delays, require extra effort on the part of the researchers, and increase costs (sometimes to the extent that the costs become excessive).

Legal restrictions were surmountable, however, and SIME/DIME has made extensive use of personally identifiable data in public records. These include records of individuals' educational performance; records of income; records of receipt of public support or benefits; and police and court records of juveniles.

Access to the school records has been restricted by both federal

and state legislation.⁷ Other laws have restricted access to such records of income as federal and state income tax returns, SSA files, and welfare office records of public support. Consensual access to such records has often been granted, but nonconsensual access has been granted only in several limited circumstances. For example, SIME/DIME was operating an experimental parallel public-support program under government supervision and therefore was granted access to welfare records. Police and court records of juveniles are usually under the jurisdiction of a juvenile court, and there seems to be more flexibility in releasing those data than would be the case with welfare or IRS files; however, the confidentiality of identifiable data would have to be strictly ensured.

Even when access to records is legally permissible the record-keeping agency may limit, place restrictions on, or even deny access to the records, thereby threatening the data quality or making the effort too costly. Generally, this can happen if the agency interprets the legal aspects of access too strictly or perceives the data collection to be too intrusive or burdensome. While such agencies are not unreasonable, they correctly see themselves not primarily as research agencies, but as agencies having an operational purpose. Thus, permission for research activities is usually viewed as an accommodation, and the agency is generally within its authority to withhold this type of data. However, a researcher may sometimes use the federal or state freedom of information laws to require access.⁸

To ensure a successful data collection effort, the researcher must obtain the cooperation of all parties involved in the data collection process. The funding agency should allow adequate time for obtaining access and be

ready to lend its influence to gain the cooperation of the agency holding the records, Researchers must be prepared to (1) give the record-keeping agency all the information it needs to make its decisions on whether to grant access, and (2) cooperate with such agencies to minimize the intrusive impact of the research being prepared to sacrifice some control over the schedule, the supervision of workers, and quality of data collected. The record-keeping agencies should be aware that those records may hold valuable data for legitimate research, and be prepared to receive and evaluate reasonable requests for access to those records.

EXTRALEGAL ETHICAL CONSIDERATIONS

Beyond the safeguards required by law, the ethical implications of social research must be considered. Generally, these concerns fall into two areas: (1) providing adequate protection for human subjects of experimentation, and (2) protecting privacy.

Social experimentation is a relatively new phenomenon, and the concerns about privacy seem to grow with the recent sophistication of and publicity about automated data systems. The researchers must consider when nonconsensual access to administrative records is appropriate and, if used, whether research subjects should be notified. In obtaining consent to access of a record, the researcher must consider how much information subjects should be given. Finally, the researcher must consider additional restrictions on access to information to prevent damage to the subject.

Although there is no agreement on how to handle these ethical questions, protecting the privacy and integrity of human subjects remains an important ethical, social, and political concern. We also wish to state that sound ethical considerations may well lead to more extensive restraints on researchers than those required by law.

NOTES

1. Kenneth C. Kehrer, et. al., "The Gary Income Maintenance Experiment: Design Administration and Data Files," (Mathematica Policy Research, 1975).
2. Barbara H. Kehrer and Charles M. Wolin, "The Impact of an Income Maintenance Experiment on Low Birth Weight," (Mathematica Policy Research, 1976); Rebecca Maynard, "The Effects of the Rural Income Maintenance Experiment on the School Performance of Children," American Economic Review, 62, February, 1977; Richard J. Murnane, "The Impact of Changing Student Enrollment Patterns on the Distribution of Teachers in an Urban School District," (Mathematica Policy Research, 1977); Rebecca Maynard and Richard J. Murnane, "Will Welfare Reform Improve Children's School Performance?" (Mathematica Policy Research, 1978).
3. Despite the approval of the head of an agency, local entities are often somewhat autonomous, and may refuse to allow any data collection or may limit the collection from records in their jurisdiction. A case in point is that of the school principal.
4. Most pertinent is the Federal Privacy Act of 1974, restricting access to all personally identifiable records held by federal agencies. There are exception for bona fide research use, and the person who is subject of the records can also consent to an inspection. 5 U.S.C. §552a (1976). There is also the Family Educational and Privacy Act, restricting access to school records, applicable to any school district receiving federal funds. 20 U.S.C. §1232g (1976). Finally, of minor interest, was the restrictions on records on juvenile delinquents who were charged in federal courts. 28 U.S.C. § 5038 (1976).
5. Neither Washington nor Colorado have anything comparable to the Federal Privacy Act of 1974, but local agencies will often restrict access, and in both states the public records law expressly exempts personal files on individuals from the general requirement for public inspection. See Rev. Code Wash. §42.17.310(a) (1977 Supp.); Col. Rev. Stat. §24-72-204 (1973).
6. The DHEW regulations requiring review of researcher proposals on human subjects is at 45 C.F.R. §46 (1977).
7. The Family Educational Privacy Act is cited in note 4, above. Note that when DHEW issued these regulations, it recognized the need for special regulations covering social science research.
8. See Federal Freedom of Information Act, 5 U.S.C. §552a (1976); Col. Rev. Stat. §24-72-201 et seq. (1973); Rev. Code Wash. §42.17.250 et seq. (1976, as amended, Supp. 1977).

DISCUSSION

KAY THODE: I would like more detail on the background decisions made with respect to providing adequate protection for human subjects when the SIME/DIME experiemnts were undertaken. I would have assumed that those questions would have been answered by the project.

HALL: All persons who were enrolled in SIME/DIME signed an enrollment agreement which explained the basics of the program--what they were agreeing to do as participants in the program. In addition, they were given a supplemental verbal explanation which explained whether or not they were eligible for financial benefits, what their guarantee level would be, what their tax rate or reduction rate would be, what their obligations were in terms of reporting, and things like that. The heads of households would sign the agreement. Also, every time a family was interviewed they had the option of refusing the interview. Of course, part of their agreement was to provide information to the experiment, so if they refused to provide information their payments would be stopped.

THODE: I am still not sure that is adequate protection for human subjects. For instance, I am very concerned that some subjects were receiving income payments as high as, say, \$280 a month at the point at which the experiment ended. I would feel that any family whose income suddenly decreased by that amount would experience a considerable trauma.

[Gary Christopherson was requested to answer the above question.]

CHRISTOPHERSON: First, disenrollment of the family from the experiment went smoothly although we were actually taking out a big chunk of money from many families' income. The thing that surprised me most is that 99% plus of the families didn't even complain when we gave them their first notice that it was about to end. They were very cognizant of the fact that they had signed a contract in the form of an enrollment agreement for a limited period of time.

There were maybe two families who were enrolled for three years and thought they had been enrolled for five years. Other than that, it was very clear-cut in just about all of the families' minds that it was a very distinct period of time. As we approached the end of the experiment we first provided all families with a notice saying, "90 days from now your income-maintenance grants are going to end." If the family had been on welfare anytime during the experiemnt, on AFDC, or before the experiment, we also sent them a letter saying, "Would you like us to notify your case worker as an assistance in transition?" We also provided them the same kind of service with public housing.

We followed up the 90-day notice with another notice of 60 days, and another notice of 30 days; and we also kept our field office open and available for use as a referral agency after their enrollment and eligibility for the grants ended.

CHARLES METCALF: I would like to make one comment. Often in the interviewing process we will be asking about criminal behavior, and we are getting reports of illegal activity that has not been detected. We maybe have succeeded in finding somebody and interviewing somebody that is currently being sought by the police. In such cases, we see an obligation not to report anything. We have some express protection from federal subpoena, but it does get us into some rather difficult ethical questions. But where we have made an explicit commitment not to reveal such information, we do not.

HALL: My own legal and ethical view is that if we collect information from people under a promise of confidentiality our obligation is to maintain confidentiality if it is legally possible. However, we cannot and should not promise that we can keep these data from all legal process. I don't think we have ever done that. In other words, we don't voluntarily report anything that we get in a confidential setting; however, if someone attacks us through a legal process we may have to give it up.

BERNIE SWOAP: If willful underreporting were discovered, were sanctions applied; and if not, why not? And were they informed ahead of time whether sanctions would not be applied?

HALL: There were procedures for recoveries of overpayment in all cases during the operation of the experiment. If a family was found to have been over- or underpaid, regardless of intent, their future payments would be adjusted. The general rule was to reduce their future grants by \$50 a month until the overpayment was repaid, as long as we could do that within a year.

SWOAP: So the sanction was limited to recoupment?

CHRISTOPHERSON: So far [in Seattle] the sanction has been limited to recoupment, with threat of stiffer penalties if recoupment isn't forthcoming.

SWOAP: How were they advised at the outset?

CHRISTOPHERSON: In the enrollment agreement they agree to report all income accurately and expenses accurately. Beyond that it is taken on a case-by-case basis.

ROBERT WILLIAMS: In Denver we did go beyond recoupment to legal action of a civil nature. In a very few instances we did make referrals to the appropriate authorities for possible criminal action.

JOHN MCCOY: If you found that there had been underreporting or misrepresenting to another agency, was there any action on your part to refer a case to that agency indicating that there was a possible source of cheating? I am thinking of the IRS records, for example,

CHRISTOPHERSON: Absolutely not, because of our confidentiality agreement with the family.

SIME/DIME DATA PROCESSING

by

Virgil M. Davis
Data Processing Manager,
Center for the Study of Welfare Policy,
SRI International

This paper outlines the scope of the data collection and processing effort underlying the Seattle and Denver Income Maintenance Experiment (SIME/DIME) and the main issues in processing these data. It provides initial information to the researcher who may want to attempt independent analysis of the data and explains, to some extent, the long delay in obtaining any final analysis of the data. Finally, the paper makes some suggestions for handling voluminous data, based upon our experiences.

DATA COLLECTION AND ENTRY

The size of the data base, the program variables and the complexity of the data are described above, particularly in the papers by Spiegelman and by Christopherson and Marshall.

As described in the Christopherson and Marshall paper, a periodic interview consists of modules of questions which may vary over time, both in their inclusion in the interview and their internal contents or structure. There is a core set of modules that is included in every interview and a varying set of modules, called non-core modules. The latter include some which were asked on an irregular basis while others were asked with some

frequency, often about once a year (such as net worth, real property, and training and education).

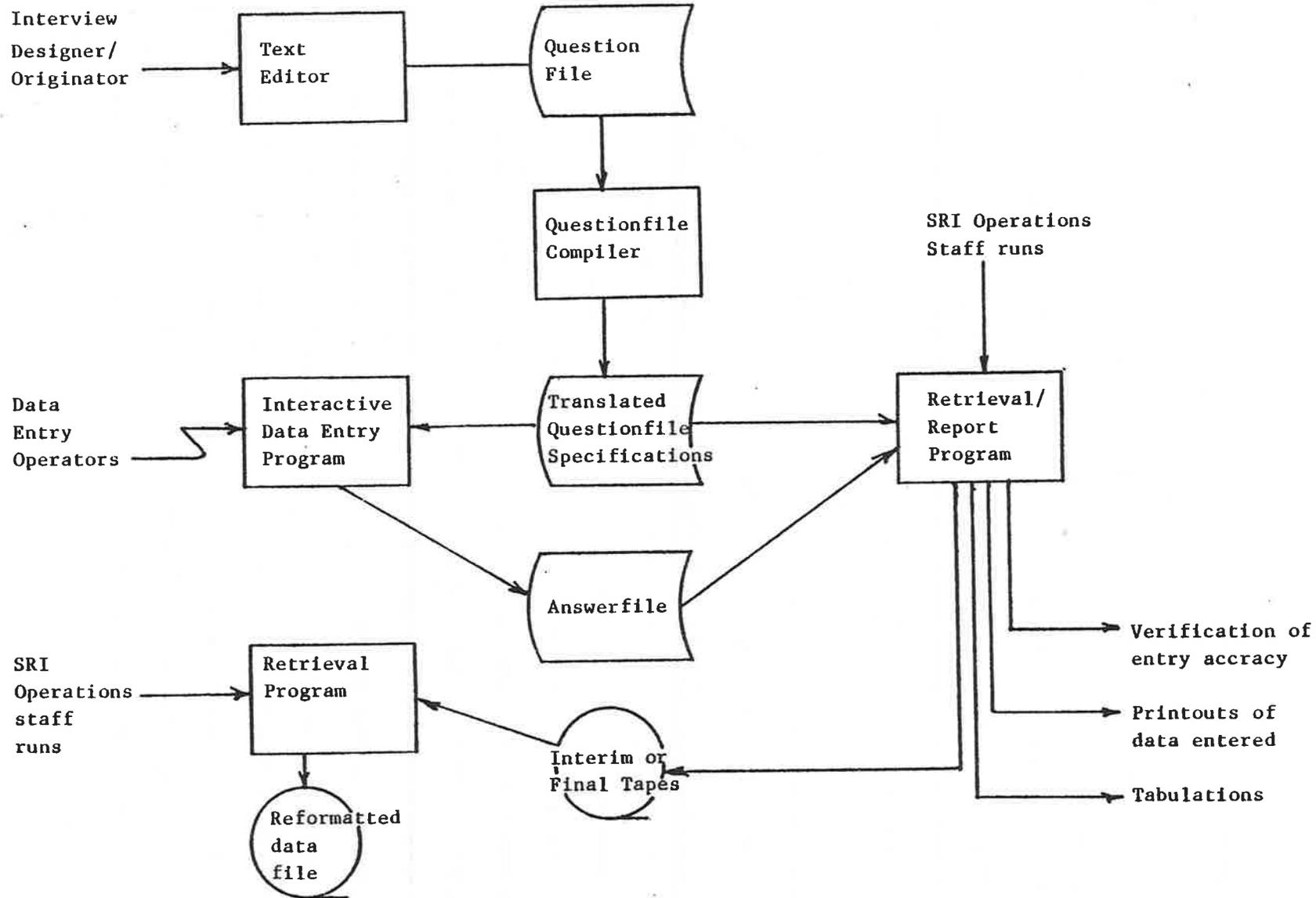
DATA ENTRY

The interview instrument is approximately 300 to 400 pages long. It contains hundreds of decision points where the interviewer must determine the appropriateness of administering, or repeating, a sequence of questions. The periodic interview has the potential for thousands of distinct answers, but on the average contains approximately 500 separate responses.

The problem of translating these interviews into machine-readable form and readying them for analysis was compounded by the anticipated variability of the interview modules, the existence of multiple sites, and hardware limitations. (At the time the experiment began, mini-computers were not as available and cheap as they are today and micro-computer-based products were unavailable.)

Since the complexity of the instrument seemed to preclude keypunching and optical scanning, data were entered interactively under control of a central computer. To do this, we developed the computer program dubbed the Electric Interview, with functional capabilities that are similar to key-to-disk entries. The specifications for the entry process (outlined in Figure 1) are defined in an interview-oriented language that allows the users to specify the prompt to be issued to the terminal, the type and limits on the response, whether an answer's boundaries may be overridden, and branching tests and sequencing instructions.

These specifications are developed by highly qualified research support



OUTLINE OF ELECTRIC INTERVIEW SYSTEM

FIGURE 1

personnel and tested prior to the official release for data entry.

Operators in Seattle or Denver offices then entered the data using hard copy, low-speed terminals connected to a fast, sophisticated central computer. The controlling Electric Interview program monitors the entry by (1) evaluating the answer for range or value acceptability, (2) performing limited inter-item checks, (3) controlling the rollback of the entry sequence to a point specified by the entry operator, (4) determining acceptability and changing previously entered value where necessary, (5) performing save and subsequent restart operations under user control, and (6) maintaining statistics on operator actions.

The data entry operator enters data directly from the interview booklet used in the field, which has experienced only a few internal adjustments and possibly some recoding during the quality control process, and therefore unlikely to have suffered major transformation errors. The program directs the entry operator through the complicated interview sequence while the operator is constantly verifying the range checks and sequence operations against the actual data. If conflicts occur, as where the operator is forced to skip recorded data or is unable to enter data, the process is stopped and the problem referred to the entry supervisor.

When the process reaches a successful completion, the result is a sequential stream of answers residing on the system as a single disk file. Each answer is recorded separately along with an indication of the answer type (text, interfer, date, real, Yes/No, etc.), the question number it is associated with, and whether it overrode the boundaries. Files are

periodically merged with their predecessors so that a complete file set exists on one tape, in addition to the current set of on-line disk files. (Sufficient sets of rotating dumps are preserved to insure the recreatability of any files retained on-line.)

During the course of an interview entry span of five months (one year for those who left the area) interview specifications may still be found to be in error, boundaries may be changed (typically widened), and anomalies may be discovered which cannot meet specifications. After reviewing a trouble report, program adjustments may be made to satisfy the first two situations. A new version of the interview is released only after the test cases have been repeated.

On completion of the entry process all files are merged and a complete machine generated inventory is produced and set to the field. These are compared against site records and, on frequent occasion, cause the interview to be reopened for additional entry.

DATA BASE EVOLUTION

The number of respondents (averaging 2500 per site), the average number of responses (500 per periodic interview), the amount of storage used to hold each answer and associated descriptive information (18 characters, or bytes), plus the header and control information on each interview's file, combined to yield approximately a full tape for each interview. In addition to the fullness of the individual data tape, we faced problems due to a sequential answer stream rather than a data-

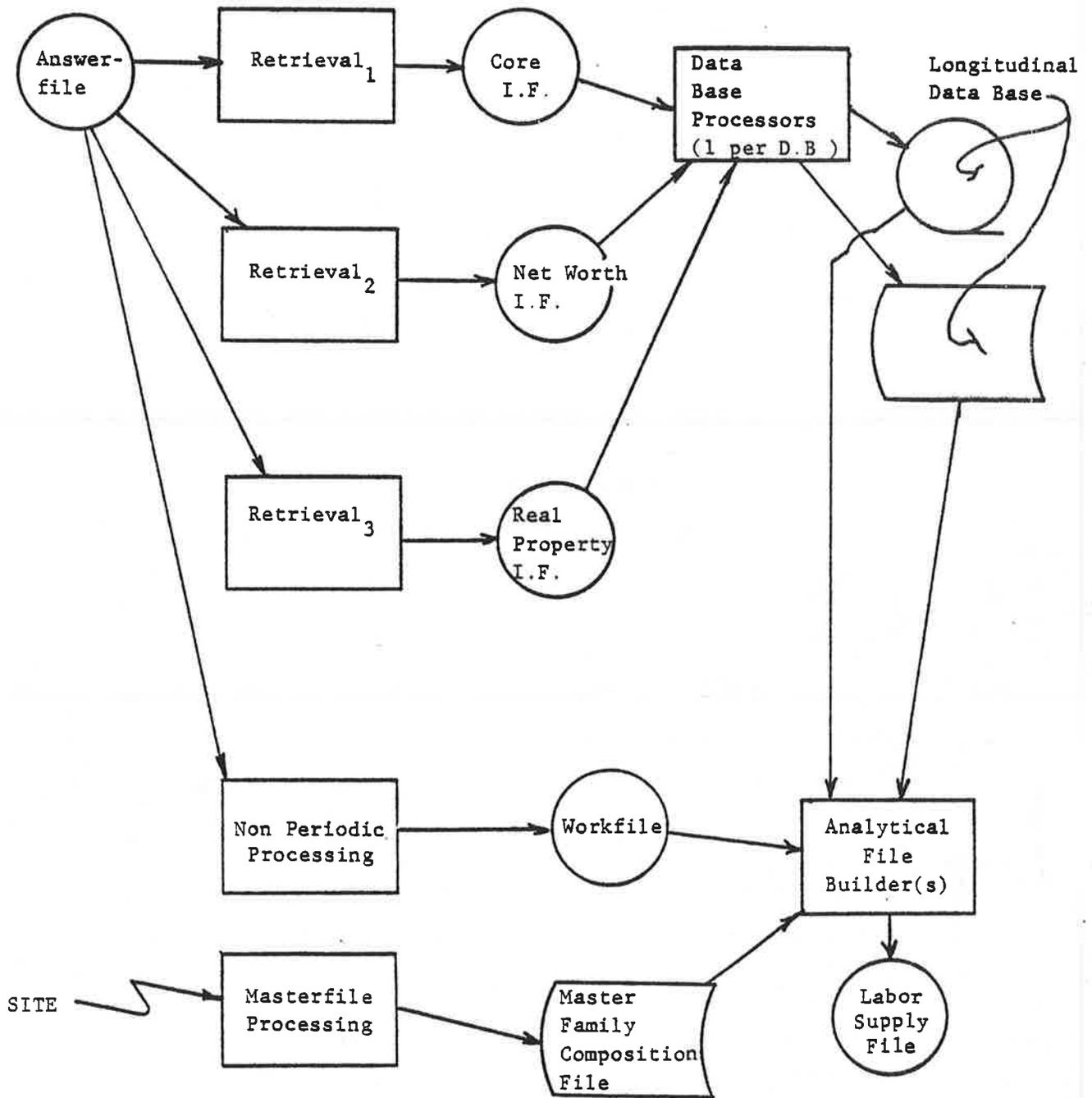
oriented structure. Also the data were difficult to use because they were physically separated into cross-sectional tapes, and data definitions were cross-sectional in nature. That is, the definition for a module's data in one tape was likely to be inapplicable to a tape containing a subsequent administration. As a result, the cost to retrieve and merge data was large and only experienced programmers could be used. In sum, the entry process solved the problem of entering complex data, but left it in a form that was incompatible with analytical access. The next step was to reestablish the inter-interview continuity inherent in the module approach of the questionnaire and make the data more accessible. Figure 2 shows an overview of this process.

Longitudinal Data

In late 1975, we decided that a more integrated and accessible data base should be developed to better represent the longitudinal nature of the data. Since we were continuously entering data and producing derivative files with data and computed variables, we had to avoid preempting or impacting existing procedures or products. As a result, we used the data entry process and the cross sectional file (Intermediate File) derived from the basic answer file as given. To develop the longitudinal files, initial emphasis was on the data derived from both core modules and non-core modules if given with some frequency (around once per year).

It was decided to build definitions using a descriptive form that was compatible with on-line data base systems in order to help impose a consistent methodology across different elements of the data base. We

FIGURE 2
ANSWERFILE PROCESSING



defined a consistent primary key to use on merging of the data. Design then proceeded on a component basis using the interview structure as the basis for the data base structure. Repeating sets of questions within an interview were reflected by repeating sub-records in the target data base. Nonrepeating items were collected into a periodic record for the appropriate information set. Thus, an information set would consist of a single periodic component and any number of repeating group components. For descriptive purposes the latter groups were referred to as transaction groups, that is events which could occur on an irregular basis within the interview and were, in essence, in effect until either the next occurrence of the event or the end of the information stream, whichever occurred first. Note that this definition allowed an event to span periodic boundaries. This removed many of the artificial events previously imposed by adhering to cross sectional interview tapes. Additionally, we were able to develop and apply linkage criteria for sub-components of the data base and therefore to link jobs and property across time.

In building this longitudinal data base, we also attempted to integrate previous processing and sources into a single descriptive document. This meant that we were required to relate the derived data base to the source document (interview), the data entry specification (logic which assigns an internal question number to the data), the data source (an intermediate file variable, record and location) and the raw questionnaire (a representative question was attached to each data item to reduce the need to go to the thirty plus collection instruments).

Aperiodic Data

Data that was collected on an irregular basis remained a problem. Rather than treat this problem in the way in which we treated the longitudinal data, we took advantage of the mechanized questionnaire as a basis for an automated retrieval and reformatting. The resulting file made the data more accessible, but the file structure and format differed for each interview processed, since it was driven by the mechanized version of the interview. This left us in the inter-periodic discontinuity in questions, compounded by discontinuities in the records.

As an initial step in removing these false discontinuities and integrating the data, we have constructed an on-line, data management description of the modules to be integrated. An interactive process prompts the building of the module one question at a time. A full description of the question, including that in the hard copy, is entered for each initial occurrence. Repetition in another interview requires only interview-unique data, for example, the local question number, and the item name assigned in the initial entry definition. It is our intent to use this descriptive information as a basis for establishing a longitudinal representation of the data and for reformatting it in a largely automated fashion.

Data Linkage

The continuity of some of the collected data across time is of considerable interest to researchers. Examples of this are jobs,

training and education programs and ownership of real properties. Unfortunately, the continuity is broken each time an interview is administered and it requires a data processing effort to reestablish it. The collected information is sometimes inadequate to unambiguously connect entities across a common boundary, such as residence, since real property addresses are not collected. Creating automated linkages is also hampered by the fact that some of the required information is of a textual nature that does not readily lend itself to precoding, such as employers. The problems are further compounded by variations in the collected data that are improper if the entities are truly the same. Thus, differences in initial (or final) dates or values imply discontinuities when they may actually represent only differences in recollections. The linkage process is forced to deal with this imprecision. The linkage criteria must be delicately balanced to be neither overly restrictive nor improperly aggressive.

In the recently developed data bases, we have been fairly conservative in linking entities across periodics, but have had reasonable success in linking jobs and properties. We may augment our automated linking of jobs by referring to the recorded interview's employer name as a further reference when a link is expected. (Job is denoted as "still working" or "previously working" at an interview's boundary.) In any event, we have some linkage in the data base, whereas they previously were established during the creation of analytical variables.

MASTER FAMILY COMPOSITION FILE

Once the data collection activities began in earnest, it became clear that the records keeping associated with tracking families (and their

individual adult members) was potentially overwhelming. Following individuals as they married, divorced, moved, and so forth was a near-impossible clerical task. Furthermore, without some form of automated support it was infeasible to perform certain checks, summaries or subset selections.

Therefore, an automated process was developed at a single computer site. The final result enabled us to (1) obtain a more logical structure to present the data; (2) set up a local clerical expert to address process or data anomalies; (3) virtually remove complicated process procedures from site consideration; (4) obtain ready research access to this important file; and (5) accomplish more sophisticated retrievals and edits.

While it is not surprising that we felt compelled to develop these monitoring and accounting capabilities, it was startling to learn how much it was needed and how difficult to do. The frequency and creativeness of family splits alone was sufficient cause to implement this facility. As a gross external indication, the number of families alone increased by 250% during the life of the experiment.

PROCEDURAL PROBLEMS

SIME/DIME was a difficult project, administratively, and operational decisions often made it difficult to maintain longitudinal consistency. The following gives a brief review of the problems that arise in such situations.

First, a site's periodic interviews did not exist as a series of interviews to be given on particular, common dates. Instead, the large population targeted for interviews required SIME/DIME to pursue two to three periodic

interviews at any point in time. Additionally, a parallel wave of longer duration but lesser magnitude, in the form of the Post Move/Residual interview, could overlap four (or more) periodic interviews per site. The problem was further compounded by the fact that the two site experiments were not synchronized. Figure 3 represents interviewing at both sites for one year, and illustrates the inherent timing problems.

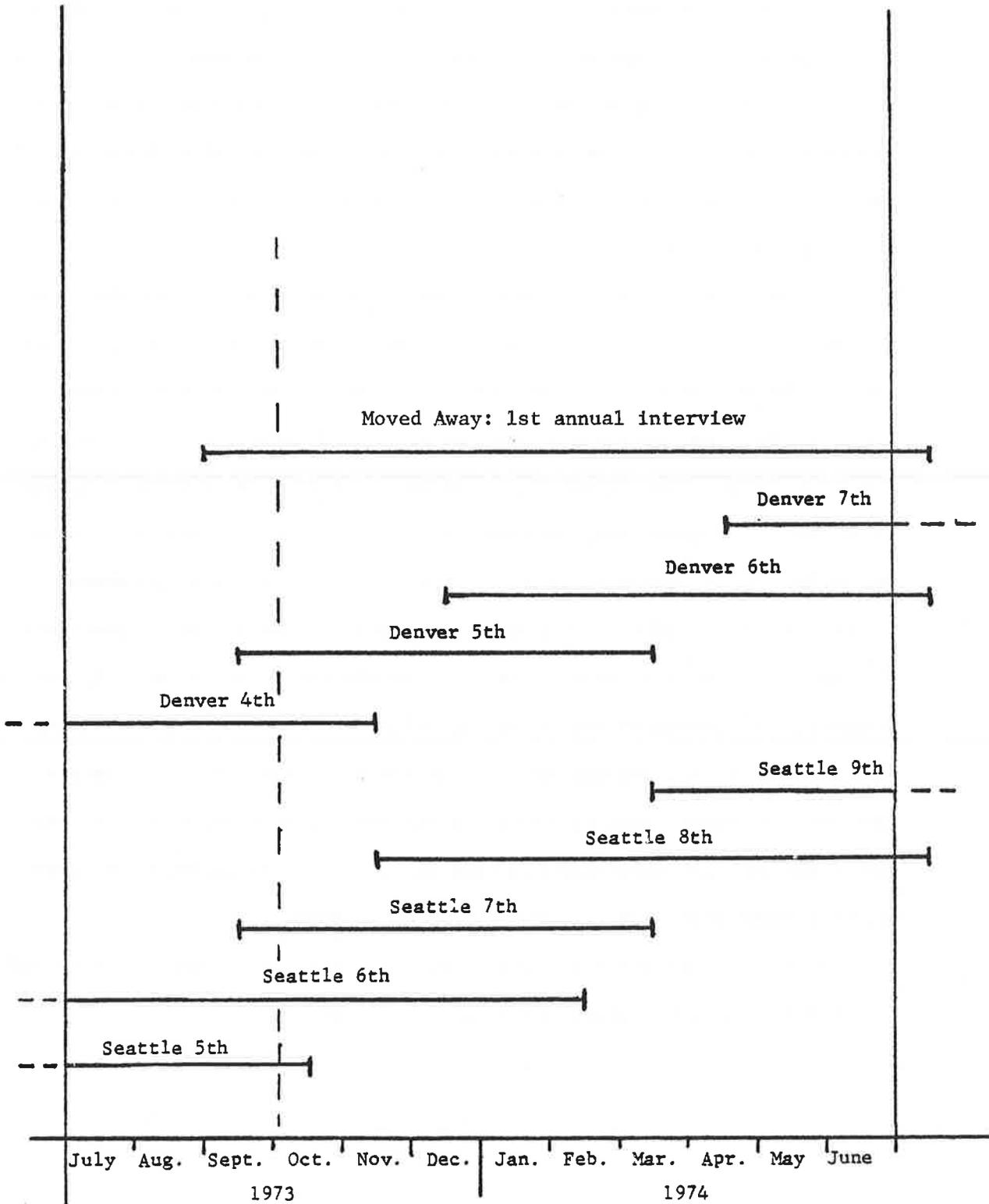
Throughout the experiment, and especially during the early interviews, question modules were revised. Thus at any particular time, several versions of a question would appear in the active interviews. These different versions were likely to be supported by different operational procedures and entry program components.

As time progressed, many efforts were made to control the operational and definitional contributions of the discrepancies in the modules. Agreement on the refinements and an increased emphasis on consistency promoted stability in the modules. Several new procedures also strengthened the data processing, increasing the continuity of the data and improving longitudinal representation.

At first data that did not meet specification caused difficulty to the whole process. Such data with unacceptable values was nonetheless there, although uncorrectable. The booklets containing this data tended to be deferred until the eleventh hour, when the restrictions would be relaxed sufficiently to let in all violators and a final burst of data would enter the system. This was a relatively costly solution considering the additional data gained, but more importantly, the unnecessary delay in a family's data entry could delay its subsequent interview.

FIGURE 3

INTERVIEWS ACTIVE IN A ONE YEAR CROSS-SECTION
(LAST HALF OF 1973/FIRST HALF OF 1974)



This procedure was finally replaced by a new process allowing the rejected data to be entered upon an approved request, as a privileged entry. The program is made available to a specific set of numbered families at one time. This version allows the entry operator to override any value that is remotely acceptable. Thus, for example, an hourly wage of \$100 could be entered but an individual identifier that failed the check digit test would still be restricted.

As the Electric Interview was developed to be used exclusively from a terminal, the testing process was labor intensive. This led us to restrict the testing of initial releases and forego testing of improved version. In early 1976 the program was modified to capture terminal input and accept input from disk files rather than terminals. This permitted construction of a library of basic test data that could be used at a simulated terminal. We could then (1) quickly run a full test of the entry program before initial installation and after any modification of an installed interview; (2) construct tests as readily as we could construct interviews; (3) perform a machine validation of the files resulting from the test, thus avoiding clerical oversight; (4) readily test the process through to the completed retrieval to insure that we would not discover file incompatible results after the complete interview had been entered; and (5) capture and rerun input streams which had caused the program to fail.

As a result of these changes, the installation of a new interview and the revision of an existing one became more routine.

EDITING ACTIVITIES

Certain editing activities contributed to the quality of the resultant data. These occurred at several points. It soon became clear that the interviewer could work better with prior knowledge of the last interview. Thus, family composition, training and education, job and net worth data were prerecorded to assist in verifying continuity or detecting discontinuity.

During the quality checking, both the staff and the machines review the interview for completeness and reasonableness. Additionally, control information filled in by the interviewer is checked. In addition, during data entry each answer must pass through an interview specific range, and some undergo limited checks against other answers. The path through the entry process is controlled so that the operator may not enter data for illogical paths.

It is our current opinion that the best opportunity for editing the collected data appears when we construct the longitudinal data base. We are now developing an editing procedure which, given the data base descriptions, can pass the file and construct longitudinal data lists for every item selected. It is our expectation that this will allow us to (1) develop time sequencing functions which can tie together items on different but overlapping time sequences; (2) develop a synchronization function to find corresponding values in different answers; (3) more easily develop item level checks, both range and inter-item, more easily since the code is operating upon single indexable items with no complex programming concerns; and (4) develop functions to evaluate reasonable item growth, fill in missing values or even override wild values.

Finally, we flagged the data (using a single digit) to show whether data contained longitudinal (inter-periodic) discrepancy, intra-periodic error (logic or range error) or modifications.

Analytical Edits

The major analytical file, the labor supply file, is edited for completeness and reasonableness as a final step in the production process. Apparently unreasonable values are retraced as far as the data base to insure that the implementation of the variable algorithms has not miscomputed a value. However, at this time there is no concerted effort to fill in missing values or correct anomalies. This probably should be undertaken soon, given the extent to which some researchers address this file.

Master File Edits

Once the Master Family Composition file was moved into Stanford Research Institute it was kept current on interview history, the creation and demise of families, and changes in family composition. Subsequently, in conjunction with the operational experts at the sites, an extensive set of edit operations were performed on this important file and it is in excellent condition.

Payments History Edit

Stanford Research Institute applies monthly consistency checks on an important file containing payments data and supporting information. Since this file was produced by a parallel process, we are attempting to verify that the content is consistent with the Master Family Composition File, and the payments per period agree with external accounting reports, and data values are reasonable and internally consistent.

LEARNED GUIDELINES

Many of the problems which we have encountered in SIME/DIME could have been avoided or alleviated by:

- insuring that the data processing effort imposes no artificial limit on the other data related activities -- particularly collection,
- insuring that the latest definition of data to be collected is consistent with the preceding definition(s) or the redefinition is necessary and sufficient,
- shortening the time from collection to entry,
- specifying the means by which the entities may be unambiguously identified as equivalent for linkage purposes,
- insuring sufficient redundancy in the critical data collected to make the interview more self sufficient,
- extending the entry edit process to use data collected previously, to determine the reasonableness of the latest data,
- specifying the way in which discrepancies can be resolved (not just recorded) as they occur, and
- providing machine generated material to be used by the staff in confirming the current entry activities and anticipating the next round of collection and verification.

Most of these recommendations appear attainable but we have had no success in shortening the time from collection to entry. This leads me

to suggest that some consideration should be given to moving the entry process closer to the interview -- thereby pre-empting or deferring some of the Q.C. activities. This recommendation presumes, of course, the existence of strong automated editing procedures.

VALIDATING SIME/DIME INCOME DATA

by

Harlan I. Halsey
Economist
SRI International

This paper describes the initial testing of the validity of the Seattle-Denver Income Maintenance Experiment (SIME/DIME) data, particularly the family income data. The success of any major social experiment such as SIME/DIME depends on accuracy of the data, the analytical method, and the sample size, among many things. The more accurate the data, the less stringent the demands in the other two areas. At this stage, of course, it is too late to affect the sample size, but knowledge of data quality can affect the analysis and interpretation of the results. Since the data will inevitably be found to be in error, the objectives of the validity study should be to assess the size of the bias, and the magnitude of the variance. This paper analyzes the accuracy of the SIME/DIME data derived from periodic interviews of subjects, conducted approximately three times a year; and subjects' income reports, made monthly or every four weeks.

Another purpose of the validity study is to shed light on the behavioral reporting phenomenon itself. Reporting is a facet of human behavior subject to influences both within the experiment and in society at large. Data are reported by individuals on many

occasions other than in these negative income tax (NIT) experiments. For example, data is self-reported on state and federal income tax forms; on the income transfer programs such as Aid to Families With Dependent Children (AFDC), the Food Stamp Program, and public housing programs; on social surveys such as the Survey of Economic Opportunity and the forthcoming Survey on Income and Program Participation; and in the Department of Labor's Current Population Survey. In fact most such data are reported by the individuals themselves. Notable exceptions include social security wage and salary data and unemployment insurance wage and salary data, which are reported by the employer. The accuracy with which data are reported and recorded affects public policy decisions on taxation, unemployment, inflation, and welfare, as well as our estimates of the income distribution and the unemployment rate.

For most self-reported data, either there is little incentive to hide the truth or the report can be easily checked. This is true, for example, of reports on the number and ages of children in a family. Income, however, is an exception. Income is often easily concealed. When self-reported, we can expect high variance and a downward bias due to such things as illicit income, and tax avoidance through under-reporting of income.¹ Recently Peter M. Gutmann has estimated that the fraction of gross national product associated with unreported transactions has risen steadily since World War II, and presently amounts to 9.4% of the GNP.² C. Northcote Parkinson has

maintained that government can tax only up to the cost of tax avoidance; that is, if it is cheaper to avoid the tax than to pay it, the individual will choose to avoid it.³ A small study of income reporting in AFDC, conducted on the DIME pre-enrollment data, indicated that families reporting income reported only 60% to the Denver Welfare Department, relative to that reported to DIME.⁴ An additional 24% of the families reporting income to DIME reported nothing at all to the welfare department.

POTENTIAL PROBLEMS WITH SELF-REPORTED INCOME

There are several special difficulties in recording and verifying income data, partly because income is highly taxed, providing incentives to under-report, and partly because income is held in extreme confidentiality, making verification difficult. The problem is compounded by the volume and nature of the data. Income is usually reduced to one or perhaps two variables for analysis, but it is actually composed of many individual payments from different sources. A validity study should concern itself with these individual payments because many different aggregations are useful for analysis and for comparison purposes and because individual income streams will be of particular interest.

Some corroborative sources only record data in the aggregate, as is done on federal personal income tax forms or unemployment insurance records. Here we are forced to aggregate in order to compare and we face possible timing error as a result. (A payment near the end of

a period may be reported in different periods by the recipient and the source.) The longer the period, the smaller this type of error becomes.

The taxation system further complicates the validation effort. Not only do different levels of government tax, and at different rates, the SIME/DIME experiments imposed taxes of 100%, and 50%, depending on type of non-wage income, and variable tax rates, ranging from 80% to less than 50% on wage and salary income. Allowable deductions also must be accounted for. All of these differing tax rates can have an impact on the incentives to report, and under-report, income. Another trouble spot is human memory. SIME/DIME interviews are conducted at intervals which stretch the memory of the respondent to recall data. (No doubt with sufficient incentives, respondents could recall income data over a period of years with high accuracy, but this would usually require the retrieval of records along with considerable mental effort.) Thus, an income stream whose payments are infrequent and which vary in size or which terminated some time ago is likely to be poorly reported relative to an income stream whose payments are current, closely spaced, and of equal size.

Family structure can also present problems. Some types of non-wage income, like AFDC, accrue to the family as a whole and not to any individual member. Further, when reporting income to a transfer agency, such as AFDC, a family may not report the current male head,

but he will be counted by SIME/DIME. This is one reason income reported to AFDC was not compared with SIME/DIME income in the validity study discussed below.

Finally, some public agencies do not retain data for long periods of time, or if they do, it is stored in ways which make it difficult to access as time passes, and regulations and definitions change making it difficult to define older data.

SUMMARY OF THE 1976 SIME/DIME VALIDITY STUDY

In 1975 and 1976 we conducted a validity study of the income data from the first two years of SIME and the first year of DIME. This section is a summary of the full report on this study.⁵

Sample Selection

We drew a 10% random sample--stratified by race and type of head, pre-experimental normal income, and financial treatment level--from the originally enrolled families for whom marital status of the head(s) remained unchanged from enrollment until January 1, 1973. In Seattle the sample contained 163 families; in Denver, 279 families. (Denver's sample is larger due to the inclusion of a Chicano group, and the earlier start in Seattle, with consequent greater attrition in this group, due to marital unions and dissolutions.)

Analysis of Wage Income

Since the purpose of this study was to validate the SIME/DIME data, our null hypothesis was that the public agency data is correct. The data were organized into three groups, or subsamples, consisting of

available data from the periodic interview, the IRF, and a public agency source. As controls do not prepare IRF's, two-way subsamples were also constructed for financial treatment families and control families separately. Within the experimental data, we assumed that the interview data is less affected by any direct incentive to misreport, so we adopted it as the standard of comparison. We were especially interested in discovering bias or large variations in reporting to SIME/DIME that could affect subsequent analysis of the data. Our results are shown in Table 1. The tests, of course, address only the question of differential reporting, and not absolute reporting accuracy. A family which consistently misreports income will not change the outcome of a validity study.

Since unemployment insurance is employer-reported, it is accurate except for employer misreporting, errors in the records of the unemployment insurance agency, and our occasional failure to find the proper record. These errors are probably small, so we believe the differences between unemployment insurance and SIME/DIME provide good measures of misreporting in the absolute.

The small mean differences between pairs of reports are overwhelmingly in the expected directions. That is, all but one of the Student's T-statistics are negative, indicating under-reporting to SIME/DIME relative to the public agency or under-reporting on the IRF relative to the periodic interview. The differences are statistically different in nine of the 30 comparisons. The largest and most

TABLE 1

WAGE EARNINGS T-TESTS AND CORRELATION COEFFICIENTS

Seattle 1971										
Income Type	FIN/ CONT	Sources	N	T	DIF	PA	%DIF	Stand. Error	Corr. Coeff.	
Gross Earnings (Annual)	3-WAY FIN	IRF-INT	25	-.94	-126	5643	-2.23%	135	.973	
		IRF-PA	25	-2.51*	-235	5643	-4.16%	94	.991	
		INT-PA	25	-.90	-109	5643	-1.93%	121	.983	
	2-WAY FIN	INT-PA	57	-.36	-45	6241	-0.72%	125	.963	
		2-WAY CONT	INT-PA	54	-.48	-91	5809	-1.57%	190	.915
Earnings Covered by Unemployment Insurance (Quarterly)	3-WAY FIN	IRF-INT	40	.10	5	1955	0.26%	51	.971	
		IRF-PA	40	-4.14*	-118	1955	-6.04%	28	.991	
		INT-PA	40	-2.60*	-123	1955	-6.29%	47	.976	
	2-WAY FIN	INT-PA	45	-1.52	-51	1579	-3.23%	34	.976	
		2-WAY CONT	INT-PA	44	-.51	-18	1436	-1.25%	34	.962
Seattle 1972										
Gross Earnings	3-WAY FIN	IRF-INT	43	-1.63	-213	7274	-2.93%	131	.968	
		IRF-PA	43	-2.72*	-307	7274	-4.22%	113	.977	
		INT-PA	43	-.66	-94	7274	-1.29%	144	.962	
	2-WAY FIN	INT-PA	47	-.82	-111	7407	-1.49%	136	.964	
		2-WAY CONT	INT-PA	45	-.12	-23	7229	-0.32%	182	.933
Earnings Covered by Unemployment Insurance (Quarterly)	3-WAY FIN	IRF-INT	53	-2.23*	-47	1450	-3.24%	21	.989	
		IRF-PA	53	-3.61*	-124	1450	-8.55%	34	.972	
		INT-PA	53	-1.85	-77	1450	-5.31%	42	.958	
	2-WAY FIN	INT-PA	53	-2.00	-81	1471	-5.51%	41	.961	
		2-WAY CONT	INT-PA	43	-2.39*	-84	1610	-5.22%	35	.966
Denver 1972										
Gross Earnings (Annual)	3-WAY FIN	IRF-INT	44	-1.37	-144	6740	-2.14%	105	.986	
		IRF-PA	44	-2.94*	-295	6740	-4.38%	100	.986	
		INT-PA	44	-1.11	-151	6740	-2.24%	136	.977	
	2-WAY FIN	INT-PA	107	-2.00*	-196	6915	-2.83%	98	.970	
		2-WAY CONT	INT-PA	75	-1.49	-214	7127	-3.00%	144	.931
Earnings Covered by Unemployment Insurance (Quarterly)	3-WAY FIN	IRF-INT	117	-.08	-2	1324	-.15%	27	.963	
		IRF-PA	117	-1.57	-46	1324	-3.47%	29	.956	
		INT-PA	117	-1.12	-43	1324	-3.25%	39	.922	
	2-WAY FIN	INT-PA	119	-1.15	-35	1308	-2.68%	30	.947	
		2-WAY CONT	INT-PA	70	-1.78	-100	1527	-6.55%	56	.893

Key: *Significance at .05 level

FIN = Financial Treatment Families

CONT = Controls

IRF = Income Report Form (Experimental families' monthly report to SIME/DIME)

INT = SIME/DIME Personal Interviews

PA = Public Agency Records

DIF = Difference between income as reported in the two sources

%DIF = DIF ÷ Public Agency Reports x 100%

statistically significant difference occurs between the IRF and the public agency reports, where differences are statistically significant at the 95% level of confidence in all cases, except for the Denver 1972 unemployment insurance/IRF difference. The extent of the under-reporting on the IRF, compared to unemployment office data, ranges between \$46 and \$307 per quarter. This is between -3.47% and -8.55% of the income reported to public agencies.

Income reported in the periodic interview is between .32% and 6.55% less than that reported to public agencies. The periodic interview/IRF differences range between +.26% and -3.24% of public agency reports. Relative to the public agency data the SIME/DIME data appears to be quite consistently reported in the mean. As expected periodic interview data is closer to the public agency data than the IRF data. There is no pattern among the T-statistic, the correlation coefficients, or the mean differences which suggests that the SIME/DIME controls reported differently than financial treatment families.

We also tested the possibility that the difference in reports could be related to the amount of income. Under the assumption that families never over-report, the more income, the larger the potential difference. Such a relationship could undermine the validity of the T-tests. To test for income under-reporting we regressed the difference in reporting on the amount of income reported to the public agency, a dummy variable distinguishing dual-headed families; and in the

combined control and financial treatment samples, a dummy variable distinguishing control families from financial treatment families, and the interaction of the control dummy with the earnings variable. No strong indication of a relationship between income level and under-reporting was found.

The coefficients were generally of the expected sign and within a reasonable range, but few were statistically significantly different from zero. The magnitude of the earnings coefficients was between +6% and -14%, with negative values indicating increasing under-reporting to the experiment relative to the public agency. The more negative earnings coefficients are always offset by positive constants, however, so that there would appear to be over-reporting at low incomes whenever there is large under-reporting at high income. The sample sizes are so small and error distributions were such that more sophisticated specifications of the regression produced unintelligible results.

In general, the regression results do not contradict our previous suppositions, nor do they support them, probably because of the small sample sizes. The largest sample sizes are in Denver, and the regressions run on these samples might indicate a tendency for controls to under-report more income to DIME as income rose, but is difficult to believe that low-income families over-report as much as the positive control dummy indicates. It is more likely that either the public agency data are not free from error, as we had assumed, or

that this is a small sample effect.

We have also conducted some preliminary analysis of the full sample, comparing earnings reported to SIME and to IRS by couples. Again, reporting is quite accurate in the mean, but individual variation is great. Figure 1 contains a point for every family headed by a couple and remaining intact from enrollment through 1972 in Seattle. As can be seen, there is considerable variance in income reporting to SIME. This must be taken into account by the analyst.

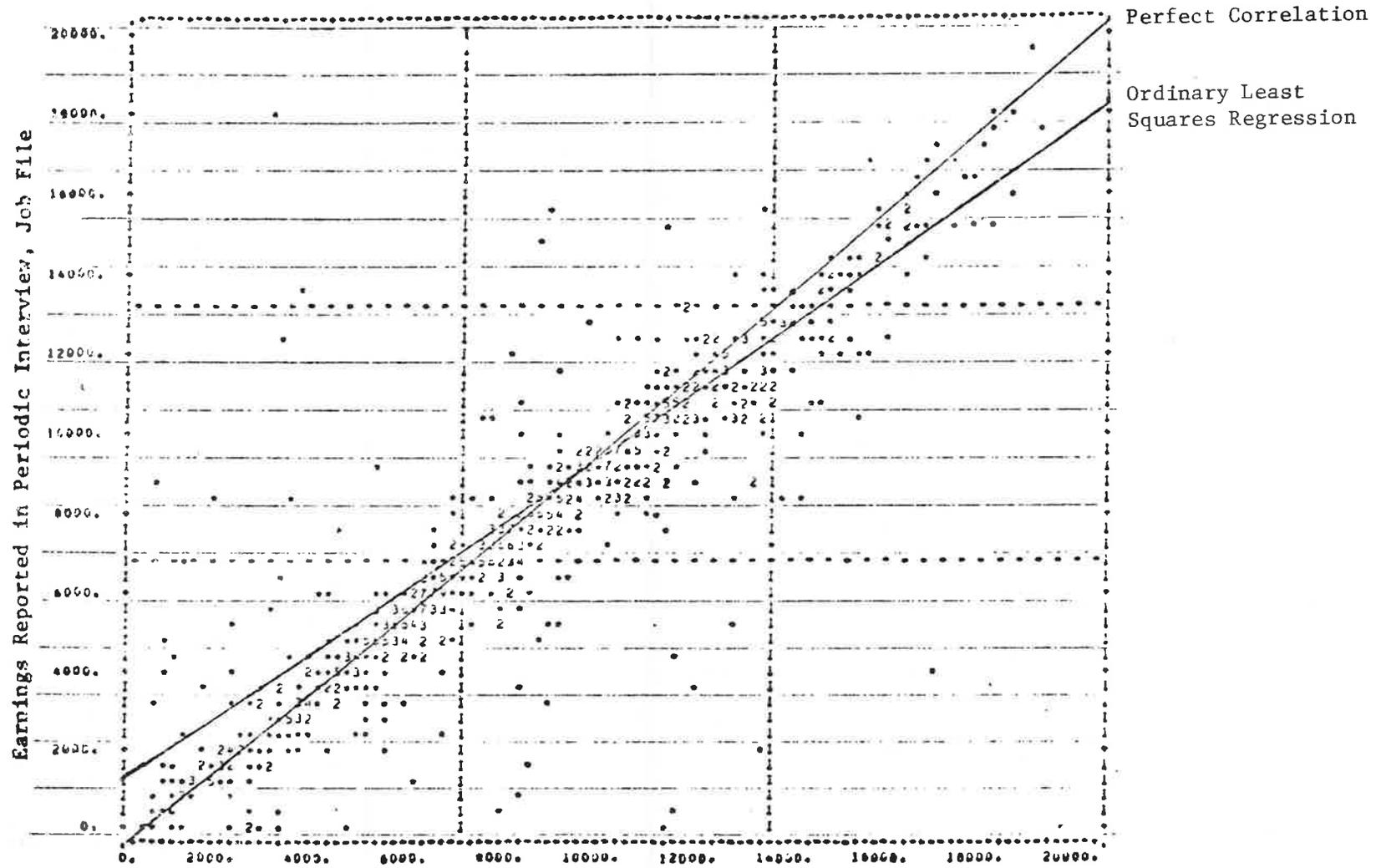
Analysis of Non-wage Income

We have also analyzed three non-wage income sources: the AFDC grant, unemployment insurance, and the Food Stamps Program; and one expense, rent in public housing. SIME/DIME taxed non-wage income from public sources at 100% and reimbursed changes in the rent in public housing by 100%. Because of these taxation incentives, we expect to see non-wage income under-reported and rent over-reported on the IRF, and to a lesser degree on the periodic interview. Here, the public agencies are the source of the grant, benefit, and rent amounts, and the issue of under-reporting does not arise.

All data was aggregated into calendar quarters and student's T-statistics of the differences in reports were constructed. The T-statistics are reported on Tables 2, 3 and 4 for Seattle, 1971 and Denver 1972 respectively. Generally speaking, the differences as a fraction of mean income or expense are much larger than was found for wage income, and wherever the T-statistic is statistically significant, it is in the expected direction.

FIGURE 1

EARNED INCOME COMPARISONS, SEATTLE, 1972



Earnings Reported to the Internal Revenue Service, Form 1040

TABLE 2

SEATTLE 1971 NONWAGE INCOME AND EXPENSES

Income Type	FIN/ CONT	Sources	N	T	DIF	PA	%DIF	Stand. Error	Corr. Coeff.
AFDC Grant (Quarterly)	3-WAY FIN	IRF-INT	22	1.41	61	452	13.50%	44	.775
		IRF-PA	22	-2.22*	-86	452	-19.03%	39	.679
		INT-PA	22	-2.71*	-147	452	-32.52%	54	.615
	2-WAY-FIN	INT-PA	27	-2.68*	-105	441	-23.81%	39	.694
	2-WAY CONT	INT-PA	33	-3.10*	-135	490	-27.55%	44	.555
Unemployment Benefits (Quarterly)	3-WAY FIN	IRF-INT	16	-.23	-14	422	-3.32%	63	.528
		IRF-PA	16	-2.78*	-140	422	-33.18%	50	.699
		INT-PA	16	-1.89	-126	422	-29.86%	67	.421
	2-WAY FIN	INT-PA	24	-2.53*	-158	506	-31.23%	62	.296
	2-WAY CONT	INT-PA	22	-1.34	-76	498	-15.26%	57	.565
Food Stamps Net Benefit (Quarterly)	3-WAY-FIN	IRF-INT	26	.34	2	43	4.65%	6	.937
		IRF-PA	26	2.29*	37	43	86.05%	16	.326
		INT-PA	26	2.14*	35	43	81.40%	17	.322
	2-WAY FIN	INT-PA	39	3.80*	58	39	148.72%	15	.056
	2-WAY CONT	INT-PA	40	2.30*	28	59	47.46%	12	.073
Food Stamps Net Benefit Excluding cases where PA reported 0 (Quarterly)	3 WAY FIN	IRF-INT	17	1.55	5	72	6.94%	3	.982
		IRF-PA	17	.55	3	72	4.17%	6	.928
		INT-PA	17	-.21	-2	72	-2.78%	7	.897
	2-WAY FIN	INT-PA	24	.52	7	76	9.21%	13	.555
	2-WAY CONT	INT-PA	26	-1.90	-17	95	-17.89%	9	.589
Public Housing Expense (Annual)	2-WAY FIN	INT-PA	17	1.78	121	406	29.80%	68	.575
	2-WAY CONT	INT-PA	20	1.10	154	464	33.19%	140	-.185
Public Housing Expense Excluding cases where PA reported 0 (Annual)	2-WAY FIN	INT-PA	14	.70	31	493	6.29%	45	.836
	2-WAY CONT	INT-PA	15	-1.00	-115	618	-18.61%	114	.259

Key: See Table 1.

TABLE 3

SEATTLE 1972 NONWAGE INCOME AND EXPENSES

<u>Income Type</u>	<u>CONT</u>	<u>Sources</u>	<u>N</u>	<u>T</u>	<u>DIF</u>	<u>PA</u>	<u>% DIF</u>	<u>Stand. Error</u>	<u>Corr. Coeff.</u>
AFDC Grant (Quarterly)	3-WAY FIN	IRF-INT	10	-.32	-23	448	-5.13%	70	.731
		IRF-PA	10	-1.49	-139	448	-31.03%	93	.381
		INT-PA	10	-1.45	-117	448	-26.12%	80	.590
	2-WAY FIN	INT-PA	9	-1.46	-130	498	-26.10%	89	.493
	2-WAY CONT	INT-PA	26	-3.03*	-150	561	-26.74%	50	.687
Unemployment Benefits (Quarterly)	3-WAY FIN	IRF-INT	15	.09	4	304	1.32%	50	.541
		IRF-PA	15	-2.07	-132	304	-43.42%	64	.277
		INT-PA	15	-1.80	-136	304	-44.74%	76	.079
	2-WAY FIN	INT-PA	15	-1.80	-136	304	-44.74%	76	.079
	2-WAY CONT	INT-PA	17	1.34	99	188	52.66%	74	.284
Food Stamps (Quarterly) Net Benefit (Quarterly)	3-WAY FIN	IRF-INT	22	1.28	12	73	16.44%	9	.904
		IRF-PA	22	.28	8	73	10.96%	28	-.091
		INT-PA	22	-.16	-4	73	-5.48%	25	.054
	2-WAY FIN	INT-PA	21	-.08	-2	75	-2.67%	26	-.007
	2-WAY CONT	INT-PA	34	1.33	18	65	27.69%	13	.154
Food Stamps Net Benefit Excluding cases where PA Reported 0 (Quarterly)	3-WAY FIN	IRF-INT	16	.53	2	101	1.98%	5	.966
		IRF-PA	16	-2.52*	-44	101	-43.56%	18	.410
		INT-PA	16	-2.64*	-47	101	-46.53%	18	.463
	2-WAY FIN	INT-PA	16	-2.53*	-45	99	-45.45%	18	.441
	2-WAY CONT	INT-PA	25	-2.26*	-20	92	-21.74%	9	.638
Public Housing Expense (Annual)	2-WAY FIN	INT-PA	17	2.37*	260	286	90.91%	110	.158
	2-WAY CONT	INT-PA	24	.74	82	464	17.67%	112	.161
Public Housing Expense (Excluding cases where PA reported 0 (Annual)	2-WAY FIN	INT-PA	12	.45	37	405	9.14%	81	.675
	2-WAY CONT	INT-PA	17	-.63	-81	655	-12.37%	127	.137

Key: See Table 1.

TABLE 4

DENVER 1972 NONWAGE INCOME AND EXPENSES

Income Type	FIN/CONT	Sources	N	T	DIF	PA	%DIF	Error	Coeff.
AFDC Grant (Quarterly)	3-WAY FIN	IRF-INT	61	1.12	17	408	4.17%	16	.867
		IRF-PA	61	.24	7	408	1.72%	27	.594
		INT-PA	61	-.37	-11	408	-2.70%	29	.563
	2-WAY FIN	INT-PA	64	-.29	-9	412	-2.18%	30	.489
	2-WAY CONT	INT-PA	37	-.91	-42	430	-9.77%	46	.434
Unemployment Benefits (Quarterly)	3-WAY FIN	IRF-INT	4	-1.13	-80	141	-56.74%	71	.905
		IRF-PA	4	.05	6	141	4.26%	115	-.161
		INT-PA	4	.61	85	141	60.28%	139	.212
	2-WAY FIN	INT-PA	6	1.41	217	94	230.85%	153	-.118
	2-WAY CONT	INT-PA	2	0	285	0	-	0	0
Food Stamps Net Benefit (Quarterly)	3-WAY FIN	IRF-INT	59	.40	3	85	3.53%	6	.750
		IRF-PA	59	1.58	17	85	20.00%	11	.275
		INT-PA	59	1.56	14	85	16.47%	9	.434
	2-WAY FIN	INT-PA	61	1.67	15	84	17.86%	9	.377
	2-WAY CONT	INT-PA	40	-2.33*	-29	119	-24.37%	12	.376
Food Stamps Net Benefit Excluding cases where PA Reported 0 (Quarterly)	3-WAY FIN	IRF-INT	47	.15	1	116	0.86%	6	.840
		IRF-PA	47	-2.15*	-16	116	-13.79%	8	.671
		INT-PA	47	-2.40*	-17	116	-14.66%	7	.726
	2-WAY FIN	INT-PA	50	-2.81*	-18	112	-16.07%	6	.748
	2-WAY CONT	INT-PA	38	-3.94*	-41	128	-32.03%	10	.582
Public Housing Expense (Annual)	2-WAY FIN	INT-PA	23	1.53	147	451	32.59%	96	.355
	2-WAY CONT	INT-PA	9	.80	164	352	46.59%	205	-.197
Public Housing Expense Excluding cases where PA reported 0 (Annual)	2-WAY FIN	INT-PA	14	-.95	-90	740	-12.16%	95	.440
	2-WAY CONT	INT-PA	5	-.49	-135	634	-21.29%	274	-.344

Key: See Table 1.

The AFDC grant is under-reported to SIME by relatively large amounts. As a fraction of the mean AFDC grant, the under-reporting ranges from 19 to 33% in 1971 and from 26 to 31% in 1972. The differences are statistically significant at the 95% confidence level in 1971 but not in 1972, where sample size is quite small, except for one, the two-way comparison of interview and public agency reports. The situation is different in Denver. In spite of larger sample sizes, the reporting difference ranges from +4 to -10% and none of the T-statistics are significantly different from zero. Both the mean differences and their standard errors are much smaller in Denver than in Seattle, indicating more accurate reporting in Denver. Hopefully, this reflects the fact that DIME started a year after SIME and benefitted from the Seattle experience.

The IRF-interview differences are never statistically significant in either site, but the mean difference drops dramatically between 1971 and 1972 in Seattle. This may indicate the overcoming of difficulties attending the start-up of the experiment. Evidently AFDC grants are relatively consistently reported to the experiment. The same consistency is observed for unemployment benefits and food stamp benefits, except for the Denver 1972 unemployment benefit where the IRF appears to be accurate and the interview data appear to be 60% over-reported.

Unemployment benefits appear to be under-reported in Seattle, while the Denver sample size is too small to be useful. Mean under-

reporting to SIME ranges between 15 and 45%. In 1971 financial families under-reporting is statistically significant, but controls under-reporting is not. In 1972 financial families under-reporting is just below statistical significance, while controls over-report in the mean but not significantly so.

Food stamps benefits at first appeared to be over-reported to the experiment by large amounts, but were relatively consistently reported between the IRF and the interview. Investigation indicated that many official food stamp records have been lost. The statistics on food stamps were reconstructed excluding cases in which the food stamps record indicated zero face value of food stamps purchased. (These statistics are presented in the tables below the first set of results.) This restricted sample provides evidence of either accurate or of under-reporting, as expected. In the 1972 results, the mean differences are negative and significant, ranging between 21 and 47% in Seattle and 14 and 32% in Denver.

The analysis of rent in public housing reports indicated the expected tendency toward over-reporting. Although over-reporting was the expected response to the experimental incentive for financial treatment families, the quality of the public housing data was again suspect, so the analysis was rerun for the sample for whom the public housing authorities reported positive rent paid. In the restricted sample, none of the T-statistics are significant, and the mean differences are smaller, though still positive, for the financial

treatment families in Seattle. In Denver, the differences, while never statistically different from zero, switch from positive to negative.

SUMMARY

Earned income reported to the experiment is under-reported in amounts between \$100 to \$300 per year relative to income reported to public agencies, and the amount is statistically significant. However, it amounts to less than 5% of mean earned income. The variance in the error in annual earned income reported to SIME/DIME is on the order of \$1,000 and somewhat surprisingly, this variance does not seem to depend on the income level.

AFDC and unemployment benefits were under-reported by \$90 to \$160 in Seattle in 1971, but differences could not be detected in 1972 in Seattle or Denver. Non-wage income is reported with a higher variance, as a proportion of mean non-wage income, than is earned income.

The wide variance in reported income is large enough to cause economically significant errors in parameter estimates in ordinary least squares regressions, so the analyst must use care in interpreting such estimates, or use more sophisticated methods.

Finally, for purposes of future validation studies, we have learned that we need larger sample sizes to study non-wage income, as the frequency of such income is low, and variation in differences between reports is relatively large. We have also learned that public files must be carefully checked for accuracy and completeness.

NOTES

1. Christian G. Hill, "Waitresses," Wall Street Journal, April 5, 1978, p. 1; Alfred Malabre, Jr., "The Outlook," Wall Street Journal, March 26, 1978, p. 1.
2. Peter M. Gutmann, "The Subterranean Economy," Financial Analysis Journal, (November/December 1977).
3. C. Northcote Parkinson, The Law and Profits, (New York: Ballantine Books 1960).
4. Harlan Halsey, et al., "The Reporting of Income to Welfare: A Study in the Accuracy of Income Reporting," (Center for the Study of Welfare Policy, Stanford Research Institute, Research Memorandum 42, August 1977).
5. The full report is found in Harlan Halsey, Bina Murarka, and Robert G. Spiegelman, "The Seattle and Denver Validation Study," (Draft Research Memorandum, June 1976, Stanford Research Institute, Menlo Park, California).

DISCUSSION

JAMES WALSH: I want to make sure I understood your variance statement. Did you say a significant number of people over-reported their income?

HALSEY: Yes. Looking at that graph with all of the points scattered on it, there are as many above the line as below the line. Using the Form 1040 income for comparison, there is a significant number that appeared to over-report.

HENRY LAHEWSKI: I am interested that you used IRS records. Did the IRS provide you with a record of specific individuals who were in the experiment?

HALSEY: Yes. We asked each family to sign a release for the IRS records.

ROBERT SPIEGELMAN: One thing you didn't point out is that the extent of under-reporting does not differ for experimentals and controls; and therefore, the under-reporting would probably not introduce any bias to the analysis.

SIME/DIME DATA FILES
AVAILABLE TO THE PUBLIC

by

Constance F. Citro
Deputy Director
Policy Studies Division, MPR

Miriam Aiken
Researcher, MPR

Kristen Puckett
Researcher, MPR

Although much of the data collected by the Seattle-Denver Income Maintenance Experiment (SIME/DIME) is still being processed, there are several longitudinal data files now available at moderate cost.¹ As we are encouraging policy analysts in government and academia to use them, this paper reviews their contents and applications. Called SIME10TH and DIME10TH, they include a version organized around families, with subfiles for members over 16 (the family-person version) and one organized around individuals who were heads of households at any time during the experiment (the composite principal person version). Both contain month-by-month data for 48 months for all families in the experiments.

CONTENTS AND FORMAT

The files contain key demographic and socioeconomic variables, including number of persons in the family and their age, race, and sex; family income from public and private transfers; family expenses for health care and support of dependents; liquid assets; homes and automobiles owned;

also, earnings and unemployment benefits, hours worked in all jobs and in the jobs with the longest and next longest hours in each month, regular and overtime wage rates, occupation and industry, spells of voluntary and involuntary unemployment, educational background and marital status for all persons 16 and over.

These variables were selected so as to permit studies of welfare participation. By no means all of the variables in the data base were put on the monthly files--items on job satisfaction and job search behavior, notably, were left out. Moreover, only very few of the many socio-psychological variables collected in the sets of questions appearing only sporadically (the "non-core" modules) were included. Nevertheless, a wide range of variables are contained on the records, which are nearly 16,000 characters long in the original version prepared for DHEW.² Each file covers four years, including the year covered by the enrollment interview through the period covered for all families by the tenth periodic interview (therefore named SIME/DIME10TH), or January, 1970 to December, 1973 in Seattle and January, 1971 through December, 1974 in Denver. By spring of 1979, another two years of data (through the sixteenth periodic interview) will be included.

Family-Person Format

The family-person version of SIME/DIME10TH reflects the organization of the data base records and permits identification of families and individual persons in the field. The family-person files contain a record for each family that ever participated in the experiments during the

four-year time span, followed by a record for each person 16 years and over who was ever a member of that family.

Figure 1 shows partial dumps for a set of records for one of the families in the file. This family is number 94804906. The item marked "G-level" is the "group level" code; in the figure, it is indicating a two-adult Chicano family. The Family Status Code is 1, repeated twelve times, for the first 12 months. This tells us that this family was one of those enrolled at the outset of the experiment; the code of 2 in the remaining 36 months indicates the family became a financial treatment family eligible for payments. As shown on the figure, the family had four persons present for all 48 months. Other information in the file gives details on the family members and on unearned income. This family had no other transfer income. There was no DIME payment in the first year, but after month 13 it was at least \$20 every month, and usually more.

The heads of this family each have a person record. The male head is person 48049028 (relationship code 01) and the female head is person 48049036 (relationship code 02). Both were present in the family for the full 48 months (indicated in the Person Status Code by a value of 02 in month 1 followed by a string of 04's). The husband was steadily employed at one (or sometimes two) jobs in 1971 and 1972, but experienced some spells of unemployment at the end of 1973 and 1974. The wife was unemployed throughout. This is a straightforward situation.

In contrast, figure 2 shows a white husband-wife family (the G-level is 4), number 94209307, that separated in month 28. The family looks very

much like the first family for those 28 months. There were four persons present, husband, wife, son and daughter. The Family Status Code shows a string of 1's (16, for 16 months prior to enrollment), followed by a string of 2's indicating a financial treatment from months 17 to 28. The male head, person 42093017, had a job (sometimes two or three) in most months, and the female head, person 42093025, had a job throughout. However, following month 28, all of the monthly demographic, income, and employment fields are filled with 9's and the Family Status Code is 3, followed by a string of 4's, indicating that this family unit was no longer intact. The Family End Code, which was 00 for family in figure 1, is 01 here, as the heads split.

Such changes in family composition were expected and were, in fact, a prime research interest. Thus, a system was set up at the outset to record a change in the number of heads; the old family number was "retired" and one or more new family numbers assigned to the individual persons involved. Hence, the file on family 94209307 was replaced by files on two families--with new numbers for each adult head, 98196900 and 98196803.

Figure 3 shows these new families, Family 98196803 contains the female head. The person number remains the same--42093025. She has with her the daughter from the previous marriage. The G-level is 3 for a single-head white family and the Family Formation Code of 01 indicates that this family started because the heads split in the previous family. The new family still had a DIME payment in most months, and the woman herself still had a job. The string of 9's represent months prior to the start of

this family. The Family Status Code is 0 for these months and then 2 for months 29-48. Similarly, the male head is in a new family, 98196900, with the son from the previous marriage.

The family-person structure lends itself very well to longitudinal studies on a variety of topics for families that did not change marital status at any time in the four years. Of about 2,700 families originally enrolled in DIME, over 1,600 continued with the same single parent or with the same couple for the entire four years. Restricting the number of years examined increases the sample size available--over 2,000 family heads remained the same in any one year of the four. Other changes in composition can be rich areas of study--birth of new children, the exit of teenage children who set up their own families, and so on. Of course, the files are also suitable for point-in-time studies of the entire sample.

The family-person files become harder to work with when the researcher wants to examine families that changed status. Tracing a family requires searching through the files to find all of the records for every other family unit made up of members of the original family. In our example, one must link the original family (94209307) to families 89196900 and 98196803 to have a continuous stream of information for the male and female head, respectively. The composite files were designed to permit this linkage.

Composite Principal Person Format

The composite principal person SIME/DIME10TH files were built subsequently from the family-person records. The composite version is organized

around those persons who were heads of a SIME/DIME family at any time during the experiment. For each such person, there is one record type that contains a continuous stream of employment and earnings data for all 48 months that was constructed by linking the records for that person to each family to which the person belonged in the family-person version. Following the principal person's own record is a composite family record that links data on demographic characteristics, benefits, expenses and assets of the families of which this person was a member. If this person was ever married, there is also a composite record of employment and earnings on other adults who were co-heads with this person (spouse or spouses). Last, if other adults were ever members of this person's family, but not a head, there is a somewhat shorter composite record combining job and earnings data for all such tertiary persons.

To demonstrate what is happening in the composite files, figure 4 shows the composite file of the female head (person number 42093025) who split up with her husband in month 28. Her own record (type 1) shows a continuous stream of job data (instead of the string of 9's appearing on the family-person file). Her composite family record (type 2) shows continuous presence in the experiment (family status is 1, followed by a string of 2's all the way through month 48) and a continuous stream of DIME grants. For the first 28 months, these went to the husband-wife family. Several variables trace the changes in family structure that she experienced: the family indicator is 1 from months 1 through 28, and 2 from months 29 through 48, documenting the change in families; likewise, the

G-level, which is now a monthly variable, is first 4, then 3. The family has four members the first 28 months and then two. Finally, the composite record for her spouse (type 3) shows that she had one spouse in months 1-28 who was usually employed during that period; the remaining months are filled with 989898's, indicating there was no spouse present, but the principal person was still active in the experiment. There is no tertiary record (type 4), as this woman had no other adults in her family.

The composite file structure, with four record types, may appear complicated. However, briefer versions can readily be extracted that merge selected data on the family, spouse, and tertiary persons with the principal person's information as required for a particular application. These composite files facilitate a wide range of longitudinal studies using persons as the unit of analysis. They should not be tabulated by families, as doublecounting would occur.

These files make possible studies of women who experience separation or divorce and subsequently head a single-parent family; of young adults who leave their parental families to set up their own homes; of patterns of labor force participation and job search for prime-age male heads regardless of how many families they belong to over time, and so on.

DOCUMENTATION

There are three major pieces of documentation for use in working with the monthly files: (1) a Record Format Description detailing the arrangement of the data on the tapes, (2) a Glossary of Variable Definitions, and (3) a printout of the distribution of responses to each item produced with a computer program called STATS.

Record Format Description

The Record Format Description (RFD) is the basic tool for getting at the data on the files, whether using a package or a specially written program. Each of the four monthly files has its own RFD that describes the technical characteristics of the tape recording, such as number of reels, number and sequence of logical records, blocking factor, etc., and then notes the arrangement of the variables in each record type. For each variable, the RFD provides a unique variable number, a unique variable name or mnemonic, the character position and length of the variable in the record, and a brief description of the variable--or, more properly, a variable title--plus a listing of codes for coded variables.

The variable number is in ascending order from the beginning of each family-person or composite file record type. This number can be looked up in the Glossary to obtain more information about a particular variable. The variable name, or mnemonic, is no more than eight characters long and can be used in statistical packages like SPSS. The field position indicates the location of the variable on the file and can be cross-referenced to the STATS printout. The length is the number of characters occupied by each field. The variable description provides a brief statement about the variable and, where applicable, codes are listed. These descriptions are not intended to be definitive and the user must consult the Glossary for a complete definition.

Glossary

The Glossary provides detailed definitions of the variables on the monthly files. There is one Glossary for the family-person files and a

second one for the composite principal person files. Each Glossary is organized by record type (family and person, or principal person, family, spouse, and tertiary, respectively), and, within record type, by broad categories of variables (such as benefits and expenses or employment history). Introductory material for each section indicates generally how that class of data was collected in the experiments and extracted for the monthly files. It also highlights inconsistencies in wording of questions across periodics and the use of codes to indicate the absence of data. There is also a listing for each variable organized by the RFD variable number.

The variable listings repeat the mnemonic name and variable title from the RFD and indicate all allowable codes, including those where no response was applicable (the N.A. codes). The listing then gives the precise question wording that was used for the item in the interviews. Any wording changes that occurred from one periodic to another are detailed. The Glossary also describes interviewer instruction or probes contained in the interviewer booklets. For variables on the monthly files that were derived by combining answers from more than one question, by recoding, or by other manipulations, an explanation is furnished of the procedures followed. Finally, there is a "comments" section, where site-specific qualifications about the variable may appear, or general warnings about the quality of the response, cautions about missing data codes, or references to other related variables.

The Glossary provides the description of each variable needed by users to make valid inferences from the data and understand both their applicability and their limitations.

STATS Printout

The STATS Program provides information about the distribution of responses to individual variables. STATS runs are available for each record type in the SIME10TH and DIME10TH family-person and composite principal person files.

For each variable, the program lists the mnemonic name and character position, as in the RFD, listing separately each monthly observation for an item such as the SIME or DIME grant. For each field, the program calculates and displays the range, mean and standard deviation (of values which are more than 0, and excluding the N.A. code values), and the number of families with responses in this range. Another column shows the number of families with a value of zero; and another shows each possible N.A. code.

For continuous variables, such as the monthly income maintenance payment, the STATS information gives a good picture of the distribution of responses. For coded variables, such as the Family End Code, standard frequency distributions are appended to the STATS printout, as the range, mean, and standard deviation values are less meaningful in these cases.

Working with the RFD, Glossary, and STATS printout, the user should be able to perform valuable, original analyses of the data. However, as a word of caution, it is only fair to note that, when dealing with a longitudinal data base of this complexity and volume, unexpected difficulties will inevitably appear. Some of those that SIME/DIME analysts have confronted are described briefly below.

POTENTIALS AND PITFALLS

Potential

The rich potential of the data can be illustrated by listing some of the data available on just one key research topic, i.e., the labor supply response to the guaranteed payments: race, number of heads, number, age and sex of all family members; calculated pre-experimental normal income range; experimental or control status of family; family stability, including not only breakups or unions among heads, but also arrivals and departures of tertiary members; work history of all members over 16, including type of work, monthly hours of work, number of jobs, salary, reasons for not working, job search behavior; the guarantee and tax rate of the DIME or SIME treatment assigned for experimental families, and the size of the payment each month; eligibility for and use of counseling treatment made available in the experiment.

Data Anomalies

A data base so rich affords many possibilities for research and policy analyses, but it also inevitably contains errors and anomalies, some of which are trivial and others which are not. Errors can occur at any stage, from the respondent's faulty recall to an interviewer's mistake in recording the information, to an error in entering the data, or to incorrect programming of an analysis file.

As described above in the paper by Davis, considerable effort has gone into locating and correcting errors due to miscoding, faulty programming, and similar "mechanical" sources. In the process of reorganizing

data into consistent monthly files for families and persons, and relating data from separate components of the data base, we found several anomalies of this type. For example, in programming the DIMEL0TH person records, we found several cases where the family formation code indicated "not applicable," although the family record for that person had valid information on family formation. In this case, the code was corrected.

As another example, we found three families that appeared to have no head. Inspection showed that in one family a relationship code was almost certainly incorrect: the code assigned a son as the head of the family. In this case, the relationship code was changed to male head. In the other two cases, only children appeared. Because there were no data to indicate that these "families" were indeed part of the experiments, they were removed from the file.

A more difficult example which required considerable detective work involved the case on the family-person file of a two-month-old infant working full-time 177 hours a month at \$5.75 an hour. There were, in fact, several persons whose birthdates indicated that they were young children, but who had work histories and separate person records in the file. The employment history and master file records for these individuals suggested that the birthdates were incorrect and that these persons were truly old enough to have a work history. As there was no basis on which to correct the birthdate as such, a warning flag was added to the probably erroneous birthdate.

Other apparent problems in the files after close inspection may reflect the bewildering variety of real-life situations confronting the

experiments' families. For example, what does one make of an adult working steadily for a wage of 15¢ an hour? Or, a person who reported working over 400 hours a month for several months in a row? Or persons who indicated they had a job and were involuntarily unemployed at the same time? Again, considerable investigation was required, looking at many other variables in the records, to see if there were reasonable explanations for these observations.

Upon a closer look, the two or three cases reporting very low hourly wages turned out to be women who described their type of work as "child care worker in a private home." These women also reported, at some time or other, receiving in-kind income. This information suggests that, while the low wages are improbable, they are not impossible, and thus cannot be easily dismissed as an error.

Similarly, inspection of the cases reporting very high work hours lends some plausibility to the responses. In one instance, a man who reported working 405 and then 426 hours in two months of 1974, and continued for the next two months at 327 and 342 hours, respectively, turned out to be a gas station manager. In the second case, a man working 425 and 411 hours for two months in a row was a local policeman. Again, these responses, while stretching the limits of the physically possible, are not improbable given the type of work reported.

Persons who indicated both employment and unemployment in the same month often had a job that ended in the middle of the month, as determined from a variable in the file on percent of the month worked. In other cases,

the person reported that a job lasted all month, but that the person was still unemployed. These cases usually were workers with construction jobs or similar employment where they may have had a job to go to and therefore considered themselves employed but could not work in the particular month because of bad weather, the next project hadn't started yet, and so on.

All of these cases can be annoying to the researcher who is trying to be sure that an analysis includes only valid responses. For many applications, it may make sense to throw out the most deviant cases. However, other researchers may need to know the anomalous but real situations that must be considered in forming welfare policy. Given that responses such as these are not obviously "errors," we did not purge them from the monthly files; on the contrary, the responses have been retained as recorded for the user to decide how to treat them.

The documentation developed for the monthly files is designed to alert the user to anomalies of this kind, so that users are forewarned and can determine whether or not to use the data. The Glossary indicates aspects of the data collection, such as wording changes in a question, that may affect interpretation of the data and also notes major inconsistencies and known problems in the data. The STATS printouts indicate for each variable where there may be an out-of-range response, and, also, through the mean and standard deviations, the magnitude of the problem posed by the anomaly. For example, the DIME10TH family-person STATS show that, although in almost every month there are persons reporting more than 400 hours worked, the mean value ranges from 145 to 173 hours, and the standard

deviation is about 50 hours. Thus, the responses exceeding 400 hours represent less than 1% of the total cases.

Mathematica Policy Research has reviewed STATS runs for every variable on the monthly files, looked at record dumps both of those files and of the source data used to construct them whenever anomalies were spotted, and consulted with its field offices and SRI International to help determine the best way to handle problems that could not easily be resolved. Nonetheless, new users will undoubtedly uncover other errors and discrepancies. They are encouraged to report these problems so that we can study them, correct the files and documentation where possible, and alert other users.

NOTES

1. The files may be obtained from Mathematica Policy Research, Inc., Suite 416, 2101 L Street, N.W., Washington, D.C. 20037 (telephone 202-833-9510). Ask for the announcement form for the public use monthly files which gives details on files, time spans, record size, and number of reels. Costs are \$560 or \$800 per file, depending on density of the file.
2. The public use versions of SIME/DIME10TH have record lengths of about 11,500 characters.

DISCUSSION

VEE BURKE: I have a fairly elementary question. I understood at the beginning that a unit had to have two members to be eligible for payment?

CITRO: Units had to have an adult and a dependent child, or a married couple.

BURKE: As a follow-up, then, in the case of the family you described that had two heads and two children and then split, one parent taking one child and one, the other. What would have happened if the mother had taken the two children and the husband taken none? Would he have been eligible for any grant?

CITRO: Yes. Once a family was in the experiment they could stay in. In practice, I believe often in that case, the experiment had trouble keeping track of that husband, but he could stay.

III

IMPACT ON WORK AND
LABOR SUPPLY

WORK EFFORT RESPONSE BY
RACE, SITE AND EXPERIMENTAL DURATION

by

Philip K. Robins
Senior Economist, SRI

and

Richard W. West
Econometrician, SRI

Several features of the income maintenance experiments make it difficult to extrapolate the experimental results to the national population. First, the sites may not be representative of the U.S. population. Second, two ethnic minorities (blacks and Mexican-Americans) are over-sampled relative to their proportions in the U.S. population, and may have different responses from other groups. Third, the experiments are conducted for a limited period of time, while a nationwide program would undoubtedly be permanent.

The Seattle and Denver Income Maintenance Experiments (SIME/DIME) were designed to discover the effects, if any, of these three potentially biasing variables. Relative to the other experiments, the SIME/DIME sample sizes were made large enough to permit separate analyses of ethnic groups (black, white and Mexican-American), site (Seattle, with a high unemployment rate, and Denver, with a low rate), and experimental duration (with three- and five-year programs at both sites). This paper reports our initial analysis of labor response by ethnic group, site, and program duration.

Data covering the first ten quarters of the experiment are used in the empirical analysis.

The models we employ measure the experimental effect by a single dummy variable. While this approach represents a considerable simplification over prior studies, it enables us to perform fairly powerful tests of differences in response among the several groups. The models include one where the experimental effect is allowed to vary freely over time (Model I) and one which imposes restrictions on the time pattern of response (Model II).

SAMPLE DATA

In SIME/DIME periodic interviews, subjects reported weekly hours of work, as well as any changes in weekly hours of work occurring since the previous interview. These data permit construction of a continuous work history for each individual. For purposes of this study, we have constructed a quarterly time series of (annualized) hours of work for the first ten quarters of the experiment and for the four quarters prior to the experiment (estimated from as much pre-experimental data as are available).

This analysis uses the tenth quarter sample and ignores data on families that left the experiment prior to the tenth quarter. If the attrition is systematically related to labor supply,¹ this may lead to bias in our estimates. From the first through the tenth experimental quarter, attrition reduced the sample by 17% for husbands (1.7% per quarter), by 13% for wives (1.3% per quarter), and by 13% for single female heads of families (1.3% per quarter).

It is important to recognize that the experimental effects reported here represent the average responses to the 11 tax and guarantee combinations

tested in SIME/DIME, and are influenced by the assignment model and do not necessarily represent the responses that would be forthcoming from any particular nationwide negative income tax (NIT) program.²

ESTIMATED EFFECTS: MODEL I

Any labor supply response model estimated from SIME/DIME data should take into account permanent differences in labor supply between experimental and control families, time effects, and potential biases caused by the assignment process. There are two reasons for this. First, apart from any experimental effects, hours of work follow an upward trend during the experiment for both men and women. During the first ten quarters of SIME/DIME, average hours of work of the control group increased by 5% for husbands, by 20% for wives, and by 12% for female heads of families. Second, the assignment model produced differences between the experimental and control groups. In the year just prior to the experiment, controls worked more hours--average hours of work of controls was greater than that of experimentals by 4% for husbands, by 28% for wives, and by 7% for single female heads of families.

The upward trend in labor supply is the result of three phenomena: first, nationwide, women were entering the labor force in greater numbers. Second, for both men and women, the employment picture in Seattle steadily improved; from February 1972 to February 1975 the unemployment rate in Seattle fell from 14.6% to 9%. Finally, by systematically eliminating families with high incomes in the pre-experimental period,³ the sample contains an excessive number of families with temporarily low incomes. As income

for these families returned to normal, the average income in the sample rose. As hours of work and income are correlated, truncation of the sample by income would influence average hours of work over time.

The pre-experimental differences between treatment and control families was a function of the assignment process; higher income groups were excluded from certain financial treatments, but not from control status. In particular, families with high incomes were not assigned to financial treatments with low breakeven levels, because it was assumed that they would not respond to these treatments. Thus, control families tend to have systematically higher incomes than experimental families.

All of these factors are potentially biasing. The simplest way to obtain unbiased estimates of experimental effects is to regress the change in hours of work ($H_t - H_p$) on the experimental treatment variable (F) and the normal income variables (E) (the estimate of income in the absence of an experiment, adjusted for family size).⁴ For the entire sample, the results, as shown in Tables 1, 2 and 3, indicate that the experiment has a substantial disincentive effect on labor supply. The table shows the net effect on hours of work comparing actual hours (which often went up) with predicted hours, which (based on the control group) would usually go up more. Male heads decrease their work effort until about the fourth or fifth experimental quarter. Their response then becomes relatively stable at a range from 147 to 187 fewer hours of work per year (an 8 to 10% reduction), compared to the hours we would expect in the absence of an experiment. For wives the response peaks in the seventh experimental quarter, when they are working fewer hours

TABLE 1

FINANCIAL TREATMENT EFFECTS ON ANNUAL HOURS OF WORK BY EXPERIMENTAL QUARTER-10TH QUARTER SAMPLE

Husbands [1, 2]											
(Standard Errors in Parentheses)											
Experimental Quarter	Total	Race			F-test for Ethnic Differ- ences	Site		F-Test for Site Differ- ences	Experimental Duration		F-test for Duration Differ- ences
		Black	White	Chicano		Seattle	Denver		3 Year	5 Year	
1	-16.8 (31.4)	-6.3 (54.9)	-24.4 (44.0)	-14.6 (72.2)	.03	21.3 (46.8)	-46.9 (41.7)	1.2	-13.9 (34.6)	-22.8 (43.4)	.04
2	-51.5 (32.2)	-53.1 (56.2)	-45.9 (45.1)	-64.3 (74.0)	.02	-7.1 (47.9)	-86.7** (42.7)	1.6	-37.1 (35.4)	-81.5* (44.5)	.96
3	-93.0*** (34.7)	-79.6 (60.5)	-87.5* (48.5)	-131.7* (79.6)	.15	-41.1** (51.6)	-134.0*** (45.9)	1.8	-68.1* (38.1)	-144.7*** (47.8)	2.5
4	-146.7*** (34.82)	-121.1** (60.7)	-128.0*** (48.7)	-242.9*** (79.9)	.90	-108.4** (51.8)	-177.0*** (46.2)	1.0	-118.0*** (38.3)	-206.3*** (48.0)	3.2*
5	-187.3*** (35.1)	-195.6*** (61.3)	-166.0*** (49.2)	-230.1*** (80.6)	.25	-177.3*** (52.2)	-195.2*** (46.6)	.07	-174.2*** (38.6)	-214.3*** (48.5)	.66
6	-160.2*** (35.4)	-108.5* (61.8)	-165.3*** (49.6)	-237.3*** (81.3)	.81	-156.6*** (52.7)	-163.1*** (47.0)	.01	-152.0*** (39.0)	-177.3*** (48.9)	.26
7	-146.7*** (36.7)	-141.5** (64.0)	-112.4** (51.3)**	-249.9*** (84.2)	.99	-125.5** (54.6)**	-163.5*** (48.7)	.28	-130.1*** (40.4)	-181.2*** (50.6)	.98
8	-172.2*** (36.2)	-147.3** (63.2)	-166.0*** (50.7)	-233.1*** (83.1)	.36	-161.3*** (53.9)	-180.8*** (48.0)	.07	-138.8*** (39.8)	-241.7*** (50.0)	4.1**
9	-154.7*** (38.1)	-192.7*** (66.5)	-112.9** (53.4)	-202.4** (87.5)	.63	-120.5** (56.7)	-181.7*** (50.6)	.66	-122.7*** (41.9)	-221.0*** (52.6)	3.4*
10	-183.7*** (38.8)	-193.8*** (67.6)	-123.0** (54.2)	-332.6*** (88.9)	2.1	-140.1** (57.6)	-218.2*** (51.4)	1.0	-165.6*** (42.6)	-221.2*** (53.5)	1.0

¹Sample size = 2,171

²Proportion of sample receiving financial treatment = .57

* Indicates significant at the 10% level.

** Indicates significant at the 5% level.

*** Indicates significant at the 1% level.

NOTE: Control variables include eight dummy variables for normal income categories, preexperimental annual hours of work, dummy variables for Black and Chicano, a dummy variable for Denver, number of family members at enrollment, number of children under 5 years of age at enrollment, AFDC benefits in the preexperimental year, age at enrollment, and three dummy variables for the manpower component of the experiments.

TABLE 2

FINANCIAL TREATMENT EFFECTS ON ANNUAL HOURS OF WORK BY EXPERIMENTAL QUARTER-10TH QUARTER SAMPLE

Wives [1,2]

Experimental Quarter	(Standard Errors in Parentheses)										
	Total	Race			F-Test for Ethnic Differences	Site		F-test for Site Differences	Experimental Duration		F-test for Duration Differences
	Black	White	Chicano	Seattle		Denver	3 Year		5 Year		
1	-16.7 (26.5)	13.8 (45.3)	-39.2 (37.6)	-12.3 (60.3)	.4	-15.2 (39.6)	-17.9 (34.9)	.00	5.2 (29.1)	-61.9* (36.3)	3.3*
2	-46.3 (29.4)	-116.8** (50.3)	-3.5 (41.7)	-32.3 (67.0)	1.6	-68.2 (44.0)	-29.4 (38.8)	.4	-16.0 (32.3)	-108.9*** (40.4)	5.1**
3	-75.6** (30.4)	-124.0** (51.9)	-17.3 (43.1)	-142.3** (69.2)	1.8	-33.7 (45.4)	-108.0*** (40.1)	1.5	-60.5* (33.4)	-106.8** (41.7)	1.2
4	-74.7** (30.7)	-128.9** (52.5)	-24.0 (43.5)	-110.5 (69.9)	1.4	-38.3 (45.9)	-102.9** (40.5)	1.1	-58.4* (33.7)	-108.3*** (42.1)	1.4
5	-76.7** (31.2)	-96.3* (53.3)	-50.8 (44.3)	-109.7 (71.0)	.4	-89.6* (46.6)	-66.7 (41.2)	.1	-62.5 (34.3)	-106.0** (42.8)	1.0
6	-116.8*** (32.2)	-152.5*** (55.1)	-54.9 (45.7)	-215.9*** (73.4)	2.1	-97.9** (48.2)	-131.5*** (42.5)	.3	-95.4*** (35.4)	-161.1*** (44.2)	2.1
7	-156.9*** (33.4)	-165.8*** (57.1)	-144.2*** (47.4)	-174.4** (76.0)	.08	-130.5*** (49.9)	-177.4*** (44.0)	.5	-138.5*** (36.7)	-194.9*** (45.8)	1.5
8	-141.0*** (33.8)	-126.4** (57.8)	-159.7*** (48.0)	-117.9 (77.0)	.2	-149.6*** (50.6)	-134.3*** (44.6)	.05	-136.5*** (37.2)	-150.2*** (46.4)	.08
9	-124.9*** (34.5)	-118.3** (58.9)	-125.6*** (48.9)	-135.2* (78.5)	.02	-106.1** (51.5)	-139.5*** (45.5)	.2	-107.9*** (37.9)	-160.0*** (47.3)	1.2
10	-113.4*** (34.9)	-118.9** (59.7)	-94.5* (49.6)	-153.4* (79.6)	.2	-81.0 (52.2)	-138.6*** (46.1)	.7	-96.5** (38.4)	-148.4*** (47.9)	1.1

¹Sample size = 2,252²Proportion of sample receiving financial treatment = .58

* Indicates significant at the 10% level.

** Indicates significant at the 5% level.

*** Indicates significant at the 1% level.

NOTE: Control variables include eight dummy variables for normal income categories, preexperimental annual hours of work, dummy variables for Black and Chicano, a dummy variable for Denver, number of family members at enrollment, number of children under 5 years of age at enrollment, AFDC benefits in the preexperimental year, age at enrollment, and three dummy variables for the manpower component of the experiments.

TABLE 3

FINANCIAL TREATMENT EFFECTS ON ANNUAL HOURS OF WORK BY EXPERIMENTAL QUARTER-10TH QUARTER SAMPLE

Female Heads [1,2]

Experimental Quarter	(Standard Errors in Parentheses)										
	Total	Race			F-Test for Ethnic Differ- ences	Site		F-test for Site Differ- ences	Experimental Duration		F-test for Duration Differ- ences
	Total	Black	White	Chicano		Seattle	Denver		3 Year	5 Year	
1	26.8 (34.9)	-1.0 (50.1)	-18.2 (55.3)	220.7*** (85.2)	3.1**	-86.1* (50.3)	127.8*** (47.6)	9.6***	25.0 (37.4)	31.3 (48.6)	.02
2	-30.6 (38.2)	-94.1* (54.9)	-11.4 (60.6)	112.4 (93.3)	1.9	-94.6* (55.2)	26.7 (52.3)	2.6	-33.3 (40.9)	-23.8 (53.2)	.03
3	-90.2** (40.0)	-140.3** (57.5)	-75.6 (63.5)	24.5 (97.8)	1.1	-119.5** (57.9)	-64.0 (54.8)	.5	-77.0* (42.9)	-123.4** (55.7)	.7
4	-66.8* (40.1)	-135.6** (57.7)	8.8 (63.7)	-44.5 (98.1)	1.5	-115.6** (58.1)	-23.2 (54.9)	1.4	-55.2 (43.0)	-96.3* (55.9)	.6
5	-102.6** (40.6)	-161.2*** (58.4)	-38.5 (64.5)	-82.6 (99.3)	1.1	-170.1*** (58.7)	-42.2 (55.6)	2.5	-96.9** (43.5)	-117.0** (56.5)	.1
6	-129.5*** (41.7)	-184.8*** (60.0)	-81.9 (66.3)	-79.8 (102.1)	.8	-198.0*** (60.3)	-68.4 (57.1)	2.5	-122.8*** (44.7)	-146.7** (58.1)	.2
7	-125.8*** (42.5)	-206.2*** (61.1)	-39.6 (67.5)	-94.8 (103.9)	1.8	-162.5*** (61.5)	-93.0 (58.2)	.7	-119.0*** (45.6)	-143.2** (59.2)	.2
8	-178.4*** (42.5)	-215.2*** (61.1)	-111.9* (67.5)	-229.3** (104.0)	.8	-151.7** (61.5)	-202.3*** (58.2)	.4	-174.4*** (45.5)	-188.5*** (59.2)	.06
9	-198.9*** (42.7)	-233.6*** (61.4)	-151.0** (67.9)	-211.2 (104.5)	.4	-210.5*** (61.8)	-188.5*** (58.5)	.07	-195.9*** (45.8)	-206.5*** (59.5)	.03
10	-205.1*** (44.2)	-238.8*** (63.5)	-156.1** (70.1)	-222.5** (108.0)	.4	-225.3*** (63.9)	-186.9*** (60.5)	.2	-212.1*** (47.3)	-187.3*** (61.5)	.2

¹ Sample Size = 1,656² Proportion of sample receiving financial treatment = .63

* significant at 10% level
 ** significant at 5% level
 *** significant at 1% level

NOTE: Control variables include eight dummy variables for normal income categories, preexperimental annual hours of work, dummy variables for Black and Chicano, a dummy variable for Denver, number of family members at enrollment, number of children under 5 years of age at enrollment, AFDC benefits in the preexperimental year, age at enrollment, and three dummy variables for the manpower component of the experiments.

(a 20.9% reduction). This shifts to 113 fewer hours per year in experimental quarter 10 (a 14.8% reduction). For single women who head households, the response increases until experimental quarter 10, when they are working an average of 205 fewer hours per year (a reduction of 18.2%), compared to their predicted hours in the absence of an experiment.

There are also apparent site differences, but the differences are not statistically significant. The larger difference in Denver may be due to ethnic differences (there are no Chicano families enrolled in the Seattle experiment) or to differences in the labor markets of the two cities.

Economic theory predicts that under a temporary program the change in work associated with increased income maintenance payments (the income effect) is smaller and the change associated with the benefit reduction rate, or tax (the substitution effect) is larger than under a permanent program. Thus, the observed responses of three-year families, relative to five-year families, would depend on the size of the income and the substitution effects and the rate of time preference.

Estimates of response by length of the experiment show significant differences for husbands in three out of ten quarters, with a greater response in five-year families. As observed responses are larger for five-year families, theory suggests that the income maintenance effect dominates the substitution effect.⁵ For women the duration differences are also larger in the five-year families, but these are significant only in the first two quarters for women in dual-headed families.

ESTIMATED EFFECTS: MODEL II

These estimated experimental effects, although clearly negative, fluctuate considerably from one quarter to the next. As Model I

does not constrain the time pattern of the labor supply response in any way, it does not provide a precise picture of this response. More precision is possible by using some simple theoretical concepts to impose a relatively restricted structure on the time pattern of response.

Consider a person who is enrolled in an NIT experiment. Upon enrollment there is a sudden and unforeseen change in the person's budget constraint. The person now desires to work a different number of hours. For a variety of reasons it is unlikely that the person will immediately adjust hours of work to correspond with these new desires. Many jobs do not have flexible hours and adjustment may have to wait until a new job is found. Furthermore, the person may not even attempt to find a new job because of the costs associated with changing jobs. However, if the job is lost for reasons outside the person's control or the job becomes undesirable for other reasons, the person may find a new job having hours consistent with his or her desires.

Model II does not assume that individuals immediately adjust to the experiment; it imposes a constraint on the time pattern of response and is based on the assumption that each person adjusts his or her hours of work by the proportion of the difference between desired and actual hours of work in every period. We then take this proportion and use it to project a long-run response to the experiment. This model--called a partial adjustment model--permits estimates of long run response as well as the speed at which the adjustment takes place.⁶

Table 4 presents estimates of the long run response to the experiment

and the half-yearly speeds of adjustment. The results indicate long-run financial treatment effects of -159, -108 and -201 hours of work for male and females in dual-headed families and women who head families, respectively. The corresponding percentage effects are -7.5%, -15.4% and -19.0%. The speeds of adjustment are .37, .28 and .22 respectively, indicating that 22 to 37% of the deviation between actual and desired hours is removed each half year. The time required for a 90% adjustment is 4.6 years for single female heads, 3.6 years for females in dual-headed families, and 2.5 years for males.⁷

TABLE 4
FINANCIAL TREATMENT EFFECTS ON ANNUAL HOURS OF WORK
AND SPEED OF ADJUSTMENT, USING A PARTIAL ADJUSTMENT MODEL
(Estimated Asymptotic Standard Errors in Parentheses)

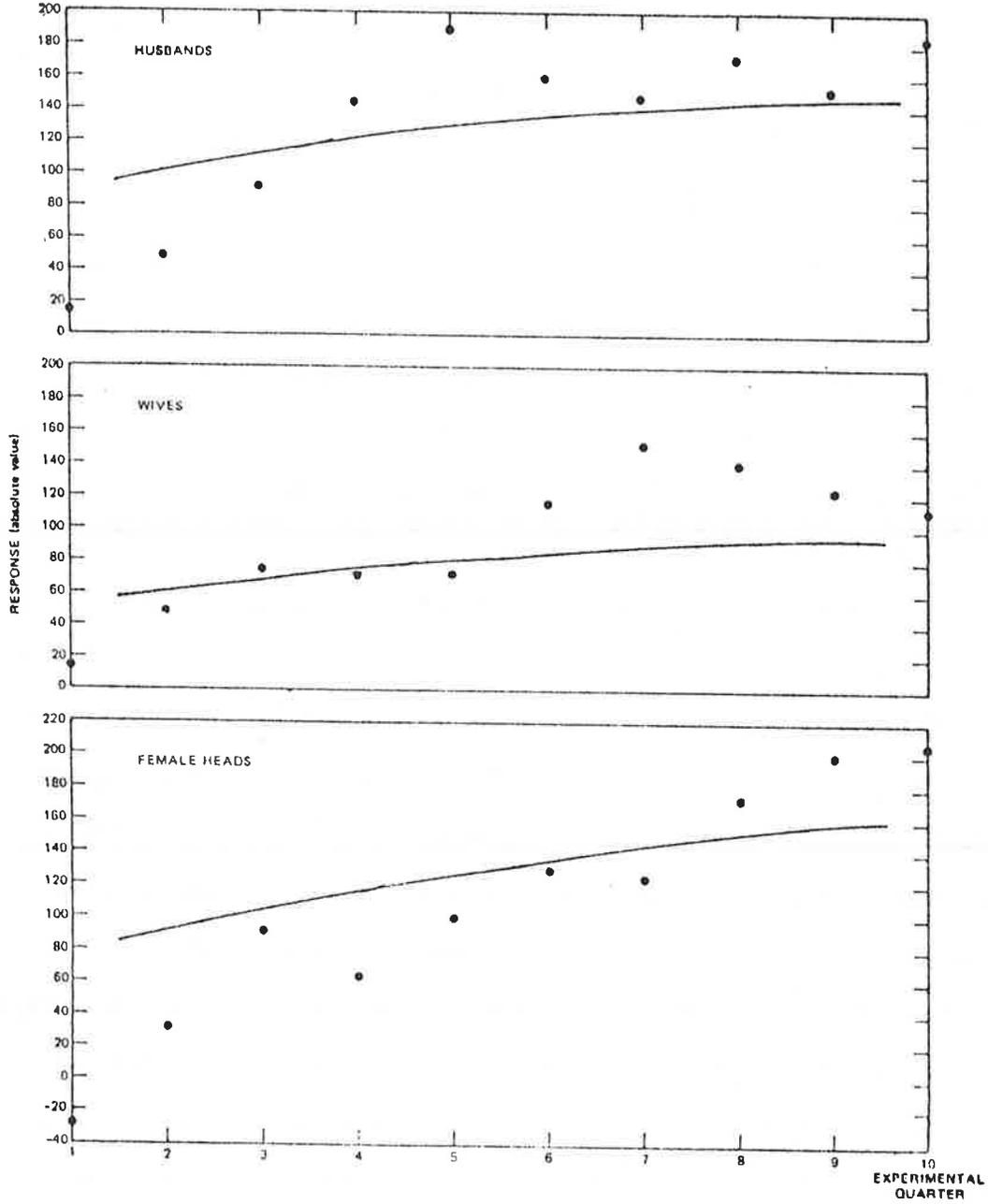
	<u>Annual Hours of Work</u>		
	<u>Husbands</u>	<u>Wives</u>	<u>Female Heads</u>
Long Run Financial Treatment Effect	-158.9*** (30.8)	-108.0*** (32.1)	-201.1*** (48.5)
Speed of Adjustment Per Half Year	.364*** (.013)	.276*** (.011)	.221*** (.012)

***Indicates significant at 1% level.

The estimated responses to the financial treatment appear to be somewhat lower than the quarterly estimates generated from the previous model. Figure 1 compares the absolute values of the quarterly responses estimated with Model I and the calculated half-year responses from Model II. As can be seen, the fit of Model II is generally good.

FIGURE 1

GRAPHS OF ESTIMATED QUARTERLY RESPONSES FOR MODEL I
AND HALF-YEARLY RESPONSES FOR MODEL II (Annual Rates)



Key: Line represents Model II, dots represent Model I.

RACE, SITE, AND EXPERIMENTAL DURATION DIFFERENCES IN RESPONSE

Having discussed the basic results for Model II, we now proceed to discuss differences in the long-run financial treatment effect by race, site, and experimental duration. Estimates of separate responses by race, site, and experimental duration are given in Table 5. The results indicate that there are no significant differences in response by race, site, or experimental duration for women, whether in dual- or single-headed families. However, for men, the differences are significant in all three cases. Black or Chicano men have a response about twice the response of white men. The response of men in Denver is about twice the response of men in Seattle.⁸ The response of men on the five-year program is 1.6 times the response of men on the three-year program.

The finding of a significant difference in response between the three- and five-year husbands is important because of its implications for the effects of a permanent program. Presumably, the larger response of the five-year husbands implies that husbands in a permanent program would have an even larger response. However, it is important to note that the estimated difference between the three- and five-year programs are for some average of the programs being tested in SIME/DIME. Since the theoretical biases of a short duration experiment have opposite signs according to whether the response is generated by the substitution effect (which is overestimated) or the income effect (which is underestimated), the total bias of a particular program will vary with the support level and the tax rate of that program. In particular, most programs that might actually be implemented are likely to have

TABLE 5

RACE, SITE AND DURATION DIFFERENCES
IN THE LONG RUN FINANCIAL TREATMENT RESPONSE

	<u>Husbands</u>	<u>Wives</u>	<u>Female Heads</u>
Total	-158.9*** (30.8)	-108.0*** (32.1)	-201.1*** (48.5)
Race			
Black	-229.9*** (45.6)	-115.5** (45.7)	-217.1*** (59.5)
White	-96.3** (38.4)	-88.7** (39.9)	-150.3*** (63.4)
Chicano	-203.8*** (54.0)	-139.5*** (53.6)	-272.9*** (85.6)
χ^2	7.867**	.818	1.941
Site			
Seattle	-99.8*** (40.4)	-114.6*** (41.8)	-209.5*** (61.1)
Denver	-202.7*** (36.5)	-103.3*** (37.2)	-194.5*** (55.6)
χ^2	5.132**	.062	.054
Duration			
3 years	-132.6*** (33.8)	-84.4** (35.1)	-190.9*** (52.0)
5 years	-214.8*** (42.8)	-156.2*** (44.4)	-226.5*** (67.4)
χ^2	3.548*	2.530	.293

NOTE: The coefficients represent the long run financial treatment effect. Estimated asymptotic standard errors in parentheses. The χ^2 statistics are for the test of the null hypotheses

of no race, site or duration differences respectively. Degrees of freedom are 2 for the race tests and 1 for the site and duration tests. For race and site, the model differs from the basic model in the text in that there are no equations for the determination of treatment, and normal income levels are included in the H* equation.

** indicates significance at the 5% level

*** indicates significance at the 1% level

lower support levels than the average SIME/DIME program. Consequently, the difference between the permanent effects of such a program and predictions from a limited duration experiment is likely to be less than the difference implied by these results.

The results for Model I indicate no race, site, or duration differences for any group. However, the results of Model II indicate the presence of race, site and duration differences for men. We can have greater confidence in the Model II results because it constrains the response over time, and thus provides a more powerful test of the null hypotheses that there are no differences.

SUMMARY AND CONCLUSIONS

In this paper we specify two different models to identify race, site and experimental duration differences in the labor supply response to SIME/DIME. In Model I the experimental effect is allowed to vary freely over time. It yields no statistically significant differences in response by race, site, and experimental duration, although in many cases the differences are large enough to be economically significant.

In Model II the time pattern of response is restricted to follow a geometric form. The results of this model suggest significant differences in response by race, site and experimental duration for men, but not for women, whether they are in dual- or single-headed families. For men, whites have a smaller response than blacks or Chicanos; those in Seattle have a smaller response than in Denver; and those on the three-year program have a smaller response than those on the five-year program. It is estimated

that the length of time required to make 90% of the adjustment to the long-run equilibrium labor supply under the NIT in this model is 2.5 years for men, 3.6 years for women in dual-headed families, and 4.6 years for female heads of families.

Extensions of this analysis could take several directions. First, estimates of income and substitution effects within a longitudinal framework could be used to simulate long run responses to specific NIT programs, as is done in Keeley et al. (1977b) and in the paper by Maxfield and Robins below. Second, combining the program participation decision with the labor supply decision would help to explain movement above and below the break-even level over time. Third, investigating alternative specifications of the response over time would help in ascertaining the precise nature of the adjustment process. Fourth, additional time periods of data, including those from the post-experimental period, can be used to obtain estimates of the long-run labor supply effects of a permanent program. All of these approaches will be investigated in future work.

NOTES

1. See P. K. Robins and R. W. West (1978, Appendix A) for a discussion of attrition on labor supply which concludes that use of the tenth quarter sample does not lead to a significant bias in the estimated treatment of effects.
2. Estimates of the labor supply response to particular nationwide NIT programs are presented in Keeley, et al. (1977b) and in the paper by Maxfield and Robins, below.
3. A family was eligible for SIME/DIME if pretransfer income (adjusted for family size) was less than \$9,000 per year in a family of four with one working head and less than \$11,000 per year in a family of four with two working heads. For a discussion of how the experimental sample was chosen, see Kurz and Spiegelman, "The Design of the Seattle and Denver Income Maintenance Experiments" (Research Memorandum No. 18, Stanford Research Institute, Menlo Park, California, May 1972).
4. Several other variables whose presence is likely to increase the efficiency of the estimated treatment effects are also included in the equation. These additional variables are pre-experimental hours of work, race and site dummies, age at enrollment, number of family members at enrollment, number of children under five years of age at enrollment, AFDC benefits in the pre-experimental year, and three dummies for the manpower component of the experiment. Equations are estimated in which the treatment dummy, F, is interacted with ethnicity, site, and experimental duration dummies. Because pre-experimental hours of work is included as an independent variable, we use experimental hours of work as the dependent variable.
5. The income and substitution effects of a permanent program can be estimated empirically using post-experimental data (Keeley, 1977). Such an approach will be investigated in future work.
6. See Philip K. Robins and Richard W. West, "A Longitudinal Analysis of the Labor Supply Response to a Negative Income Tax Program: Evidence from the Seattle and Denver Income Maintenance Experiments (Research Memorandum No. 59, Center for the Study of Welfare Policy, SRI International, December 1978b), for more details on this model.
7. These figures are calculated as $\log(.1)/\log(1-\gamma)$, where γ is the speed of adjustment.
8. This result could be due to the fact that there are no Chicanos in Seattle; thus reflecting race rather than site differences. However, tests for the black and white sample only (Chicanos excluded) indicate that the site difference is not due to the presence of Chicanos in Denver. The site difference is significant at the 10% level and the Denver response is still twice as large as the Seattle response (-85 versus -182).

DISCUSSION

UNIDENTIFIED PARTICIPANT: It sounded as if there were a lot of value judgments made in the development of the allocation model.

ROBINS: I think there will always be a lot of value judgments made, and they were totally unanticipated. People who developed the sample allocation model thought that you could go out and pick people to fill the cells as the computer programming model told them to: they found out it wasn't so easy. Also, you had to know the answer before you started, so some estimates of the work effort response had to be made before the experiment even began in order to select a sample.

UNIDENTIFIED PARTICIPANT: Did any of the work on labor supply response in the New Jersey or the other experiments consider the assignment model problems that you are considering?

ROBINS: I don't think so. I know that the issue was addressed recently by John Cogan at the Rand Corporation: he analyzed New Jersey data. After examining the sample allocation model, his impression was that it didn't have as serious implications for the New Jersey sample as it does for ours. In our sample there are large differences in the kinds of people assigned to various experimental plans, so gross comparisons across plans will be like comparing apples and oranges.

LEE BAWDEN: In the other experiments they did control for that, but it's not the same model you used; they controlled by accounting for pre-experimental hours of work.

ROBINS: They controlled for it additively but not interactively for the treatment. That is, they accounted for control-experimental differences, but not differences among experimental plans.

BAWDEN: . . . and the normal wage rate.

ROBINS: Right.

ROBERT LERMAN: Would you like to take this opportunity to say whether you would have preferred a straight random procedure, for those that are thinking about future experiments?

ROBINS: I would suggest that is a very dangerous procedure to allocate families to different treatment on the basis of characteristics on which responses so heavily depend. I would also say that it's useful to have a cost-efficient design, and costs of an observation can vary with the treatment, but I think allocation can be done in such a way that you don't lose so many degrees of freedom in the analysis.

LERMAN: That means move more toward a random sample.

ROBINS: Yes.

WEST: I agree with that. It would make life easier for us.

LERMAN: Normal people could read the results.

ROBINS: Yes, the papers contain many references to random sampling, and many people probably believe sample selection was completely random. Yet it's random only within a small space. We have had some conversations with some of the people who have participated in the Health Insurance Experiment with the Rand Corporation, and they have decided to use a less restricted random sample.

RONALD DEAR: Would you repeat briefly why you didn't use more people?

ROBINS: The budget for the experiment was fixed and they wanted to obtain a sample which yielded the most information in statistical terms. So they exploited the fact that a high-income family costs less in a given program than a low-income family, because their payments would be smaller. Therefore, they took the high-income people and assigned them to the more expensive programs and took a lot of the low-income people and assigned them to the least generous programs.

DEAR: Purely an economic consideration?

ROBINS: Yes. Unfortunately, they failed to do one thing. After coming up with this highly stratified sample, they did not calculate whether the sample sizes that were obtained were large enough to obtain statistically reliable results. The unrestricted random sample would have been smaller, but reliable results would have been possible because the number of experimental parameters to analyze would have been much smaller.

BAWDEN: Did you stratify samples by income?

WEST: Yes, SIME/DIME samples were stratified by families' predicted income, adjusted for family size.

JIM WALSH: Your data indicate that control families have systematically higher incomes. How does that affect the conclusions?

WEST: Some of the families that were supposed to be assigned to the least generous programs had high incomes. So the experimental designers said, "Well, these families are probably above the breakeven level and won't be useful as experimentals." So they didn't assign them to those experimental plans, yet comparable income families were assigned to control status. We ended up with higher income for control families than for experimental families.

ROBINS: It's a very complicated subject, but essentially the assignment models considered literally hundreds of different experimental treatments, because there was a different experimental treatment, theoretically, for blacks, for whites, for Mexican-Americans, for one-parent families, two-parent families, and for each income level. Since the number of families within some of these cells was very small, and since the families were not comparable, you would have to estimate a different experimental effect for each of them. But the sample cells would be much too small to be statistically effective. Richard West, in research done a few years ago, determined that something like 20,000 families would be required to obtain statistically significant effects in an unrestricted assignment model where there were that many effects to be estimated.

KAY THODE: Well, what is the implication of that?

ROBINS: The implication is that you have to rely much more heavily on theoretical considerations in measuring experimental response and make more assumptions on how people behave. Instead of letting the data tell you what's going on, you have to rely on theoretical considerations. By being sparse in parameterization of the treatments, we allow a lot of people to give us a little information, rather than a few people to give us unreliable information about a lot of variables.

Another way to obtain more efficient estimates is to make use of the longitudinal nature of the sample. That is, to take into consideration that people's labor supply may change over time, either because of the experiment or because of other reasons. Using these additional observations should yield "larger samples." Now, if everybody's labor supply remained constant over time, and if everybody immediately adjusted to the experiment in the first quarter, then longitudinal samples would be useless. However, we know that we can exploit the time dimension to get better estimates of the effects. This work is really exploratory--in the sense that the model is a much simpler model statistically than the ones we described in the paper--yet they are very suggestive. We do find statistically significant differences for husbands in all three of these strata; for Seattle versus Denver (with larger responses in Denver) and race (with larger responses for blacks and Chicanos). As there are Mexican-Americans in Denver and not in Seattle, the site effect may be confounded by the ethnicity effect, but we also tested for that, and concluded that the site effects are real. We also find larger effects for five-year families, compared to three-year families.

Economic theory does not indicate whether five-year families will have larger responses, but that's what we found. SIME/DIME is also testing a 20-year program, but unfortunately there are few families enrolled--I think 170 families in Denver, with all racial groups and all family structures. We may not have a large enough sample to ascertain any precise effects on 20-year families.

LERMAN: Excuse me, for duration effects, you mean there are no a priori reasons for believing the three-year response will be smaller?

ROBINS: Exactly. I doubt seriously whether a more powerful test could be performed with the SIME/DIME data. Results presented in this paper now are results for the experimental families. They are not results for the national population. Further, they are average responses over all experimental families--including families who are above the breakeven level as well as families below, families on the \$4,800 program as well as on the \$3,800 program, and so on. So our results are highly specific to the sample, and further analysis is needed to take these results and extrapolate them to the nation.

I think it's also important to note that all these results represent total effects of the experiment: we do not eliminate individuals from the sample if they change their marital status. This is a very different approach compared to the other experiments. So these responses may reflect marital status changes as well as effects that have operated through the financial treatments alone. If there is any interaction between marital status changes and labor response, they are included in the aggregate effect. Some of the research work we are doing now is trying to untangle these two. I don't know whether we will be successful, but I think it is important to note if we had taken another sample--where families kept a constant marital status--we might have gotten different answers.

BERNIE STUMBRAS: May I ask how big a portion of the sample changed marital status?

ROBINS: I think the sample size was reduced by about one-half, including not only family status changes, but also by people moving out of the area.

WEST: Roughly around 20% of the sample changed marital status in the first two years of the experiment, but for more detail you should refer to the presentations on marital stability.

ROBINS: Because we used the tenth quarter sample, we lost families due to attrition, as discussed in the paper.

Another thing which distinguishes this experiment from the others is in the way we measured hours of work. We measured labor supply, or hours of work, on a continuous basis. We take account of all changes in hours of work that occur over short periods of time for individuals. Therefore, the measure of labor supply we use, although sometimes given by year, is annual totals and take into account movements in and out of the labor force, movements from full-time to part-time to overtime, movements from one job to another, etc. More widely used techniques for gathering labor supply data rely on questions like, "How many hours did you work last week?" "How many weeks did you work last year?" Then the analysts extrapolate those responses to an aggregate annual total. I think that is a distinct advantage of the SIME/DIME data, the hours of work measures are very precise.

CURTIS ALLEN: Do you have the discontinuities broken down so you can show them separately--part-time, full-time, in and out of work?

ROBINS: So far, the analyses that we have performed have just looked at, in addition to these duration issues, only the aggregate annual reduction in hours of work. We have not looked at whether this is concentrated among over-time workers, reducing their overtime, or among part-time workers dropping out or among full-time workers reducing to part-time workers. But we have a study we will be performing in the next year that will address these specific questions.

ALLEN: Do you have wage information also, so that you can play with an annual wage income figure as well as hours?

ROBINS: Oh yes, we have very comprehensive economic information on each individual.

ALLEN: Have you calculated that yet so that you can see whether reduction in hours that you find here is matched by reduction in income?

ROBINS: We have not yet looked at the effects of the experiment on earnings.

ALLEN: Do you have occupational information?

ROBINS: Yes, and we are also looking at the effects of the experiment on changes in occupational levels. Not all of this is being reported at this time.

LERMAN: The eleventh and twelfth are the last two quarters. Wouldn't that affect the response?

ROBINS: You would think so, but if you look at the three-year/five-year differences, there is no differential decline for the families, even by the tenth experimental quarter. You would have expected the three-year families' response to peak earlier and perhaps be declining at the tenth experimental quarter, but in fact, we don't really observe that.

LERMAN: But in the eleventh and twelfth we do?

ROBINS: Yes, but it's the same for both the three-year and five-year families. There will always exist an experimental effect, although it is probably not very large even after the experiment ends.

LERMAN: For the same reasons you would expect a difference between three- and five-year families I would also expect a bigger difference in those results for white families as they face lower average interest rates and there is some preliminary evidence that estimates of time are somewhat longer for white families, even correcting for income. Are you finding that kind of result?

ROBINS: Yes. Whites have a lower response than blacks and Chicanos.

CHARLES METCALF: So one, as a result, should feel less comfortable about the smaller response of whites because you are not getting all of it.

ROBINS: I think that's a good point. The danger, of course, is in treating the three-year/five-year and black/white variables as independent of each other. The statistical tests within this model do not suggest that there are blacks interacting with duration and so on. But looking at these strata separately, there does not appear to be any significant differences by either race, site, or experimental duration, except that in a few quarters the experimental duration effects are significantly larger for the five-year families.

LERMAN: In just eyeballing Figure 1 for husbands, there seem to be a lot of dots above the line. In fact, for long-run response in almost every quarter except the fourth, they are well above the line.

ROBINS: We estimated another model in a longer version of our paper and it seems to give an even better fit, and has slightly different implications.

LERMAN: For the 20-year sample, say you only have 180; I don't think that's so small by other standards You could ask the question about three and longer, and you would have 180 extra observations over that period. I realize there are problems with that, but still

ROBINS: One problem is that the 20-year sample wasn't created until after the experiments began, so there are data processing problems associated with this sample. Some families were converted from control status to the 20-year program while some families were converted from other treatment programs. It's very complicated. So we decided here not to try to deal with some of these problems in the way the 20-year sample was created.

LERMAN: Okay, I didn't know that.

METCALF: There may have been substantial effects of the assignment models on duration, however I would be extremely surprised if that could account for anything like a 60% differential. At most, it can cut it down by maybe 45 or 50%, something like that.

Second, if that 60% were true, the conclusion we would draw is not that we should be using a 60% larger estimate, because the true estimate would be still larger, and you would have to do some sort of projection beyond the five-year estimate in order to get the true effect.

ROBINS: However, it is not necessarily a linear projection.

METCALF: No, but your best estimate is going to be bigger.

My third observation is that you correctly stressed that your second model depends on the particular way you constrained it and if that particular time constraint is wrong, then your estimates might be wrong. But on the other hand, if you look at the primary case where you have unconstrained results (male heads on experimental duration), it's sort of misleading to observe the result is significant only for three periods, because, in fact, the five-year effect is bigger than the three-year effect in ten of the ten possible quarters. It wouldn't take a very powerful order statistic to be significant on that.

Finally, the average five-year response, unconstrained, is 53% more. Given that the difference occurs in every period virtually, when you put constraints on it, it is going to produce a pretty strong statistical result.

ROBINS: In the modeling framework I reported, we found that the effects were larger for the five-year families. However, they were not statistically significant. In fact, we found that in every case that the income effect was higher for five-year families and that the tax effect was lower for five-year families. Since SIME/DIME has relatively high support levels, the effects of a permanent program with lower support levels may not be larger than our estimates.

METCALF: One unrelated point: It makes sense that you are finding that the results for wives are fairly large and similar in size to husbands, and for female heads in separate households you are not getting much effect. That confirms what we expected from the nature of time horizons. If there is any type of joint decision-making going on in the two-headed family, you would expect similar results for husbands and wives.

ROBINS: Yes.

LERMAN: In the case of wives, it might be possibly confounded by the extent of marital splitting.

ROBINS: It's a problem.

METCALF: These are the people on the fringe of being eligible--the two-worker families. In New Jersey we didn't have very many working wives because if the wife was working, that made the family too rich to be in the experiment, and the response of a wife withdrawing from the labor force could not be observed. Here you have a higher income cutoff, but still it's going to confound the wife effects.

ROBINS: There was an explicit framework adapted to acquire observations in which there were two workers. One of the income strata used was for two-worker families.

JODIE ALLEN: Do you suggest your results are underestimates, in that they were averages of Seattle and Denver, and possibly because of the fact that most of the sample was only eligible for three years?

ROBINS: Possibly. I think it is important to emphasize that the current welfare proposal in Congress has significant differences from the experimental program, including a work requirement and a job creation program.

LERMAN: Even using the low responses that you reported this morning, and generalizing those results to the national population--taking a 5 or 6% reduction in labor supply for husbands, and something like a 20% reduction for wives--much more than 10 or 20% of the cost of some of these NIT proposals will go to offset labor supply reduction. So if you are talking about a 75% guarantee and a 50% tax, the government would be spending \$8 billion, 2.4 billion of which would offset earnings reduction induced by the program itself. Even if 6% is only 6%, you can have different interpretations as to whether that's large or small.

ROBINS: Let me respond in two ways: First, thus far what we have done is to compare a given response function among alternative programs. It would also be very worthwhile to take a given program and compare alternative estimates. Secondly, there is one glaring fact about all of these numbers that we have been stressing, which may come back to haunt us. These are numbers which are generated from statistical models, and associated with these numbers are standard error confidence intervals. What may appear to be a 25% additional cost due to earnings response may be, in fact, only somewhere between 5 and 40%. It's very important to remember that these are just estimates and these aren't the facts.

METHODOLOGICAL ISSUES IN
LABOR-SUPPLY ANALYSIS

by

Michael C. Keeley
Senior Economist, SRI International

For many years there has been interest in replacing the existing complex welfare system in the United States with a nationwide negative income tax (NIT) program. The feasibility and desirability of an NIT, however, depend on its effects on aggregate hours of work or labor supply (and its cost). Interest in predicting these aggregate effects has motivated considerable research.

Previous studies analyzed existing data, usually cross-sectional, using what is now a fairly well-developed methodology. Unfortunately, in part because of a variety of statistical problems inherent in analyzing such data, the range of estimates in these studies is disturbingly large and of limited usefulness to policymakers. Consequently, a new approach to labor-supply research is being followed--social experimentation.

In principle, a controlled experiment such as the Seattle and Denver Income Maintenance Experiments (SIME/DIME) affords the opportunity to overcome most of the problems inherent in non-experimental research, primarily because in an experiment the budget constraints of individuals are experimentally varied in a measurable way. That is, experimental

treatment is assigned randomly and does not depend on response. In practice, however, large-scale social experimentation--a relatively new research tool--requires the development of a new methodology for analyzing data and for making predictions about the effects of permanent nationwide NIT programs. This methodology must exploit the advantages of experimental data and at the same time account for the specific characteristics of the experimental design.

The experiments have highly stratified samples, they test a wide variety of treatments that are not assigned on a simple random basis, including a "null" or control treatment that represents the status quo, or for some, exposure to existing welfare system, and they follow a given group of people (who divorce, marry, have children, and die) over time. This paper presents a methodology that deals with these problems.

Although there are many sorts of behavioral responses that are of interest, including potential effects of the NIT on marital status, number of children, job choice, schooling, health, migrations, and criminal activity, this paper focuses on the methodological issues associated with using experimental data to analyze effects on hours of work or labor supply. Many of these methodological considerations, however, also apply to other behavioral impact studies.

As noted above in the paper by West, experimental data have a number of advantages over non-experimental data for estimating the labor-supply responses to alternative nationwide NIT programs, but there are a number of characteristics of the design of the experiments that

must be accounted for when measuring response, Because the cost of carrying out a nationwide NIT experiment was believed to be prohibitive, the experimental sample is stratified by family size, race, site, family type, and "normal" income. In stratified sample is a sample in which certain types of persons are either over- or under-sampled from the population from which the sample is being drawn. In SIME/DIME, Blacks, Chicanos, families headed by women, low income families and families on welfare are over-sampled.

Because the experimental sample is so highly stratified, a methodology for extrapolating experimental results to the nation is necessary. To do this we estimate the labor-supply response of the experimental families in each strata and then apply these estimates to a national data base. This takes into account differences between the highly stratified samples that are used in the experiment and the national population.¹

There are however, a variety of methodological problems involved in estimating the national response, including the method used to assign experimental treatment; participation in welfare programs by experimental treatment; participation in welfare programs by experimental sample members; proper measurement of experimental treatment; the limited duration of the treatments; differences between the experimental plans and the kind of NIT plans that would exist in a national program; and the proper measurement of labor supply.

INCOME SUPPORTS AND WORK INCENTIVES

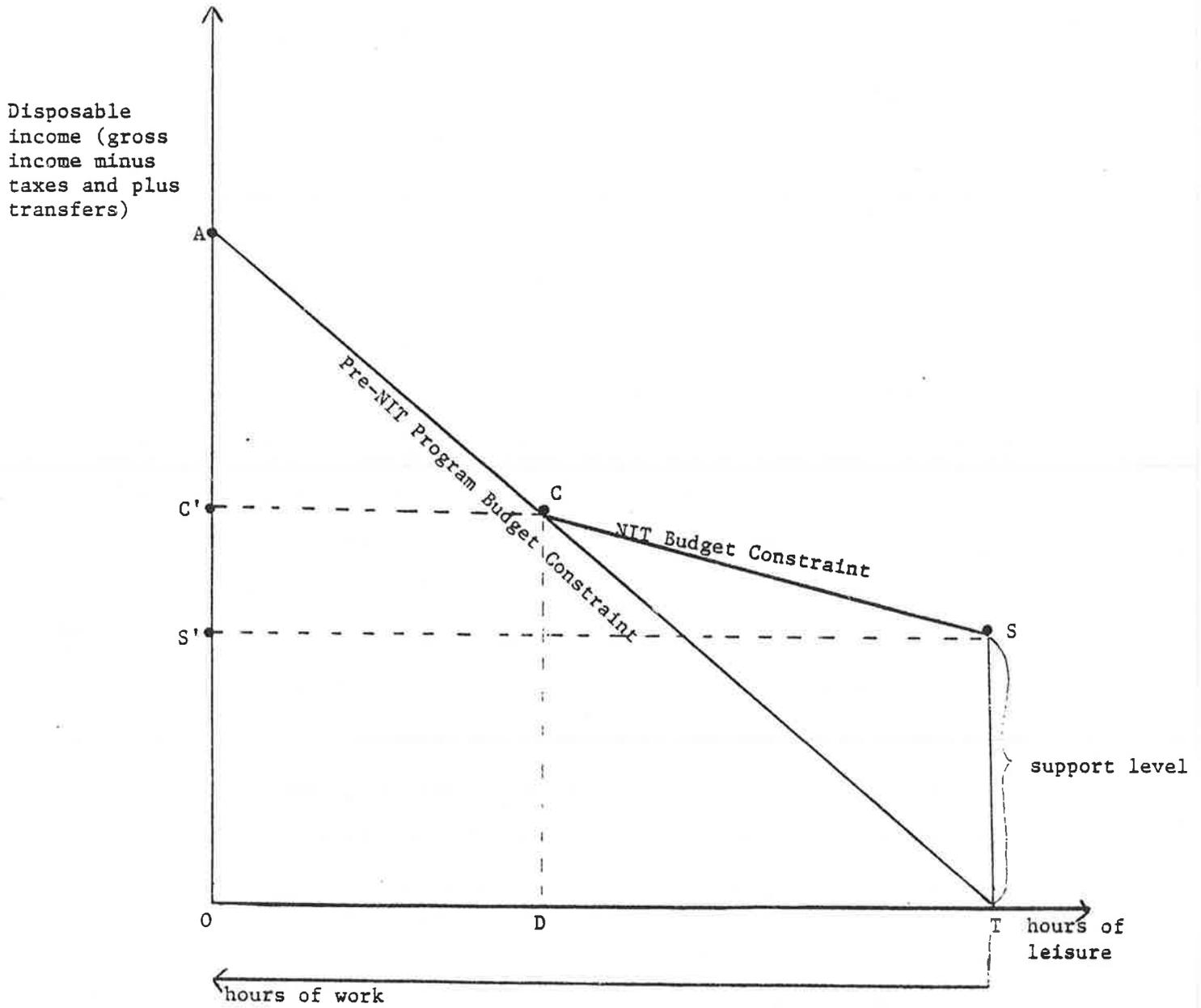
An NIT program may be characterized by a support or guarantee level and a tax rate. The support level is the grant the family receives if it has no other income and the tax rate is the rate at which the grant declines as earnings increase.

Figure 1 is a graphical depiction of how an NIT program affects a typical person. The vertical axis shows disposable (or take-home) income, which is gross income after taxes have been subtracted and transfers added. Leisure or non-working time is depicted on the horizontal axis with OT being the maximum leisure possible per period of time (168 hours per week, for example). Since all time is spent either working or not working, hours of work vary from zero at point T to a maximum of TO at point O. The trade-off between leisure or non-working time and disposable income is depicted by the budget constraint, line AT. The slope of this line is the net or after tax and transfer wage rate. (The net wage rate is equal to the gross wage rate times one minus the marginal cumulative tax rate.) This is the rate at which disposable income increases as hours of work increase.

An NIT changes the budget constraint by providing a support level TS and, usually, by increasing the tax rate. Thus, with an NIT program, a person who did not work would have a disposable income of S'. Persons working with incomes less than C' would receive grants from the NIT program that are less than S'--giving them a total income of S' plus earned income less the "tax." Since the NIT tax rates under consideration

FIGURE 1

THE EFFECT OF AN NIT PROGRAM
ON THE BUDGET CONSTRAINT



are usually much higher than the pre-existing cumulative tax rate for working low income persons, the NIT segment of the budget constraint is less steeply sloped reflecting the reduced economic return to working caused by the high tax rate of the NIT. That is, the NIT tax works just like any income tax in reducing the net wage rate or the monetary return from working. Point C is the break-even level and for someone working D hours, with income C', the NIT program does not change disposable income (holding constant hours of work).

To summarize, for persons below the breakeven level, an NIT program (1) increases disposable income and (2) reduces the economic return from working (the net wage rate per hour). Both of these factors are expected to reduce hours of work. At the extreme, in an NIT program in which the grant was offset dollar for dollar by earnings (a tax rate of 100%), there would be no economic incentive to work. At lower tax rates, disposable income falls as hours of work decline. The estimation of the effect of these factors on hours of work or labor supply is an important objective of the research.

Labor-supply effects are of particular importance because large reductions in labor supply due to an NIT would have a substantial impact on the cost of the program. The greater the reduction in hours of work, the greater the cost.

THE SIME/DIME ASSIGNMENT MODEL

In all of the NIT experiments, including SIME/DIME, assignment to the experimental treatment plans or to control status is not done on

a simple random basis. Instead, a mathematical model known as the Conlisk-Watts (Conlisk-Kurz in the case of SIME/DIME) Allocation Model² is used to determine the allocation of experimental treatments (including the control group, sometimes referred to as the "null" treatment) based on the family's race, "normal" income level, and number of heads.

In principle, this model should increase the number of observations possible with a given budget by accounting for the fact that the cost of an observation varies systematically with both household and treatment characteristics. (In fact, however, the opposite may be true.³) The case of one treatment illustrates the idea behind the model. If a treatment observation does not cost any more than a control observation, then the sample would be equally divided between control and treatment status. On the other hand, if treatment observations are more expensive than control observations, then the ratio of treatment to control observations should equal the square root of the ratio of control to treatment costs. For example, if treatment observations are four times as expensive as controls, two-thirds of the sample should be assigned to control status and one-third to the experimental treatment. This basic idea is generalized in the NIT experiments which use many treatments whose costs vary with family characteristics.

In particular, a given treatment is less costly in terms of expected NIT payments the higher the income of the family (as can be seen by referring back to Figure 1). For families with high enough

incomes to put them well above the breakeven level, the expected NIT payment is zero. Similarly, for most lower income families, NIT programs with either higher support levels and (depending on response) lower tax rates are more expensive.

The use of this assignment model results in a sample in which assignments to treatment or control status are not independent of pre-experimental household characteristics, has several important implications for measuring and interpreting experimental responses.

The assignment model used in SIME/DIME resulted in a sample in which families with low normal income levels are more likely to be assigned to programs with low support levels (or low breakeven levels). Persons with high normal incomes are more likely to be assigned to programs with high support or breakeven levels. The resulting assignment to control status, however, was just the opposite: Low normal income families are less likely and high normal income families are more likely to be assigned control status. Thus, the distribution of income of controls differs from the distribution of income of persons assigned to an experimental program.

One important implication of this assignment procedure is that simple mean differences in labor supply (or any other variable) between persons receiving treatment and controls (or between any two different treatment programs) are not indicative of a response. Instead, at the minimum, all comparisons of differences in labor supply among families on different programs, including the control group, must control

for initial differences in labor supply caused by the assignment model. This may be done using multivariate statistical procedures, if all the variables used to assign treatment are known.

A more serious problem arises if response to an NIT program depends on the same variables used to assign treatment. Economic theory, however, indicates that labor-supply response is likely to depend on pre-experimental income because the magnitude of the change in disposable income caused by the NIT depends (negatively) on pre-experimental income. This would be true of any other response that depended on income. At the minimum, any kind of expected response above the breakeven level is likely to be smaller than response below the breakeven level. Whether or not someone is above the breakeven level depends to a large extent on pre-experimental income--the same variable used in the assignment model. Thus, any comparisons of labor supply among the experimental treatments, including the control programs, must account for the different distributions of pre-experimental family characteristics used to assign treatment so that it can be determined if differences in response are due to the NIT treatments. For example, differences in response between persons on high support and low support programs are due not only to differences in support level but also may be due to the different distributions of income of persons enrolled in these programs. In fact, it is possible that the NIT-caused change in disposable income varies inversely with the support level when this assignment procedure is used.

Keeley and Robins⁴ show that in a model in which the effects of the NIT support level and the NIT tax rate are estimated without control for

interactions between assignment variables and response, both support and tax effects are about one-half what they would be if random assignment of treatment had been used. They also show that the distribution of response by income level is more profoundly affected by the nonrandom assignment procedure than it would be had the experiment used random assignment procedures. Similarly, response of high income persons is larger. This is because the assignment model gives smaller "treatments" to low income persons and larger "treatments" to high income persons than would occur with random assignment. From a policy point of view, such a misleading distribution of response would lead to serious errors in extrapolating experimental response to the nation as a whole because the national distribution of income differs substantially from the distribution of income in the SIME/DIME sample.

MEASURING EXPERIMENTAL TREATMENT

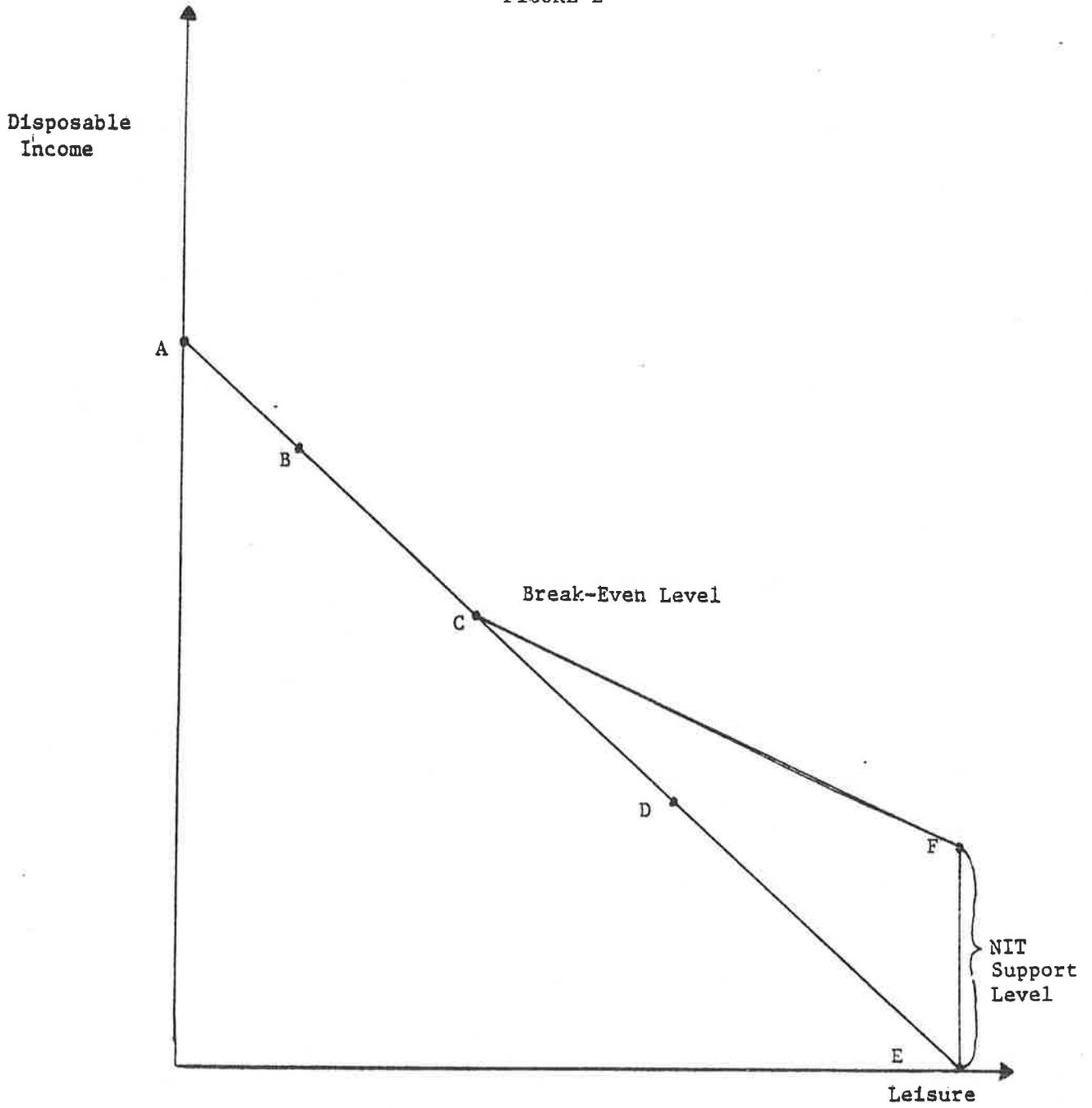
Because comparisons of treatment and control groups are not free from bias under the assignment model, the analysis of labor supply response utilizes economic theory to construct a response model that accounts for the interactions between "normal income" and response.⁵ In addition, a variety of other statistical procedures are used when estimating response in order to eliminate artificial differences in labor supply among the different experimental treatments and the control treatment that were caused by the nonrandom assignment process.

It is necessary to measure precisely the experimental treatment for each family. This is complicated in SIME/DIME because there are eleven

different NIT programs, as well as three manpower programs (the effects of which are not discussed in this paper). Although each NIT program is well specified, economic theory indicates that response to a program depends on the characteristics of the family. Thus, a given program represents different treatments to different families depending on characteristics such as the wage rates of the family members, non-labor income, tastes for work, family size, and eligibility for and participation in other welfare programs.

To see why this is so, consider Figure 2. A pre-experimental budget line AE is depicted along with an NIT program ACF with support level EF. These budget lines depict the trade-offs between work (or leisure) and disposable income. First consider a person initially located at point E prior to the initiation of the program. Although this person experiences an increase in income and a decrease in the net wage due to the NIT, he does not reduce his labor supply because he is already out of the labor force.⁶ Similarly, consider someone above the breakeven level at position B. Although he might respond, the probability of response is low and persons sufficiently far above the breakeven level will not respond. For these persons, the NIT represents a null treatment. However, for persons below the breakeven level, the NIT program represents a variety of "treatments" depending on hours worked. Each person has his net wage reduced by a given amount, but the change in income varies from zero at point C, the breakeven level, to a maximum equal to the support level, EF, at zero hours of work. Thus we see that a given NIT treatment varies for

FIGURE 2



persons of a given wage rate but different hours of work. A similar argument applies to persons with different wage rates and given hours of work.

A similar problem occurs for persons on welfare prior to the experiment because SIME/DIME requires these persons give up welfare payments if they wish to receive NIT payments. For example, a person who is previously enrolled in a welfare program identical to the SIME/DIME program that replaces it undergoes no change in circumstances and essentially is receiving a null treatment. For such an individual, no response is expected. Other persons previously on welfare receive much smaller treatments than persons not on welfare--because the support component of the welfare program is subtracted from the NIT support and because welfare programs have implicit tax rates similar to the NIT programs under consideration. Because the change in disposable income and in the net wage rate is smaller for a person on welfare, we expect his or her response to be much smaller. Thus, a given NIT program represents different treatments to different persons depending on their pre-experimental receipt of welfare benefits.

Finally, in all the tested NIT programs, the support level varies with family size and, in the declining tax programs, the tax rate varies with family size. Thus, different size families with the same income assigned to the same NIT program have different changes in their budget constraints and hence receive different treatments.

To summarize, we see that the actual treatment a family receives depends not only on the NIT program to which it is assigned, but also

depends on the family's number of children and other members, income, wage rates, hours of work, and participation in other welfare programs.

In the labor supply research, the effects of welfare, family size, and differences in initial hours and wage rates on response are controlled for by directly measuring the change in each individual's budget constraint that is caused by the NIT. The procedure used measures the change in terms of three variables: (1) whether the person is below or above the breakeven level of income, (2) if the person is below, the change in disposable income at initial hours of work caused by the NIT, and (3) the change in the net wage rate caused by the NIT. With this formulation we not only account for the nonrandom assignment of treatment but also derive estimates that may be used to predict labor supply response to programs other than those being specifically tested in SIME/DIME. Also, it is possible to predict response for populations that differ from the SIME/DIME sample in characteristics that affect response such as family size, receipt of welfare, and income.

There is, however, another difference between an experimental NIT program and an actual NIT program that must be accounted for: An experimental program is temporary and an actual NIT program would be permanent. According to economic theory, the labor supply response to changes in income caused by the NIT would be much less in a temporary program. Alternatively, the effect of changing the net wage (holding constant disposable income) would have a larger effect in a temporary program. This is because in a temporary program it is possible to substitute future as well as current leisure for income. For example, for a person

planning an extended vacation, it is less expensive, in terms of foregone income, to take the vacation during the experiment than after the experiment ends. Consequently, some people will shift leisure time planned for the post-experimental period to the experimental period. The net result of these two opposing effects depends on their relative magnitude.

SIME/DIME takes several approaches to account for the temporary duration of the experiment. First, by making treatment duration an experimental variable, it is possible to use experimental data to measure the effect of varying program duration. Second, the collection of post-experimental data permits analysis of how post-experimental labor supply varies with treatment. This should make it possible to obtain estimates of the response to a permanent program.⁷

Once unbiased and statistically significant estimates of the effects of experimental changes in income and wages on labor supply are obtained, it is still necessary to develop a methodology for predicting nationwide responses. Nationwide responses will differ from the sample responses because of differences in income, explicit and implicit tax rates, hours of work, wage rates, race-ethnicity, family size, and welfare benefits between the SIME/DIME sample and the national population. To use the SIME/DIME results as a basis of national estimates, SIME/DIME estimates of response are applied to a representative random sample national data base found in the Current Population Survey. A micro-simulation computer program (MATH), developed for these experiments, is used to perform the calculations.⁸

CONCLUSIONS

To conclude, social experimentation is a powerful new research tool that is being used to provide estimates of the labor-supply responses and costs of alternative NIT programs. Although experimental methods have many advantages over non-experimental methods, a different methodology must be used for analyzing response to experimental programs that explicitly accounts for the design. This is largely because the experimental samples are highly stratified and because experimental treatment is assigned only on a stratified random basis. Other aspects of the experiment, such as varying duration of treatment, also need to be taken into account. In this paper, an outline of a methodology that accounts for these problems is presented. The methodology outlined in this paper enables us to fully exploit the advantages of experimental data (primarily, exogeneity of treatment) but at the same time accounts for the unique aspects of the experimental design that cause problems in the analysis.

NOTES

1. See Beebout and Maxfield above, for a more detailed discussion.
2. John Conlisk and Mordecai Kurz, "The Assignment Model of the Seattle and Denver Income Maintenance Experiments" (Research Memorandum no. 15, Center for the Study of Welfare Policy, Stanford Research Institute, July 1972); Michael C. Keeley and Philip K. Robins, "The Design of Social Experiments: A Critique of the Conlisk-Watts Assignment Model" (Stanford Research Institute, Research Memorandum no. 57, November 1978).
3. Keeley and Robins, id.
4. Ibid.
5. The details are found in Michael C. Keeley, Philip K. Robins, Robert G. Spiegelman, and Richard W. West, "The Estimation of Labor Supply Models Using Experimental Data," American Economic Review 68, no. 5 (December 1978): 873.
6. A special statistical technique, called tobit estimation, is used to account for the fact that response cannot be larger than initial hours of work.
7. Michael C. Keeley, "Using Post Experimental Data to Derive the Effects of a Permanent Income Maintenance Program" (mimeo, Stanford Research Institute, 1977).
8. For an account of the MATH model, see West, Research Using Experimental Data, above, and Beebout and Maxfield, Generalizing Experimental Results with Microsimulation, above.

EMPLOYMENT ISSUES FOR LOW-INCOME
WOMEN: VALUES AND CHOICES

by

Linda Drazga
Barbara Devaney
Margaret Grady
Manika Sukhatme

Mathematica Policy Research

The papers in this book and other published materials show considerable interest and concern in the effect of income assistance programs on employment decisions. However, the specific effects on single women who are heads-of-households (including women who were never married, divorced, and separated women and widows) and their families have not been a primary focus, even though they make up a substantial proportion of the families requiring income assistance, and their responses would have major policy implications.

At the outset of our study, we strongly suspected that these women respond to incentives in an income assistance program which encourage or discourage participation in the labor market--both intended and unintended. In the case of low-income single mothers where the traditional role of women (i.e., maintaining the family) conflicts with the economic role of provider, the disincentives may be aggravated. It is a common belief that the mother's employment is harmful to the family; a negative stereotype of the working woman's performance as mother has resulted and is accepted by many of these working women themselves.¹ These beliefs need to be tested. Since most welfare reform contains work requirements, it is important to discover whether there are in fact individual and social consequences of women working outside the home.

Most analyses of female labor force participation have been dwarfed by studies of women's decisions to collect Aid to Families With Dependent Children (AFDC). These studies are dated, however, and do not reflect the rapid influx of women into the labor force in recent years. (The number of working mothers with children under age eighteen has more than tripled since 1950.) Among single women, more than half of those with preschool-aged children and about 66% of those with school-aged children are employed.

The focus of this paper is on the low-income single woman who is head of her household and her decision whether or not to work outside the home. The results presented in this paper are from preliminary analyses of data from the Seattle and Denver Income Maintenance Experiments (SIME/DIME). We examined some of the problems in looking at the value and allocation of time by these women and present some preliminary analysis on the probability of employment outside the home. We also looked at the social and psychological impacts of the decision to work or not on the woman and her family; and the demographic characteristics of women who work and women who do not work, with special attention to factors that policy-makers can directly manipulate to help reduce dependence on welfare, such as educational level.

DATA COLLECTION

Portions of this paper are concerned with the social and psychological impacts on the women, with data from "non-core modules." As discussed in the paper by Christopherson and Marshall, above, periodic interviews contained core questions asked over each time. Core questions concerned

benefits, expenses, and income that came into and left the home. In addition to these, non-core questions were asked and were different with each periodic. They were grouped in approximately 50 modules, or sets, with one module presented at each periodic interview. They were asked at predetermined schedules and some were repeated, so that changes over time show the effects of marital status changes, employment decisions, and other variables. These have yielded a wealth of information relating to social, psychological and economic behavior, including aspirations, attitudes, beliefs and perceptions.

Some are employment-related, as you might expect. For instance, one module includes questions such as: Are children better raised depending on whether or not the mother works? Are women happier if they work? Should a woman work if she has children under the age of six, or over the age of six? There are also modules on job satisfaction and stability, with questions about such things as satisfaction with job wage level or, for those not working, the wage level they would accept.

One module inquires into attitudes on children working: Is it good for the child? If the child does work, how much should he/she work? Should school have priority over work?

Other modules ask about attitudes toward welfare; the social ethics of work; perception of socio-economic status; discrimination; sex roles; and physical and mental health. There are also standard background modules on individual characteristics--educational history for all family members, ethnicity, religion, and marital history.

For the most part, little analysis has been done of the non-core modules, perhaps because the data is just becoming available. This paper is one of the first, and looks at effects through the tenth periodic, by which time most of the some 50 non-core modules have been asked once or twice.

There are some problems with the non-core modules. First, the questions were not necessarily asked at the same time. For example, the health status module may have been asked for everyone at the second periodic interview, but for one individual this may have been in the fourth experimental month or the eighth experimental month, so measures are not necessarily from the same point in experimental time. The data analysis must take this into account. Second, the focus of these modules overlaps. There may be two modules that have questions you want to incorporate into a measure or scale, but were presented in interviews at different points in time. Other problems occur in the use of the non-core modules, but these were the major ones we were aware of in this study.

THE SAMPLE

The data for this paper is drawn from the control group in SIME/DIME for an eight-month period in 1971, the first year of the experiment. This yielded a sample size of 463 women, all low-income due to the original stratification of the experiment.

The average age of the women is 33 years, ranging from 16 to 58 years. The average number of persons per household is 2.9, and the average number of children is 1.6; 96% of the sample has one to four children. Approximately 80% reported no earned income at the beginning of the experiment;

those who were employed reported hours of work which were more or less uniformly distributed between only a few hours a month and full-time employment. Among those employed at the outset, a little more than a third--39%--worked through the time period examined, 37% collected welfare only, and about 25% were either fluctuating between work and welfare, or did both simultaneously. (About 5% received a combination of wage and welfare income throughout.)

An intervening variable which affects the well-being of single women who head households is the presence and age distribution of children. It is generally accepted that women with young children are less likely to undertake full-time employment outside the home, and this holds for the sample examined. Only 22% of women with children under age 6 work continuously, while 52% of these mothers did not work at any time during the year. Work priorities changed, however, for women with children older than six; 50% of these worked continuously and only 25% did not work at all. (Note the near-reversal of the work pattern when children become school-aged.)

Even though the analysis period was limited to eight months and the sample to households headed by single women only (labor force and/or AFDC status changes could be expected to occur with a marital status change), nearly 1/3 of the women changed their labor force and/or AFDC status. To observe whether any demographic characteristics were related to stability in a category, the stratifying variables were further subdivided by duration. This resulted in seven analysis categories: (1) working all eight months, ineligible for AFDC; (2) working all eight months, eligible for but not receiving AFDC; (3) working all eight months, changing eligibility status

but no AFDC; (4) on AFDC all eight months, not working; (5) on AFDC all eight months, working; (6) on AFDC all eight months, changing work status; and (7) changing work and AFDC status.

PROBLEMS WITH THE VALUE AND ALLOCATION OF TIME

Before discussing the data more fully, we wish to present some of the issues in analyzing a woman's decision to work at a job, that is, enter the labor market, or stay at home. Traditional modeling efforts have regarded the decision to go on welfare as an extension of the work-leisure models. This raises obvious problems with characterizing such home activities as cooking, cleaning, and amusement of children as "leisure". According to these models, leisure (or non-work) time is purchased at the person's wage rate. Availability of welfare income lowers the cost of this "non-work" time by making a certain amount of money available--the welfare grant--regardless of work effort, and is expected to affect the work-"leisure" choice. Traditional analyses would predict that if the wage rate available to a woman in the labor market is high, she is more likely to work, but if the family's other income is high, she is not as likely to work. In an analysis of the time allocation of both women living with spouses or households headed by a woman alone, this does not adequately explain the woman's behavior.

The problem is two-fold. First, the work-"leisure" choice does not exhaust the alternatives for the allocation of time. Women choose not simply between work and leisure, but rather among work in the labor market, work in the home, and leisure. Second, the wage rate for a working person

usually serves as a suitable proxy for the value of his or her time. There is a marginal rate to consider--the rate by which her market wage rate exceeds the value of her time spent in productive activities in the home. That is, she may decide that it is worth more to her to work at home than to work outside the home.

Because she does not work, the data on the value of time for non-working women are not available. Despite this, there has been an increasing awareness of its importance in recent economic research on female labor supply.² Approximately one-half of the population of married women and single women who head households choose to spend their time solely in non-market activities in a given year. While this may be the result of cultural and family role norms, as well as of labor market discrimination against women, it may reflect a choice that takes into account the value of a woman's time in home production. This factor should be included in time allocation models.

An analysis of the value of nonmarket activities raises important policy implications. Studies of the effects of such welfare programs as the negative income tax (NIT), food stamps, and AFDC often focus on the decrease in labor supply associated with the assistance. However, the socioeconomic consequences of decreased labor supply depend on the type and value of the nonmarket activities engaged in by the labor force dropouts.

Recently, appropriate models have appeared,³ which consider both time and money costs of goods and services in the home. Family utility and happiness are viewed as results of home-produced commodities--such as

satisfaction from children, good health, and recreation. The household is assumed to maximize its utility from the home-produced goods--limited by technology, time, and income constraints. For example, whether or not a dishwasher is owned. The merit of this model is that it provides the comprehensive treatment of time allocation decisions needed in the case of women who traditionally allocate a significant portion of their life solely to nonmarket home production activities.⁴

The home production model implies that a woman will participate in the labor force if her potential market wage exceeds the value of her time in the home. Unfortunately, data on the price women place on their time spent at nonmarket activities are not available. In the case of a nonworking woman, it can only be assumed that her home time is worth more than the wage she could receive if she worked instead.

One way to circumvent the problems associated with this lack of data for nonworking women is to determine wage functions or proxies for all women by their socio-economic status, regardless of whether they work outside the home or not. Some variables that determine that status are: work experience, education, hours of market work, years of home experience, husband's income, the number of children less than six years old, and the number of children aged 6-16.

Some of our preliminary results using economic models and expanding the sample from just controls to include single women who are heads of households that receive NIT financial treatments include estimates of the probability of a woman working outside the home, or in terms of the home

production model, that a woman's market wage exceeds her home wage. The NIT payment appears to lower the probability of a woman working outside the home, and, hence, a positive effect on the value of her nonmarket wage rate. The effects of the NIT tax rate, however, are statistically not significant.

Using the socio/economic variable ideas described above it was found, as expected, that the amount of education positively affects labor force participation, suggesting that an increase in a woman's education is associated with a larger increase in her market wage than in the value of her time at home. The probability of a woman working outside the home is negatively related to the number of children, husband's income, and other household income.

EFFECTS ON WOMEN AND CHILDREN

The single mother's decision to work has policy consequences not only for the economy but also for the family. Some studies suggest working mothers experience fatigue, high anxiety, and feelings of inadequacy. Others found higher self-esteem, confidence and mental/physical health for working mothers than for nonworking mothers.⁵

The analysis reported here is a preliminary assessment of the well-being of single working mothers as compared to single mothers receiving welfare payments. The relationships between the variables affecting the well-being of these women will be extended in a future analysis of the specific tax treatment effects for the women in the experimental group.

Psychological Effects

Looking at these results, the behavior of the low-income mothers in

the sample reflects our cultural bias that women with small children stay home. However, the employment preferences of these women reveal conflicts. Regardless of the individual woman's actual employment status, the women with children under age six are more or less equally divided in their preferences for full-time, part-time, or no work. However, when all dependent children are over six years, approximately 70% of the women preferred full-time work. Even though most mothers (regardless of the age of children) desired at least part-time work, equally high proportions of both working and welfare mothers agree that children are raised better when a mother does not work. Because these are single women and there is no person to help them in the home, for instance a father, the implications about these attitudes disapproving of employment when the children are young are not straightforward. Not surprisingly, however, welfare-only women are slightly more likely to maintain that nonworking mothers do a better job of raising children.

When asked if women who work outside the home are happier, women who work continuously are less likely to answer affirmatively. The women who don't work, on the other hand, answered that women who work are happier. While this result may indicate disillusionment with working, the conclusions must be carefully considered. The level of self-satisfaction and social adjustment is affected when a woman is torn between what is desired and what is expected. Some theories hold that an individual will eventually reduce perceived discrepancies between real and ideal aspirations. One way to assess social adjustment is to study over time how the individual aligns

these incongruities.⁶

The satisfaction these mothers feel because of the time they spend with their children is also an important consideration. Previous research indicates that the amount of recreation time spent with their children, except for time spent watching television, does not differ significantly between working and nonworking women.⁷ Some work with families with two parents, however, indicates that working mothers are most likely to cut down on their hours of sleep in order to spend more time taking care of their children. SIME/DIME data suggest that working mothers are more likely than welfare mothers to report that they spend too little time with their children, and list fatigue or work demands as the causes. However, 25% of welfare women in the sample feel that too little time is spent with children and list lack of money as the cause.

Health Effects

Previous work measuring physical health shows that working women are in general healthier and report fewer symptoms of physical illness.⁸ However, it is not clarified whether working helps one stay healthier or if healthy women are more likely to work. The longitudinal SIME/DIME data may help to address this question.

For mental health, previous research has found no significant difference between working and nonworking women, although the data show tendencies for better psychological adjustment of nonworking women.⁹ The preliminary results from SIME/DIME support this, but the data suggest that the consideration of additional variables may affect the relationship. In particular,

the number and ages of children appears to affect the levels of mental adjustment for women who rely on welfare, whether totally or in part. The numbers of cases are small, but the results are consistent.

DEMOGRAPHIC CHARACTERISTICS AND THE WORK/HOME DECISION

According to life-cycle theories, a woman's labor supply is bi-modal, that is, women generally work outside the home while they are young, drop out of the labor force during the child-bearing years, and re-enter as the children grow up. In the sample, approximately 1/3 of the working women, not on AFDC, are over 40. Another 1/3 of working women, not on AFDC, as well as those on AFDC, are between 30 and 40. There is no cluster of labor force participation in our study for women under 20 as is generally assumed. This could be attributed either to the small sample size or to the fact that the women in the sample are low-income women who could be expected to have difficulties finding employment, particularly when they were young and inexperienced. Of the women depending entirely on AFDC, nearly 70% are between 20 and 40. Among the women who vary between AFDC and labor force status, these results are not so clearly pronounced. However, in general, this sample seems to support the bi-modal theory.

Other variables expected to influence a woman's decision to work outside the home or not are presence and spacing of children. It was not possible to completely control for spacing. However, the results indicate 70 to 80% of the working mothers--on welfare or not--had no children under six. Women who remained ineligible for welfare were the least likely to have children under six, while women on AFDC and not working were most

likely to have children under six. Very few working mothers--on or off welfare-- had more than one child less than six, but nearly 1/3 of the welfare-only women had two or more.

Working, AFDC ineligible women, were most likely to have only one child under 16. Among working, AFDC eligible mothers, on the other hand, the probabilities of having one, two, three, and four or more children under 16 were fairly evenly dispersed. Patterns among the other status categories were not so clearly pronounced. This seems to indicate that welfare mothers have larger families. However, there are also more women in the childbearing years on welfare. At this point it is not possible to establish whether there is a causal relationship between the age of welfare mothers and family size.

As noted above, most women, regardless of status, preferred to work at least part-time, and full-time where there were no preschool children. There was a "grass is always greener" tendency for working women to believe non-working women were happier, and vice versa. In answers to job-related questions, most women had positive attitudes about their current or last job. Those who did not were always on AFDC, always on AFDC and working, and those who changed status.

The women also rated job characteristics. Working, AFDC ineligible women rated job security most highly. Women in the other categories rated wages more highly, but also rated job security fairly high.

The results of this preliminary analysis show a clear distinction between welfare and working women (although not as clear between AFDC eligible and

ineligible) in terms of demographic variables. There were, however, very few differences in the psychological variables. A possible implication of these results is that women choose welfare over work, not to avoid work but because their demographic and life-cycle characteristics make work an undesirable alternative. Further testing will be needed before such results can be stated conclusively.

FUTURE RESEARCH

Issues addressed by additional ongoing research not reported in this paper that are both important to working women and in evaluating the NIT include: attempts to increase the number of women working outside the home by providing adequate, inexpensive and quality child care arrangements; job training and education; and the effects (including the personality, school performance, and future aspirations) of an NIT program on the children of these low-income women.

The low-income woman's decision to work or not to work reflects a full range of social and economic considerations. To evaluate the goal of an NIT program--whether that goal is to increase labor force participation, or to increase the personal/interpersonal levels of satisfaction, or to increase the quality of family life--particular consideration is required of the determinants of the low-income woman's values about work and welfare and the influence of these values upon her employment choices.

NOTES

1. For further discussion see: "A Job for Every Welfare Mother, But What About the Kids?", National Journal (March 4, 1978): pp. 341-344; Robert W. Smuts, Women and Work in America (New York: Schocken Books 1971); S. Feld,

"Feelings of Adjustment", F. Ivan Nye and Lois W. Hoffman (eds.) The Employed Mother in America (Chicago: Rand McNally 1963); and L. C. White, "Maternal Employment and Anxiety Over the Mother Role", Louisiana State University Journal of Sociology (Spring 1972).

2. See, for example, Reuben Groneau, "The Effect of Children on the Housewife's Value of Time", Journal of Political Economy 81, no. 2, pt. 2 (March/April 1973): pp. S168-S199; Groneau, "Intrafamily Allocation of Time: The Value of the Housewife's Time", American Economic Review 63 (1973): pp. 634-651; and Groneau, "The Measurement of Output of the Non-Market Sector: The Evaluation of the Housewives' Time", in M. Moss (ed.) Measurement of Economic and Social Performance Studies in Income and Wealth 38 (New York: National Bureau of Economic Research 1974); Gary Becker, "A Theory of the Allocation of Time", Economic Journal 75, no. 299 (September 1973): pp. 493-517; Robert Michael, "Education in Nonmarket Production", Journal of Political Economy 81 (March/April 1973): pp. 306-327; and James Heckman, "Shadow Prices, Market Wages and Labor Supply", Econometrica 42 (1974): pp. 679-694.

3. Some of the most interesting studies in this area are those by: Reuben Groneau, "The Effect of Children on the Housewives' Value of Time", Journal of Political Economy 81, no. 2, pt. 2 (March/April 1973): pp. S-168-S199, and "Intrafamily Allocation of Time: The Value of the Housewife's Time", American Economic Review 63 (1973): pp. 634-651; Jacob Mincer and Solomon Polachek, "Family Investments in Human Capital: Earnings of Women", in T. W. Schultz (ed.) Economics of the Family: Marriage, Children and Human Capital (Chicago: University of Chicago Press 1974) and Willis (1973).

4. For instance, the labor force participation rate of married women, husband present was 43.9% in March 1975 suggesting that a large portion of this population subgroup specializes in home production at some parts of the life cycle. Additional data on longitudinal patterns of female labor supply is provided by Francine Blau, "Longitudinal Patterns of Female Labor Force Participation", in Parnes, et al (eds.), Dual Careers: A Longitudinal Analysis of the Labor Market Experience of Women, vol. 4 (Columbus, Ohio: Center for Human Resource Research (1975): pp. 25-55.

5. See, for example, S. Feld, id., and F. Ivan Nye and Lois Hoffman (eds.) Working Mothers (San Francisco: Jossey-Bass Publishers 1974).

6. A study of the effects of this incongruity on well-being is possible with the longitudinal SIME/DIME data, and should be considered in the planning and evaluation of any welfare program. Such an analysis will be forthcoming.

7. F. Ivan Nye and Lois W. Hoffman (eds.), id., 1974.

8. F. Ivan Nye and Lois W. Hoffman (eds.), id. (1974); and S. Feld, id., and F. Ivan Nye and Lois W. Hoffman (eds.), id. (1963).

9. The mental health of working and nonworking women was found to be comparable in studies reported by: F. Ivan Nye and Lois W. Hoffman, id. (1974) and S. Feld, id. (1963).

DISCUSSION

(Of the four authors, Linda Drazga and Margaret Grady attended the conference and responded to questions.)

KATE JESBERG: You spoke of the demographic characteristics of women who work. I am wondering if you looked at availability of transportation.

GRADY: One question asks the respondent to rate how important distance to work is. No one seemed to rate it very high. There were seventeen characteristics altogether. No one seemed to attach importance to any of them other than earning a lot of money and job security.

PATSY CARTER: One thing I have seen done commonly to assign a wage value to home activities is to substitute the costs that would accrue if you were trying to replace the work done in the home--if you had to hire a homemaker or a childcare person. Was that possibility considered?

DRAZGA: This is Barbara Deveney's section of the paper. She estimated a wage rate for women who did not work, by looking at their demographic characteristics and with a regression equation determined the wage. It is my impression that she felt this was the most accurate approach.

MASAKO DOLAN: A few years ago, I saw where the Social Security Administration attempted to do that in essentially the same way--attributing a certain amount per hour cost for chauffeuring, for child care, for cooking. Did they look at that at all?

GRADY: I don't think so.

BETH HARRIS: I question one of your assumptions--that a woman's wage rate is a suitable proxy for her time.

GRADY: Perhaps her time is really worth more. If she only worked part-time and she could have worked full-time, the value of her at-home time can be assessed on the basis of the additional amount of money she would get were she full-time. At this point, it doesn't raise questions of discrimination.

DRAZGA: When you are working within the system as it is now, the woman sees her time as valued at a certain amount. Whether or not she is being discriminated against is not the question--all you're looking at is what her time will buy on the market. It may be that the price is not right, but that is not part of the theoretical argument here.

DOLAN: I can see where you would use that formulation in terms of a person trying to figure out whether economically it makes sense to go to work or not, but it does really funny things when you talk about the value of non-market work.

DRAZGA: I don't think the problem has been solved.

CARTER: In terms of the child care material here, did you go after any information about child abuse or neglect relative to working or not working?

GRADY: As far as I know, the questions were limited to "Do you spend enough time with your children?"

CARTER: Is there a future possibility of getting information about agency intervention, court intervention, the involvement of collateral services in those families relative to whether people were working or not working, whether they had money or not? Could you tie in after the fact in any way?

PEGGY THOITS: One of the big limitations with the data is that we don't know what other agencies were involved with our families. I have been especially interested in health agencies and mental health agencies. It turns out, for ethical reasons, we cannot obtain records on individuals--especially from mental health agencies. However, you can get data in the aggregate, if you can supply them with categories like black subjects and low income subjects in the sample. Jim Short and I have been working in an effort to collect delinquency records, essentially arrest records, on a select sample of juveniles. And in answer to your abuse and neglect question, we have had, in fact, records of kids who are abused and neglected. Because the sample is so small, there is no analysis possible.

CHRIS SWIFT: I have a concern. You may write a hopeful statement that the results of the analysis will guide policymakers toward income assistance programs which provide financial security, etc. However, a recent article in a socio-economic journal had a quote by Daniel Patrick Moynihan in which he was concluding from the initial evidence of SIME/DIME that greater income increases family instability. His solution is to drop the income of women on AFDC so that they won't have an incentive to leave their marriages. I think we ought to be concerned about how this evidence is going to be used by policymakers.

HARRIS: What do you see as your responsibility in terms of that kind of conclusion being made (referring to Ms. Swift's comment)?

GRADY: There should be more information available. Marital instability could perhaps be a good thing; maybe these are families where the wife is getting beaten every Friday night and she might be better off getting divorced.

SHORT: I think that this is a terribly important question. I really think that activist organizations are going to have to take an active role in the interpretation of and what is going to be done with these findings. If we don't, others are going to. The Carter policy of family stability, come hell or high water, is just simply wrong. We have got to convince them of that.

SUSAN LANE: When you look at that nonmarket time issue and compare it against the prevailing wages for women in the labor market, it is comparing it against an already depressed wage structure. That simply adds to the problem because you have not broken out of that traditional mode of viewing womens' work-- which is consistently to undervalue it, both in the market and out of it. And I think that is the reaction that you are hearing in a lot of the comments that we are making here. I would hope that there would be a way to re-analyze, to re-look at some of that data in light of some of the other studies that have been going on regarding comparable work. I think we have got to look at some new ways of looking at the world and at womens' work.

CARTER: I have a general question. I have been impressed in general with the effort that has obviously been made in all of the papers to take a fairly objective position politically in what is clearly a hot area. However, I am wondering if anybody in the project is addressing directly the issue of politics and ethics. That is to say, how is this material going to be used? We have talked about it a little here. How is it going to be used? Who is going to use it? What are they going to get? I know a little bit about the way the press has been dealt with very cautiously. How do you respond to that or feel about it? I know it is hard, but it is very important to me.

DRAZGA: Well, it is difficult because when you are doing research, you do try to be value-free and to report the results as best you can without putting your own values into them. I think it is certainly true that every researcher has his or her own values



UNEMPLOYMENT INSURANCE AND
THE DURATION OF UNEMPLOYMENT

by

Henry E. Felder
Senior Economist
SRI International

A major concern about any income maintenance program is whether it will encourage individuals to rely on it rather than work for income--that is, whether it will have a "work disincentive effect." One of the largest publicly funded income maintenance programs in existence is the current unemployment insurance system. During the most recent period of high employment (1975-1977), benefits were available for as long as 65 weeks in many states. They averaged \$5,200 yearly; highs in some states were \$134 weekly, or \$7,968 on a yearly basis. Unemployment insurance now costs the public \$20 billion nation-wide. Policymakers have naturally become concerned about the impact of unemployment insurance on the work/leisure choice of the individual.

Recent evidence suggests that the receipt of unemployment insurance reduces total work effort by increasing the duration of unemployment.¹ One could argue, however, that the longer period of unemployment permits

individuals to obtain a better job in the end. This paper explores these possibilities and reports on a study of the work disincentive effect and the impact of the receipt of unemployment insurance on future wages (our measure here of a better job),² using the control group in the Seattle and Denver income maintenance experiment (SIME/DIME) sample. The analysis reported here shows that unemployment insurance significantly increased the duration of unemployment for some groups in the sample, and that the longer durations of unemployment were not associated with higher wages on return to work.

DATA USED IN THE STUDY

From the SIME/DIME control group, we selected those who reported involuntarily unemployment at any time during the observation period. This selection procedure produced 375 observations in Seattle and 174 in Denver. These data were supplemented by data from the Washington and Colorado unemployment agencies. These agencies supplied such information as the weekly benefit amount and the maximum benefits payable for the period 1971-1974.

BACKGROUND

Economic Conditions

When unemployment is high, the duration of a period of unemployment may be longer, and there is a greater likelihood that the individual will exhaust his benefits. When there are large increases in the insured unemployment rate³ in a given state, federal funds for extended benefits are made available. Thus, the receipt of unemployment insurance may vary

for otherwise identical individuals over time and place.

The two metropolitan areas from which the data were drawn had very different economic conditions during the period studied. From January 1970 to June 1974, Seattle experienced unemployment rates of 6% to 14%. The average unemployment rate in Denver was slightly more than 4%. Thus, we would expect the average time out of work to be longer and to be smaller in Seattle since the supply of unemployed people would permit employers to hold wages down. Denver, in contrast, experienced mainly "frictional" unemployment--unemployment that results from the normal flow of workers into and out of jobs.

Structure of the Unemployment Insurance Systems

The principal features of an unemployment insurance system are: (1) the weekly benefit amount, (2) the maximum benefits payable (the total amount available to an individual in a year for all spells of unemployment),⁴ and (3) the duration of benefits. A summary of these elements as they appear in Washington and Colorado law are presented in Table 1. During the period under observation, they were constant. The table shows little difference in these elements between the two states. However, major differences in the operation of the two systems occur in the proportion of unemployed individuals who actually receive benefits. Individuals with sufficient wage credits are denied benefits for a variety of reasons: voluntary withdrawal; fired for misconduct, not able and not available for work, various labor disputes (strikes, walkouts, etc., but not including lay-offs), refusal to accept suitable work, and so forth.

In Washington, on average, less than 10% of all claimants were denied benefits; In Colorado, the denial rates ranged from 30% to 50% of all who filed claims.

Assuming that the unemployed individuals in both states are similar, this difference can be attributed to administrative policies. In fact, Washington law at this time permitted covered individuals to voluntarily quit a job under a broad set of circumstances and remain eligible for payment. Other differences may also be found. Such policies probably affect individual decisions to leave or find work, with individuals in Washington more willing to be out of work than in Colorado.

All-in-all, theory would predict that the average duration of unemployment should be higher in Washington.

TABLE 1

MAJOR PROVISIONS AND UTILIZATION OF UNEMPLOYMENT INSURANCE
LAWS FOR STATES OF WASHINGTON AND COLORADO⁴

	1974	
	<u>Washington</u>	<u>Colorado</u>
Weekly benefit amount		
Minimum and Maximum	\$17-86	\$25-102
Duration of regular benefits		
Minimum and Maximum	8-30 weeks	7-26 weeks
Maximum benefits payable	\$2580	\$2652
Beneficiaries eligible for maximum duration (30 or 26 weeks)	60%	48%
Beneficiaries exhausting total benefit available in year	39%	30%

EMPIRICAL FINDINGS

If unemployment insurance operates as a work disincentive, we should see longer periods of unemployment and fewer hours of search among those who are covered relative to those who are not. If the payoff to longer search is a higher wage rate, then, in the long run, the individual and the economy may be better off.

In Table 2, we see that males in Seattle who receive unemployment insurance tend to have longer periods of unemployment than other unemployed males. However this is not true for insured males and females in Denver and females in Seattle, all of whom experience shorter periods out of work. For contrast, Table 2 also compares those who do and do not receive welfare payments.

Table 3 shows the hours per week spent searching for work, separately, for those who do and do not receive unemployment insurance benefits or welfare payments. This table shows that females who receive unemployment insurance spend more time in job search than females who do not. For males the results are mixed.

Ideally, to determine the work disincentive effects of unemployment insurance benefits, we should select two groups who differ only in whether or not they receive benefits. Unfortunately, for the analyst, those who do not receive unemployment insurance are usually quite different from those who do. An individual may not receive benefits because he worked in an industry that was not covered such as agriculture, state and local government or domestic service. Those who quit jobs without

TABLE 2
 DURATION OF UNEMPLOYMENT (DAYS)
 1971-1974
 (Standard Error in Parentheses)

	<u>Seattle</u>		<u>Denver</u>	
	<u>Male</u>	<u>Female</u>	<u>Male</u>	<u>Female</u>
Unemployment Insurance	193.7 (219.6)	320.3 (314.9)	89.6 (77.2)	208.2 (180.0)
No Unemployment Insurance	169.2 (216.1)	400.8 (333.9)	112.4 (179.0)	255.1 (263.5)
Welfare	387.9 (302.2)	469.2 (328.3)	233.9 (262.5)	329.4 (293.9)
No Welfare	135.6 (158.6)	299.9 (311.3)	69.3 (90.2)	205.8 (225.6)

TABLE 3
 HOURS SPENT IN SEARCH PER WEEK

	<u>Seattle</u>		<u>Denver</u>	
	<u>Male</u>	<u>Female</u>	<u>Male</u>	<u>Female</u>
Unemployment Insurance	13.8 (9.6)	7.5 (5.8)	21.2 (13.5)	9.4 (7.6)
No Unemployment Insurance	14.7 (9.4)	5.9 (5.7)	16.7 (10.3)	9.3 (8.1)
Welfare	15.5 (9.4)	6.3 (5.7)	17.4 (11.3)	7.6 (6.3)
No Welfare	13.8 (9.6)	6.7 (5.8)	15.8 (7.8)	10.4 (8.8)

good cause or were fired for misconduct cannot receive benefits. Non-recipients may also have insufficient covered employment in the year prior to unemployment. There may also be important differences associated with race, age, education and job histories. For these reasons, the simple comparisons contained in Table 2 are unlikely to provide a robust indication of the work disincentive effects. Consequently, we used multivariate regression analysis⁵ to estimate the differences in work disincentive and economic returns of job searching.

Effect of the Intensity of Job Search on Duration of Unemployment

For all groups, the days spent on job seeking was inversely related to the duration of unemployment. This result suggests that individuals who spend more time looking for work are likely to find it. The results indicated that for every hour increase in the time spent searching, there was a reduction of unemployment of about five or three days, for Seattle and Denver males, respectively. The comparable figures for Seattle and Denver females are ten and nine days, respectively. Since time spent searching has such a dramatic effect on the duration of unemployment, we shall examine both variables closely.

Effect of Unemployment Insurance on Duration of Unemployment

The receipt of unemployment insurance has an effect on the duration of unemployment among Seattle males and females, but has little effect on other aspects of the job search. The results suggest that those in Seattle who receive unemployment insurance benefits will have a significantly

longer duration of unemployment than those who do not. This result will be true regardless of the way in which we measure unemployment benefits. Among Seattle males, each \$10 increase in the weekly benefit amount increases the duration of unemployment by 13 days; for Seattle females, 33 days. Maximum available benefits and the amount of benefits as a percentage of earnings also show an increase in duration of unemployment as coverage is made more generous.

Unemployment insurance had a negative effect on the hours spent in job search only for Seattle males. This suggests that increases in unemployment insurance benefits induces a greater demand for leisure and a subsequent reduction in search among Seattle males. The observed reduction in search is interesting since to be eligible, recipients must "actively" search while receiving benefits.

The results show no relationship. However, a smaller percentage of unemployed individuals receive benefits in Denver, compared to Seattle, and this may be part of the reason for the lack of significant results. (The figures are summarized in Tables 4 and 5) In any case, we cannot say anything conclusive about the disincentive effects of unemployment insurance in Denver.

Effect of Demographic Characteristics on the Duration of Unemployment

The demographic characteristics of the individual played a very small role in the duration of unemployment and the change in wages at the end of the employment period. One exception that was both consistent with expectations and significant was the relationship between age and hours

TABLE 4

NUMBER OF INDIVIDUALS WHO RECEIVED
UNEMPLOYMENT INSURANCE BY
SITE AND SEX

	<u>Seattle</u>		<u>Denver</u>		<u>Totals</u>	
	<u>N</u>	<u>UI</u>	<u>N</u>	<u>UI</u>	<u>N</u>	<u>UI</u>
Male	213	132 62%	79	5 6%	292	137 47%
Female	166	61 37%	95	5 5%	261	66 25%
	<u>379</u>	<u>193</u> 51%	<u>174</u>	<u>10</u> 5%	<u>553</u>	<u>203</u> 37%

TABLE 5

NUMBER OF INDIVIDUALS WHO RECEIVED WELFARE
BY
SITE AND SEX

	<u>Seattle</u>		<u>Denver</u>		<u>Totals</u>	
	<u>N</u>	<u>WEL</u>	<u>N</u>	<u>WEL</u>	<u>N</u>	<u>WEL</u>
Male	213	41 19%	79	20 25%	292	61 21%
Female	166	70 42%	95	36 37%	261	106 41%
	<u>379</u>	<u>111</u> 29%	<u>174</u>	<u>56</u> 32%	<u>553</u>	<u>167</u> 30%

of search for males. Men spending an amount of time on job search were around 35 years old in Seattle and 44 years in Denver. Also consistent with our expectation was the finding that black males have a significantly longer duration of unemployment than do white males. Other demographic variables show no regular pattern and apparently are not major determinants of job search behavior or of the relative economic returns of the search.

Effect of Wages on Duration of Unemployment

As expected, the previous wage rate is a good indicator of the wage rate the unemployed individual will be willing to accept in a new job. We found that it has no significant effect on the duration of unemployment, but did effect job search. Among males, as the previous wage rate increased, the hours spent in job search decreased. The result for the duration of unemployment is contrary to a prior study reporting a significant positive relation between the duration of unemployment and the previous wage rate.⁶ Further research is required before a more definitive statement can be made regarding the relationship between the duration of unemployment and the wage rate.⁷

For all groups, those who enjoyed higher total earnings in the four quarters prior to unemployment also underwent shorter periods of unemployment. This finding is consistent with a more stable work relationship for those who had higher earnings. However, in light of this, it is somewhat surprising that the pre-unemployment wage rate bore no significant relation to the duration of unemployment. However,

the measure of past earnings over four quarters also reflects the extent of past unemployment, and so could be biased.

Effects of Nonwage Income on Duration of Unemployment

The impact of receiving welfare when an individual has a period of unemployment operates in the same manner as the receipt of unemployment insurance--it reduces the cost of being unemployed. We found receipt of welfare benefits had a more consistent relationship to duration of unemployment than to hours of job search.

The receipt of welfare significantly increased the duration of unemployment for males but not for females. It is surprising that this effect is not observed for females, who are not subjected to the same pressures as males to seek employment. We had anticipated that welfare would have stronger work disincentive for females.

A more efficient way to compare the relative impact of unemployment insurance with the impact of welfare is to calculate the income elasticities. The income elasticity is the percentage change in the duration of unemployment resulting from a percentage change in the nonwage income, where both the duration of unemployment and the non-wage income are evaluated at their means. The elasticities are shown in Table 6. The table shows an elasticity of 29%, at \$40.93 weekly benefit for Seattle males. Doubling this to approximately \$82.00 per week would lead to a 29% increase in the duration of unemployment. The results presented in Table 6 suggest that the response to unemployment insurance is much stronger than the response to welfare among the Seattle sample, but the

TABLE 6

UNEMPLOYMENT INSURANCE AND NONWAGE INCOME ELASTICITIES

	Males			
	Seattle		Denver	
	Mean Value	Elasticity	Mean Value	Elasticity
Duration of unemployment (days)	184.028		110.949	
Welfare income (dollars/day)	0.966	0.128	1.193	0.234
Weekly benefit amount (dollars/week)	40.929	0.290	10.873	0.106
Maximum benefits payable (total dollars)	1226.365	0.425	107.223	0.036
Benefit/earnings ratio (percentage)	0.370	0.138	0.098	0.097
	Females			
	Seattle		Denver	
	Mean Value	Elasticity	Mean Value	Elasticity
Duration of unemployment (days)	365.196		252.621	
Welfare income (dollars/day)	2.092	0.057	1.759	0.035
Weekly benefit amount (dollars/week)	19.552	0.175	6.432	0.010
Maximum benefits payable (total dollars)	586.104	0.142	77.806	0.024
Benefits/earnings ratio (percentage)	0.208	0.157	0.093	-0.009

opposite is true for the Denver sample. In addition, the specific form of the unemployment insurance variables appears to make a difference in the response among different groups. From this we conclude that the receipt of unemployment insurance or of welfare leads to an increase in the duration of unemployment, but the increase is relatively small and the magnitude of the response differs across groups. We found that all other forms of nonwage income had little impact on the duration of unemployment.

Job Search Success

The results show that the hours per week spent in job search and the duration of unemployment have no significant impact on the wage rate after the unemployment period ends. This is true of all groups in the study, except Denver females where an inverse relationship exists between the duration of unemployment and the post-unemployment wage. Further, demographic characteristics play a limited role in the wage adjustment of the individual.

Conclusions

The major conclusions of this analysis are that the receipt of unemployment insurance has a significant positive effect on the duration of unemployment for Seattle workers, but not on Denver workers. Among most groups, the receipt of welfare income also leads to longer periods of unemployment. From this we conclude that various forms of nonwage income--especially unemployment insurance and welfare payments--produce wage disincentive effects. Further, we can conclude that some of these nonwage income sources induce an increase in leisure consumption, as measured

by decreases in the hours spent in job search. We also found that neither increases in job search nor longer unemployment periods lead to a better paying job. Finally, the difference in response between Denver and Seattle point out the importance of the legal and/or economic conditions that exist in the region in which the individual resides.

NOTES

1. S. Marston, "The Impact of Unemployment Insurance on Job Search," Brookings Papers on Economic Activity (1975); A. Hoken and S. Horowitz, "The Effect of Unemployment Insurance and Eligibility Enforcement on Unemployment," Public Research Institute, Washington, D.C. (April 1974); R.I. Cresslin, "Unemployment Insurance and Job Search: Empirical Relationships and Interpretations," (Comm. of Ways and Means, U.S. House of Representatives, July 1975); R. Schmidt, "The Theory of Search and the Duration of Unemployment," (Working Paper No. 7317, University of Rochester Graduate School of Management, Rochester, August 1973); G. Chapin, "Unemployment Insurance, Job Search, and the Demand for Leisure," Western Economic Journal (March 1971); C. Lininger, "Unemployment Benefits and Unemployment Duration" (University of Michigan Institute for Social Research, Ann Arbor, 1963).
2. E.S. Phelps, et al. (eds), Microeconomic Foundations of Employment and Inflation Theory, New York: W. W. Norton & Co. (1970).
3. The insured unemployment rate is an administrative count of the ratio of workers who are receiving unemployment insurance during a reference week to a moving average of the number of workers who are in covered (insured) employment.
4. Table 1 was taken from "Staff Data and Materials on Unemployment Compensation," Committee on Finance U.S. Senate, June 6, 1975, Tables 2, 3, and 9.
5. The statistical estimates and other details are available in the unabridged version of this paper: Henry E. Felder, "Unemployment Insurance and Duration of Unemployment" (SRI International Mimeo, Stanford, May 1978). The following variables were used in the regression analysis: duration of unemployment (days); postunemployment period wage rate (dollars/hour); hours spent in search per week; age of individual; =1, if Black (percentage); years of schooling; pre-unemployment quarters; net worth (=0, if total earnings in the four pre-unemployment quarters; net worth (=0, if negative); average daily amount of welfare income received during unemployment period (dollars); average daily amount of all other nonwage income received during unemployment period (dollars); weekly benefit amount (maximum) of unemployment insurance (dollars); maximum benefits payable at start of unemployment period (dollars); ratio of WBA to weekly earnings = WBA/PRW (Average weekly hours of work or 40, if hours missing); =1, if public employment agency used; =1 if private employment agency used (percentage); =1, if friends and relatives used (percentage); =1, if want ads used (percentage); and unemployment rate in SMSA at beginning of unemployment period (percentage).
6. R. L. Crosslin, "Unemployment Insurance and Job Search: Empirical Relationships and Interpretations," (Comm. of Ways and Means, U.S. House of Representatives, July 1975).
7. A major difficulty with the interpretation of the interaction of the wage rate and the duration of unemployment is the definition of the wage rate should go into the estimation. The model used here suggests that the prior wage is a proper specification for estimation. Another possibility is an estimated wage; this approach was also tried in this model but with no success.

THE SIME/DIME MANPOWER PROGRAM

by

Arden R. Hall
Economist
SRI International

The Seattle and Denver Income Maintenance Experiments (SIME/DIME) are the only ones among the federally-funded income maintenance experiments to include a manpower program. A manpower component was included in SIME/DIME in order to determine how much the expected decline in work response due to income support could be offset by providing information about the labor market and by subsidizing the cost of training and education. The general presumption was that raising wage rates through training and lowering job search costs would induce greater individual work effort.

The functions that the manpower component were to perform derived from three hypotheses:

- (a) Many individuals have deficiencies in their knowledge of the labor market, and in their understanding of their own capacities in relation to the labor market.
- (b) Lack of information results in these individuals making less than optimal decisions with regard to their own work plans.
- (c) Subsidization is necessary to achieve the socially optimal investment in education and training among low-income individuals.

This paper describes the SIME/DIME manpower component, provides a preliminary analysis of how it was used, and a preliminary assessment of its effects, measured by the number of people taking training.

THE MANPOWER TREATMENTS

One of three different manpower treatments, or control status, was assigned to each family in the experiment. Assignments were made randomly within stratified classes of families (defined by race, headship and economic status). All three manpower treatments included unlimited free use of a counseling center for all members of the eligible family aged 16 or over. These centers, operated by community colleges and staffed by experienced professional counselors, provided information about educational and training opportunities and financial support, as well as information about the labor market. Considerable resources were committed to the centers. If a client was interested in a particular occupation, a specialist attached to counseling was available to prepare a report on that occupation.

Counselors attempted to follow a consistent pattern in their interactions with clients, which included three definite steps: (1) self-assessment by the client; (2) provision of information by the counselor; and (3) formulation of a plan of action. These steps were the focus of five group counseling sessions, with additional individual counseling to supplement the process.

The culmination of counseling was the client's plan of action, a formal statement of intent, based on the counseling. It would discuss a plan for training, job search, or job retention, or a combination of these elements, spelled out in some detail. After this plan was written, the counselor might provide some assistance to the client in implementing it,

but usually would not provide intensive counseling. If the client accomplished the objective of the plan, or decided to change it, the counseling process could begin anew.

SIME/DIME made an incentive payment of \$5 for the completed initial interview, as well as payments of \$5 per session to cover inconvenience and transportation costs, up to a maximum of \$25.

The counseling centers also administered subsidy programs that were available to two of the three subsamples receiving manpower treatments. One treatment provided 50% and the other 100% of all direct costs of schooling or training--including tuition, fees, required books and equipment, and child care during the parent's school attendance. Direct cost did not include living stipends or payment for time taken off from work.

The subsidy was available for any course of schooling or training so long as it represented preparation for an occupation or career. This limitation was interpreted liberally, and in fact included everything from driver's training to graduate business school. In general, participants were allowed to choose the educational institution, but the rules required pursuit of the least expensive, adequate option available. Thus if a certain course of training were available at a local school but the participant preferred to attend a school in another area, his or her subsidy would be based on costs at the local school.

Eligible participants were required to write a plan of action before they could receive subsidies. (This requirement provided some administrative control over the subsidy program, and incidentally provided us a great

deal of information about the intentions of those who responded to the subsidy program.) Counseling centers also audited claims for subsidy payment and disbursed funds.

UTILIZATION OF THE MANPOWER PROGRAM

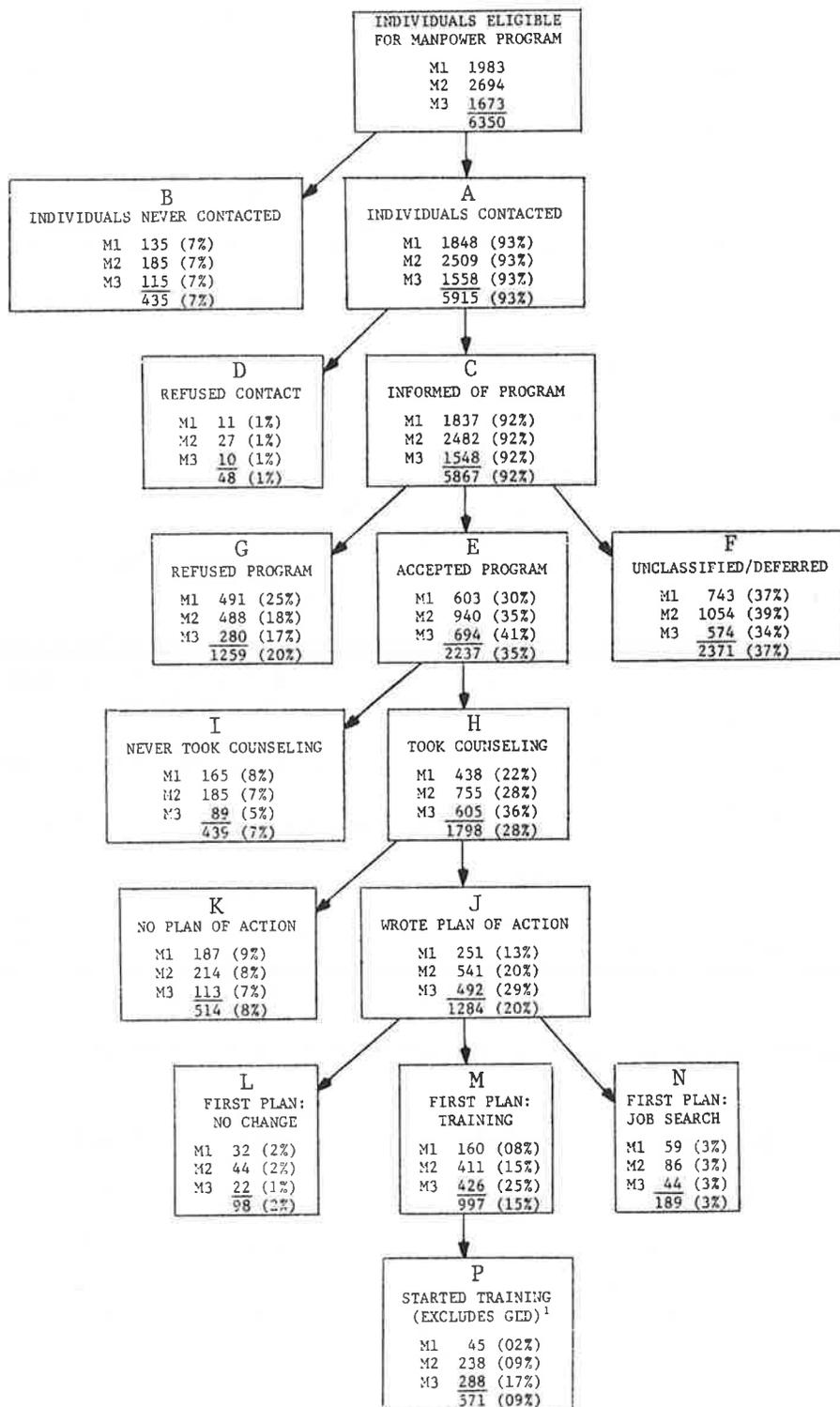
The manpower programs were available to 6350 participants. To present a brief summary of their responses, we have combined data from Seattle and Denver, and have concentrated upon the initial response to the program. The initial response represented the greatest use of the program and provides a clear picture of the cause-effect relationship between the program, and schooling and other outcomes. A discussion of subsequent responses would considerably lengthen the paper so it has been omitted.

Figure 1 illustrates the flow of people through the steps in the counseling and training process, and gives the number of people taking each step, as well as the percentage of the original treatment group that made a particular response. Thus 22% of all eligible for the M1 treatment (counseling only) or 438 people, took counseling.

Of the 6350 people eligible for manpower programs, all but about 8% were contacted by counselors and given an explanation of the program. At this point about 20% of the sample decided not to use the program. This group was contacted again on a regular basis to remind them of their eligibility, and some subsequently used the counseling center. Some clients were "deferred" by the counselor making the contact, generally because they were not yet 16. Of those who accepted the program, a high proportion actually used the counseling center. The proportion of those who used the

FIGURE 1

RESPONSE TO INITIAL CONTACT WITH THE MANPOWER PROGRAM



¹Data are not available for the 104 individuals who selected the GED as the only training objective.

center was highest among those with a 100% subsidy for education (the M3's) and lowest for those who were eligible only for free counseling (the M1's).

Response of Families to Counseling-Only Program

Writing a plan of action represented completion of the counseling process for the M1's, but it was a prerequisite to subsidies for M2's and M3's. Not surprisingly, a greater proportion of M2's and M3's who took counseling completed a plan of action.

We have classified the goals reflected in these plans as follows:

- * No change. The individual decided as a result of counseling that no change in his career plan was desirable.
- * Training. The individual decided to enroll in some kind of schooling or job training course. This course of action may have included part time or full time work as well as education.
- * Job search. The individual decided to enter the labor market or look for another job if he was already employed.

As Table 1 indicates, training was the usual goal of the plan of action but the proportion of those choosing training was much higher for those who could expect a subsidy. For all treatment groups, no change was the least likely goal to be chosen.

TABLE 1
DISTRIBUTION OF PLANS OF ACTION BY GOAL AND TREATMENT

	<u>No Change</u>	<u>Training</u>	<u>Job Search</u>	<u>Total</u>
M1	32 (13%)	160 (64%)	59 (23%)	251
M2	44 (8%)	411 (76%)	86 (16%)	541
M3	22 (4%)	426 (87%)	44 (9%)	492

Among those who chose training as a goal, only a fraction actually started schooling or job training. Unfortunately, no record was kept of those who pursued a General Education Degree, but only 104 or 10% of the training plans had this as the only goal. Thus, the number shown in Figure 1 for training is only slightly too low.

The numbers given in Figure 1 only represent the response to initial contact with the manpower program; total response was considerably greater. An example of the difference will illustrate the magnitude of subsequent response. Of 3678 individuals who did not accept the program at the initial contact, 1022 or 28% took counseling after a subsequent contact. Some of these people prepared plans of action and enrolled in a school or training course.

The most reasonable measure of usage for the M1 treatment is attendance in a counseling session. Table 2 gives the participation rate of M1's by sex, family type and occupation. Table 2 shows a slightly higher participation rate by women, and we can see that it is the single female heads rather than the wives who are responsible for the higher overall participation of women. Table 2 also shows that members of female-headed families generally made more use of the counseling center than members of husband-wife families, except for Chicanos, who have equal overall participation rates for both groups. Table 2 also implies that, overall, whites made slightly more use of the centers than did blacks or Chicanos.

The greatest differences in participation rates are by occupation. None of our nine farmers or farm laborers used the counseling center after

TABLE 2

COUNSELING PARTICIPATION OF M1'S FOLLOWING
THE FIRST CONTACT, BY SEX, RACE, FAMILY TYPE AND OCCUPATION

<u>SEX</u>	<u>PARTICIPATION RATE</u>			
Females	24%			
Males	20%			

<u>RACE AND FAMILY TYPE</u>	<u>PARTICIPATION RATE</u>			
	Male Heads	Female Heads	Other	Overall
Black female headed families	--	40%	8%	26%
Black husband & wife families	24%	22%	7%	19%
White female headed families	--	39%	9%	28%
Black husband & wife families	23%	24%	5%	20%
Chicano female headed families	--	35%	10%	22%
Chicano husband & wife families	32%	21%	4%	22%

<u>CURRENT OCCUPATION</u>	<u>PARTICIPATION RATE</u>
Professional	28%
Managerial	18%
Sales	23%
Clerical	30%
Craftsmen	26%
Operatives	31%
Operatives, transport	23%
Laborers	28%
Farmers and farm laborers	0%
Service workers	32%
Private household workers	32%
Never worked	6%

their initial contact while over 30% of the clerical workers, operatives, service workers, and private household workers talked to a counselor at least once. Among workers there is no striking pattern of participation rates by occupation, although it may be significant that the participation was very low among managers. The very low rate of participation among those who had never worked may reflect their intention not to work in the future and a decision that employment counseling was of little use to them.

Response to the program was relatively rapid and peaked in the period 15 to 30 days after the initial contact. Almost half of those who had accepted the program participated in a counseling session within 30 days and two thirds had talked to a counselor within a half year after contact. Only 42 people or 10% of those who eventually visited the counselor did so after an interval of more than six months. Those who felt a need for counseling usually took advantage of the counseling program quickly. The total response by the entire sample of M1's who took counseling was also rapid: only about 16% visited a counselor after six months.

Response of Families to Training Subsidies

The definition of usage of the education subsidy program (M2 and M3 treatments) will be the enrollment of an eligible individual in a course of formal schooling or job training. Among those eligible, about 12% received a subsidy. (After additional contacts, the participation rate rose to 23%.) With very few exceptions, the response to the 100% subsidy was considerably stronger than the response to the 50% subsidy.

Table 3 shows the characteristics of the 12% responding to the initial contact by sex, race, family type and occupation. The pattern is quite similar to that for those eligible for counseling only. The table shows a slightly higher enrollment rate among women than men and the highest rates for female heads of families. Blacks and whites seem to have participated in the subsidy program in greater numbers than Chicanos. While some of these differences in participation rates are large in percentage terms, there are no large absolute differences in participation rates evident between men and women or among family types.

When the eligible population is broken down by occupation, differences in participation in the education subsidy program become more pronounced. As in the counseling program, few people who have never worked made use of the program. Apparently, people who did not intend to work had no use for employment-related training, even when subsidized.

A somewhat surprising pattern of participation rates can be seen among those who had an occupation. It would seem that those with the fewest skills would be most likely to want training, but the results suggest the opposite. Participation rates are highest for professionals and managers and lowest for laborers, farmers and farm laborers, and private household workers.

Two possible reasons for these results are suggested. First, a person with relatively high ability will probably receive a greater ultimate return from schooling so he will be more likely to participate in the program. Previous occupation is a good predictor of ability and the professionals and managers probably had greater average ability than the laborers and

TABLE 3

RATE OF PARTICIPATION IN TRAINING BY M2'S AND M3'S
FOLLOWING THE FIRST CONTACT, BY SEX, RACE, FAMILY TYPE AND OCCUPATION

<u>SEX</u>	<u>PARTICIPATION RATE</u>		
	<u>M2</u>	<u>M3</u>	<u>M2 and M3</u>
Female	10%	18%	13%
Male	8%	17%	11%

<u>RACE AND FAMILY TYPE</u>	<u>PARTICIPATION RATE</u>							
	<u>Male Heads</u>		<u>Female Heads</u>		<u>Other</u>		<u>Overall</u>	
	<u>M2</u>	<u>M3</u>	<u>M2</u>	<u>M3</u>	<u>M2</u>	<u>M3</u>	<u>M2</u>	<u>M3</u>
Black female headed families	--	--	17%	29%	4%	7%	11%	18%
Black husband & wife families	9%	16%	10%	22%	3%	8%	8%	16%
White female headed families	--	--	16%	27%	2%	10%	10%	21%
White husband & wife families	11%	24%	9%	21%	1%	4%	8%	20%
Chicano female headed families	--	--	14%	13%	5%	2%	9%	7%
Chicano husband & wife families	8%	18%	11%	7%	1%	10%	8%	12%

<u>CURRENT OCCUPATION</u>	<u>PARTICIPATION RATE</u>		
	<u>M2</u>	<u>M3</u>	<u>M2 and M3</u>
Professional	23%	41%	30%
Managerial	13%	37%	20%
Sales	11%	21%	14%
Clerical	15%	25%	19%
Craftsmen	13%	19%	15%
Operatives	9%	19%	13%
Operatives, transport	10%	40%	19%
Laborers	9%	8%	9%
Farmers, farm laborers	0%	20%	11%
Service workers	11%	23%	16%
Private household workers	9%	15%	11%
Never worked	2%	6%	4%

household workers, and therefore could expect greater benefits from additional training. Second, the SIME/DIME income limitation (of less than \$11,000) makes the professionals and managers in our sample atypical of their counterparts outside the sample. They must be people who have been relatively unsuccessful in their occupation, and therefore likely candidates for additional training. People in our sample in less skilled occupations should be more typical of their occupation and less likely to feel a need for training to improve their position in their occupation. Unfortunately, both of these reasons for the high participation rates of professionals and managers are speculative and additional research will be required to obtain a better understanding of participation differences by occupation.

If an individual accepted the program, he or she was still unlikely to use the subsidy. Of those who accepted the program, 61% never took training. If the individual did take training he or she was most likely to begin soon after contact. The response peaks in the first quarter; 164 people or 28% of those who took a training program began more than a year after the contact; only 41 people or 7% began more than two years after contact. The total response by the entire sample of M2's and M3's who took training was similar to the group who accepted at the first contact. Most enrolled in the first quarters; about 30% of those who took training did so after one year and 8% began two years after contact. This result suggests that the success of a manpower program in attracting participants can be evaluated soon after it begins.

Goals of Individuals in the Manpower Programs

An indication of the way people intended to use the manpower programs can be obtained by studying their plans of action. Most plans included an occupational goal and a degree objective, as well as information about the educational institution to be used. Of the 1,284 individuals who wrote a plan of action, 997 (78%) stated some type of training as their objective. Table 4 gives a picture of the occupations held at contact and occupations desired.

TABLE 4
OCCUPATION AT CONTACT AND OCCUPATIONAL GOAL OF TRAINING

OCCUPATION GROUP	Occupation at Contact		Training Goal	
	NUMBER	PERCENT	NUMBER	PERCENT
Professional	63	7%	364	41%
Managerial	21	2%	57	6%
Sales	25	3%	19	2%
Clerical	208	23%	129	14%
Craftsmen	62	7%	104	12%
Operatives	134	15%	42	5%
Operatives, transport	36	4%	6	1%
Laborers	45	5%	7	1%
Farmers and farm laborers	1	*	1	*
Service workers	227	26%	110	12%
Private household workers	17	2%	1	*
Never worked	54	6%	--	--
Undecided	--	--	53	6%
TOTAL	893	100%	893	100%

*Less than 1%

The 104 individuals who elected to pursue a General Education Degree as their only training objective are not included in the above table. From Table 4 it appears that the individuals in our sample overwhelmingly desired

to prepare themselves for jobs in the professional occupations. Of the individuals in our sample who wrote training plans, 41% specified occupations in professional fields while only 7% of the individuals held jobs in that area at the time of contact. There were increases in the managerial and crafts areas also. The distribution of individuals who wanted to prepare for clerical, operative, and service fields is lower than the distribution of individuals in those fields at contact. These figures point to the intent of the sample members to upgrade their occupational status by means of formal training.

Table 5 shows the degree objective of training and the type of institution at which the individual planned to pursue that degree. Most manpower program participants decided to pursue the AA degree, to upgrade existing skills, or to enroll in or continue a college program, with equal numbers pursuing a degree or no degree. Other objectives--such as graduate school, medical training, or training for skilled or semi-skilled trades--were much less common.

Table 5 discloses a difficulty in the analysis of the plans of action. Over a third of those who wrote a plan of action did not specify the institution they wished to attend. These plans are obviously incomplete. Moreover, to receive a subsidy those eligible for M2 and M3 treatments were required to specify an institution in their plan. This requirement reduced incidence of plans with unspecified institutions among M2's and M3's but even among those eligible for a 100% subsidy, 26% did not specify an institution. The high percentage given here may reflect the way in which plans

TABLE 5

DEGREE OBJECTIVE OF TRAINING AND TRAINING
INSTITUTION SELECTED FOR FIRST PLAN

<u>Degree</u>	<u>Manpower Treatment</u>			<u>Total</u>	<u>Percent</u>
	<u>1</u>	<u>2</u>	<u>3</u>		
General Education Degree only	20	45	39	104	10%
AA	20	81	89	190	19%
BA/BS	28	46	60	134	13%
MA/Phd/MD	1	3	11	15	2%
RN/Other Medical	4	15	20	39	4%
Skills Upgrade	41	82	73	196	20%
Trade Union or Apprentice	3	5	2	10	1%
Skilled or Semi-Skilled License/Certificate	14	49	44	107	11%
College, no degree	18	59	54	131	13%
Skilled or Semi-Skilled, No License	5	17	26	48	5%
Other	<u>6</u>	<u>9</u>	<u>8</u>	<u>23</u>	<u>2%</u>
TOTAL	160	411	426	<u>997</u>	100%
<u>Training Institution</u>					
Community College	19	137	170	326	37%
College Or University	9	35	51	95	11%
Technical/Trade	1	10	10	21	2%
Other/Out of State	<u>16</u>	<u>55</u>	<u>55</u>	<u>126</u>	<u>14%</u>
Not Specified	140	366	387	<u>893</u>	100%

Note: Degree other than "General Education Degree" may also include a General Education Degree. There were 104 individuals listing both.

were written. Plans were written when the counselee had a clearly specified goal, but not necessarily the means to attain it. Those who did not specify an institution may have been unable to immediately pursue their objective.

Among those who specified a training institution, the overwhelming favorite was a community college. This is a surprising result for those eligible for subsidies because these people, especially the M3's, could afford to pay the higher costs of more expensive institutions. It is particularly surprising because, as explained below, the subsidies did induce increased enrollment. The fact that the manpower programs were administered by the community colleges may be one explanation for the popularity of community colleges. Counselors were instructed not to influence clients in their decisions, but they may have been more familiar with community college programs.

Since the community college was the most popular training institution, the amounts of subsidy paid were relatively small. Of the 526 M2's and M3's who started training, 514 received at least one subsidy payment. The average subsidy under the 50% subsidy program was \$382 and under the 100% program, \$956. On the average, the full cost of training for the latter was \$192 more than the average for the group receiving only a 50% subsidy.

Of all those preparing plans of action, there were 189 individuals who first planned to search for a job. Table 6 gives the distribution of occupations these individuals had at contact and the distribution of occupational goals of the job search plan. People apparently hoped to move up the occupational ladder, with more planning to search for jobs in profes-

sional, managerial, clerical and crafts occupations than were members of those occupations at contact. It should be noted that 7% of those who planned to search for a job had never worked previously. These people represent only 1% of those in the entire sample who had never worked. About 4% of this group planned training.

TABLE 6
OCCUPATION AT CONTACT AND OCCUPATION
OF JOB SEARCH FOR FIRST PLAN

OCCUPATION GROUP	Occupation at Contact		Occupation of Search	
	NUMBER	PERCENT	NUMBER	PERCENT
Professional	6	3%	22	12%
Managerial	7	4%	12	6%
Sales	8	4%	4	2%
Clerical	34	18%	46	24%
Craftsmen	26	14%	29	15%
Operatives	23	12%	16	9%
Operatives, transport	6	3%	5	3%
Laborers	17	9%	15	8%
Farmers and farm laborers	--	--	1	*
Service workers	45	24%	28	15%
Private household workers	4	2%	--	--
Never worked	13	7%	--	--
Not Specified, other	--	--	11	6%
TOTAL	189	100%	189	100%

* Less than 1%.

Finally, there were 98 individuals who wrote plans which specified no change.

RESPONSE TO THE MANPOWER PROGRAM

The ultimate objective of these manpower programs was to improve the client's success in the labor market and they must ultimately be evaluated in these terms. However, success in the labor market varies widely for

reasons apart from the manpower program. It will take quite a long period of observation and more data than now available before the labor market effects can be measured with any accuracy.

Until then, we have concentrated our analysis upon the demand for education that can be attributed to the experiment. Results obtained so far indicate that the manpower program did induce people to take additional schooling and training.

Theoretical Framework

The analysis regards education as a type of investment. The cost of the investment is the total of lost earnings while attending school and direct costs. The individual's gain will be the increased income after training. In making a decision about school attendance, the individual must determine whether the cost of the investment is justified by the dividends to be gained. The thing that people invest in has been called "human capital" by economists.

An individual's decisions about training determine the level of their human capital. Investment in human capital--in the form of training and education--goes on all the time, independent of any manpower experiment. People go to school until they decide that the cost, including lost earnings, outweighs the benefits. When that point is reached they leave school and devote themselves to work. The manpower program, and particularly the subsidies, should induce people to take additional schooling or training by lowering the direct costs of the investment.

However, the experiment was not expected by the participants. Thus, the strength of the effect of the subsidy should depend on the point in the individual's career it was offered. For the person who is nearly finished with his planned schooling, the effects of the inducement should be strong. A younger person is likely to have planned to attend school before he knew of the subsidy. Since the experiment only lasted three years for most people, the subsidies would not be expected to have the effect upon younger people that a permanent program would have. For the oldest strata in the sample, one might expect the lowest response to the subsidy. While the cost of attendance should be the same for older people (or greater if he or she has a higher wage), the dividends from the investment will not flow for so long, because this oldest group has fewer remaining years of work in which to reap the benefits of education. However, some additional factors would counter this effect. Those who have left school several years ago are likely to have seen the value of their skills deteriorate in the market. Also the average education of the work force has been rising. The worker who left school several years ago finds he has increasingly well-trained competition in the labor market, even if he has maintained his own skills. For these reasons older people might be more responsive to the subsidies than would at first be expected.

Results

This digression about the theory underlying our analysis of the demand for schooling induced by the manpower subsidies simplifies the information summarized in Table 7. This table presents estimated effects for each of

TABLE 7

PREDICTED EFFECTS OF THE MANPOWER
PROGRAM UPON DEMAND FOR FORMAL
SCHOOLING

		Controls (predicted quarters of attendance)	Predicted experimental effects			Controls (predicted quarters of attendance)	Predicted experimental effects		
			<u>M₁</u>	<u>M₂</u>	<u>M₃</u>		<u>M₁</u>	<u>M₂</u>	<u>M₃</u>
		<u>Seattle Husbands</u>				<u>Denver Husbands</u>			
In School	16-25	4.984	-.579	.125	.573	4.186	-.592	.131	.616
	26-45	3.659	.840	1.401*	1.852*	2.869	.818	1.397*	1.883*
	46+	2.294	.766	1.333*	1.824*	1.652	.666	1.186*	1.657*
Out of School	16-25	.717	-	-.048	.163	.438	-.145	-.033	.114
	26-45	.309	.221	.339*	.547*	.172	.141	.220*	.363*
	46+	.104	.095	.149*	.252*	.052	.054	.086*	.149*
		<u>Seattle Wives</u>				<u>Denver Wives</u>			
In School	16-25	3.269	-.223	-.095	1.335	2.862	-.213	-.091	1.320
	26-45	3.686	-.319	-.719	.704	3.266	-.310	-.693	.699
	46+	3.233	-.309	-.690	.698	2.827	-.295	-.653	.680
Out of School	16-25	.207	-.054	.273*	.879*	.150	-.041	.215*	.719*
	25-45	.280	-.084	.138	.691*	.206	-.064	.109	.565*
	46+	.201	-.063	.106	.556*	.146	-.048	.082	.446*
		<u>Seattle Female Heads</u>				<u>Denver Females Heads</u>			
In School	16-25	5.135	.474	.243	.787	4.604	.504	.256	.845
	26-45	4.526	.888	.157	.758	3.977	.922	.159	.784
	46+	2.947	.915	.152	.772	2.443	.873	.142	.734
Out of School	16-25	.529	.029	.717*	1.134*	.373	.022	.571*	.922*
	26-45	.354	.135	.502*	.832*	.241	.100	.385	.652*
	46+	.108	.052	.214*	.378*	.068	.036	.150*	.272*

* Significant at the 5% level.

the manpower treatments upon the amount of formal schooling (as distinguished from job training) taken during the first two years of the experiment. The table presents the predicted quarters of schooling (out of a possible eight) we estimated that individuals would take in response to the experimental treatment, as well as predicted levels of school attendance for controls. The predictions are broken down by experimental site, by type of person, age, and by whether or not the person was in school when the experiment began. This last attribute was included because we hypothesized that it might be easier for a person who was already in school to take additional courses than for someone who was out to begin again. The asterisks indicate estimates which were statistically significant, that is, where differences were not due to chance alone.

One thing is immediately apparent from the table. The M1 treatment (counseling only) did not have a consistent effect, but the two treatments that provided subsidies clearly induced people to take additional schooling. The effect can be seen for husbands, wives, and single family heads of families, both for Seattle and Denver, for all age groups, and for both those who were already in school and those who were not. Also, as we would expect, the more generous M3 treatment has a consistently larger effect than M2. For some groups the 100% subsidy induced more than one quarter of additional school attendance for every person in the group.

When the table is examined more closely some additional insights emerge. First, the effects of the subsidies are quite consistent between the two experimental sites.¹ However, there are differences in effect

depending upon whether the individual was in school before the experiment began or not, and these depend upon sex. For men, the treatments seem to have had a greater effect upon those who were already in school than on those who were not. For women, both wives and single female heads, the pattern of effect is the opposite. The subsidies had more effect in inducing women back into school than it did in inducing those already in school to attend longer. The effect of the experiment upon different age groups also varies by sex. The strongest effect of the subsidies was upon young women, and the magnitude of the effect declines with age. Thus the response of women coincides with our simple picture of individual decision-making. Unfortunately the men in the experiment were not so cooperative. For them the subsidies have the least effect upon the youngest category of men and the greatest effect upon those in their prime working years. Thus, for men, the subsidies seem to have been most useful as a means to upgrade skills already in use in the labor market.

FUTURE ANALYSIS

We are now analyzing three years of experimental data and examining the effects of the subsidies on attendance in training courses as well as in formal schooling. A separate study is being made of those who were members but not heads of experimental families and their reaction to the manpower program. Beyond that, we plan to examine the data in ever greater detail, to learn what kinds of people made use of the program and what kinds did not, the types of schooling and training courses taken, and the occupations for which people were preparing. We will attempt to determine whether

there were existing manpower programs in Seattle and Denver that affected response to the SIME/DIME program. As the experiment was an unexpected event for the participants, we also want to study the time pattern of their response, in order to be able to predict what the response would be to a permanent program. Finally, we will follow families for two years after the experiment ends and will study the effect of schooling and training upon labor market success among those who were influenced by the experiment.

NOTES

1. The consistency in these results is enforced somewhat by the way the underlying model is specified. However, before this particular form of the model was chosen, earlier research had shown consistency of experimental effects in the two sites.

DISCUSSION

ROBERT SPIEGELMAN: What further analysis are you going to do of the persons who changed their plan of action?

HALL: Well, we will look beyond the initial response: that research has been proposed and will be done early next year. It will be similar, but expanded version of what we presented here.

R. EMMICK: Did you say that you're going to analyze the responses to the training program based on the different assignments in the SIME/DIME program?

HALL: We've done that implicitly. The assignment was based upon family type, race, site, and economic status at the time of the experiment, and we have included all those variables. We've found not much of an effect due to initial economic status, and we've come to the conclusion that that is probably not important. Now there are differences in effect by race and family type, and those can be interpreted either as an effect of race and family type, or to some extent an effect of the assignment. So, yes, I think that the measurements of the effect are pretty accurate, but there is this problem of distinguishing between effects of the assignment and the overall effects of these variables.

MARY JANE CRONIN: What were the rules as far as the child care and expenses were concerned: whatever incurred?

HALL: Let me see if I can remember. I think that it was available to a mother or a father for school attendance. In terms of amount, I think there were some limitations, but they were generous.

MOSES NEWSOME, JR.: Was this supposed to be good counseling, or typical counseling?

HALL: Well, it was provided by counselors, some of whom were associated with the community colleges. My impression is that there's really a wide variety of counseling programs in terms of their goals and methods and so forth, and it certainly was not intended to be bad counseling, but it would be better or worse depending upon your view of the proper role of counseling.

SPIEGELMAN: But it wasn't any extraordinary innovative counseling.

EMMICK: Was migration ever suggested?

SPIEGELMAN: It certainly wasn't precluded.

EMMICK: I know you wanted people to stay in the cities.

CRONIN: Did you restrict the labor market information to any locality or area?

SPIEGELMAN: No, I think the labor market reports tended to focus on locality, but they could have included information about other labor markets.

HALL: I think that depended on the occupation that the person was investigating.

JOANNE BREKKE: I'm thinking particularly of women. What, if any, additional things were done to encourage a person to take a try at working?

SPIEGELMAN: I wouldn't say "encourage", but the counseling program was designed to be a series of counseling sessions in which the initial sessions had the term "self-assessment". There was a considerable amount of information gathered about the individual.

BREKKE: In the training, was there any time limit? When somebody's looking for a Master's or Doctor's degree, they must make a long commitment into the future.

HALL: There wasn't any commitment to extend the experiment to a longer period. They would receive the subsidy for the years that they were eligible, and when they were disenrolled, the subsidy would end. There was a program with the duration of three years and another of five years, and the subsidies followed that time frame. It turns out there was no 100% available to people who were on five-year treatment, so there was no M3 group in the five-year program. So our program would not encourage you to go into a Ph.D. program or medical training, or something like that.

SPIEGELMAN: I can't visualize at the moment a program of that degree of generosity.

DR. NEWSOME: Could you just elaborate on this training? What was the nature of the training choice that was available to a person? Could you get on-the-job training?

HALL: Yes, I think so, if it was a formally recognized program. One could not start up one's own on-the-job training program and ask for a subsidy.

DR. NEWSOME: In predicting the utilization of a general program, would you include costs on the education side or do you just take the tuition costs of, say, community colleges?

HALL: We have to consider that, and we have talked about it. We're not at the point of trying to calculate what those total costs would be.

DR. NEWSOME: You were indicating that older people might be less likely to invest in education. I think that is changing because of trends affecting senior citizens. I know at Howard University we have a lot of older people going back to school, seeking second careers or just psychological gratification as opposed to economic benefits.

HALL: I think you make a couple of points. First of all, the model that underlies this looks at education and training as an investment. In other words, we're assuming here that people don't derive any satisfaction from going to school or taking training except through higher wage rates and maybe steadier employment. That's a big assumption, but it seems to be the most reasonable assumption, if one has to make any assumptions, since we limited the subsidy to employment-related schooling and training.

Second, the model discussed here is relatively simple. It does not take in the entire variety of human experience and motivation. Since the results seem to suggest that we're quite successful in inducing older people to go to school, we're trying to think about theoretical models that seem to be more compatible with that result.

DR. NEWSOME: Did you analyze full time versus part time enrollment?

HALL: No, we haven't done that yet, essentially because of data limitations. But I would like to tackle it.

IV

COST IMPLICATIONS OF
THE LABOR SUPPLY RESPONSE

GENERALIZING EXPERIMENTAL RESULTS
WITH MICROSIMULATION

by

Harold Beebout
Director, Policy Studies Division, MPR

and

Myles Maxfield, Jr.
Senior Economist, MPR

This paper continues the discussion begun in the paper by West and describes in greater detail how we can use data from the Seattle and Denver Income Maintenance Experiment (SIME/DIME) to predict national program costs, caseloads and distributional effects. SIME/DIME results have been heavily used in the evaluation of legislative proposals for welfare reform. For example, these results have been useful in predicting cost and caseload of the Ullman welfare proposal now being considered by Congress. This is due, at least in part, to the development of microsimulation as a tool for generalizing experimental results.

The generalizations are based on the deceptively simple idea that we can take estimates based on SIME/DIME data, as described in the paper by West, and apply them to families representing the entire country, and to proposed income maintenance plans.

Unlike social experiment, proposals for new social programs are usually national in scope, are to be implemented at a current or future time, and often contain a large number of potential variants. Evaluation of a

proposed income maintenance program requires information about national impacts in a future time period for a program that probably does not contain the same features as those tested in the experiments. Such general predictions must somehow be derived from highly specific results--such as, for example, the response of black male secondary members of families with income of \$3,000, or some similar narrow result. This paper presents some options to the problem and outlines the microsimulation approach chosen in SIME/DIME.

APPROACHES TO GENERALIZING

Strategies for generalizing experimental results over time, place, and program variants necessarily start with the results estimated in specific experiments. The strategy that probably seems most natural to researchers is generalizing directly to population strata with which the regression equations were estimated. A national estimate can then be based on the proportion of each experimental strata within the national population for the analysis year.

This approach sounds straightforward, but in practice it runs into at least four serious difficulties. First, a stratification that arrays the population by existing support level and benefit reduction rate, by family status, and by employment status, just to mention a few characteristics, yields hundreds and possibly thousands of subgroups, or cells. Second, even with tens of thousands of cells, much information is lost. Continuous variables, such as age, have to be grouped and coded into categories and it is nearly impossible to account for all the interactions between the response of an individual person to a program. Third, each time it is desirable to

change a stratification, which may be each time a program variant is evaluated, the whole set of cells must be remade. Fourth, the detailed data on the national population required to generate the estimate of the number of persons and families in each subgroup for the analysis year is only available from projections made by microsimulation models.

An alternate methodology is microsimulation. The approach is conceptually simple. The determinants of the behavioral response to a negative income tax (NIT) are statistically estimated using the experimental sample, as described in the paper by West. These regressions are used to predict the expected response to an NIT of each individual in a second sample representative of the nation. The predicted responses are then inflated by the sample weight and summed over the sample to produce national or regional estimates.

Two specific microsimulation models have been used for making program estimates based on responses estimated from SIME/DIME data. A model developed at the Department of Health, Education and Welfare (DHEW), discussed below in the paper by Kasten, Greenberg and Betson, was used to estimate the cost and caseload of the Better Jobs and Income Program (BJIP). A more flexible generalization model has been constructed within the Micro-analysis of Transfers to Households (MATH) system used by a number of government agencies, discussed below in the paper by Edson and Maxfield.

MATH may be used to estimate the budget cost and distributional impacts of most income transfer programs, including Aid to Families with Dependent Children (AFDC), Supplemental Security Income (SSI), Food Stamps, and federal and state taxes on income. The model, because of its general transfer program modeling capabilities, can also be used for analysis of proposed transfer

programs such as the BJIP and special-purpose current transfer programs such as the School Lunch Program. Estimates can be prepared for future years by aging the data base--that is, projecting the sample of families forward in time.

Sample weights and number of children per family are adjusted according to demographic projections by the Bureau of the Census. Income and work experience variables are adjusted to be consistent with either the economic growth rates or the absolute levels predicted by a macroeconomic model.

Next, a transfer program is simulated on the national, and perhaps aged, sample of individual family records. If, for example, one wants to estimate the impact of replacing AFDC payments and food stamps with a uniform NIT, one must first simulate current transfer programs since they are not recorded accurately in currently available national data bases. In this example, the sample of families is passed through a set of computer statements designed to replicate the major features of each state's AFDC eligibility procedures. The families eligible for an AFDC payment are flagged and their potential benefits computed. The food stamp entitlement of the family is similarly determined. The subsamples of eligible families which participate in each program are stochastically chosen from the sample of eligibles using historically observed aggregate participation rates.

Potential benefits are then computed for each family in the national sample holding labor supply constant in a similar fashion by following the eligibility rules of the new program. The labor supply response of each member of the national sample is computed, based on SIME/DIME findings of differences between responses to current programs and the NIT. These

differences, plus the demographic and economic characteristics of each family and person in the sample, are components of the regression equations described in the previous paper. The regressions predict the expected labor supply response to the NIT for each person in the national sample. The NIT payments are then recomputed in the simulation model to reflect the labor supply response of each individual.

The impact of replacing AFDC and Food Stamps with a NIT on direct government expenditures can be estimated by summing the benefits over all the families in the national sample participating in AFDC and Food Stamps, inflating by the sample weights, and comparing the weighted sum to the totals obtained by the same summation of benefits for the new NIT program. Lost tax revenues resulting from reduced work effort can also be computed.

The distributional impacts of the welfare reform can be estimated by tabulating from the output file from the simulation the number of families gaining and losing program benefits together with the estimated changes in earnings resulting from the labor supply adjustment. This approach permits detailed tabulations showing the gainers and losers in any of the dozens of demographic or economic characteristics available in the data base.

LIMITATIONS OF GENERALIZATION

There are limits to generalizations based on this microsimulation model, mostly in the statistical analysis. The behavior of the experimental subjects must be decomposed into and statistically related to basic causes, such as

the NIT tax rate and payment level, the business cycle, and family characteristics. The observed labor supply behavior may be safely generalized only along those dimensions, which is to say only a sample of families which differ from the experimental families only in those basic causes of behavior that were included in the regression model. If labor supply behavior observed in the experiment is partially determined by something that is not included in the statistical analysis, like the NIT accounting period, then the statistical results cannot be used to predict responses to a plan that differs from the experiment in that dimension.

These models require validation, which is best done by applying them to data from other experimental sites, such as New Jersey and Gary, Indiana. The sample characteristics and point in the business cycle of these other experiments differ from those of the Seattle and Denver experiments in several measured, and probably unmeasured ways. Similar estimates of behavioral parameters produced by a variety of samples would be evidence that no relevant non-NIT characteristics are omitted from the model and would partially validate the generalizations.

A second aspect of generalization which requires validation is that the non-NIT causes of behavior are defined and measured in the experimental surveys in the same way as they are in the generalization survey. Income, for example, may include the subsidy value of low-rent housing in one survey and not in the other. One survey may include the income of previous family members who left the family prior to the survey and others may not. The benefit reduction rate of the experimental NIT program may not be per-

ceived by the program participants in the same way as they are for other transfer programs.

A final assumption needing validation is that both the regression model and the simulation model completely and accurately imitate both the new program and the behavior of those affected. For example, the assumption that labor market behavior responds immediately to the NIT tax rate and payment level requires further testing. Viewing people as moving from one static equilibrium before the experiment to another static equilibrium two years later may not be realistic. The theoretical models of labor supply need to be expanded to take into account the time over which the experimental stimuli and responses move. Questions regarding recipients' quickness to perceive the NIT tax rate, and to respond to a new tax rate, and the permanency of the response need to be addressed.

In sum, social experimentation and generalization are complementary tools for evaluating social programs and policies. Microsimulation is a versatile and, relative to alternative methods, accurate method of generalizing experimental results. For SIME/DIME purposes, additional work is needed to validate the non-NIT variables included in the models, the compatibility of variable measurement, and the theoretical models of behavioral responses.

NATIONAL LABOR SUPPLY EFFECTS AND
COSTS OF A NEGATIVE INCOME TAX

by

Myles Maxfield, Jr.
Senior Economist, MPR

and

Philip K. Robins
Senior Economist, SRI

In this paper we use the labor supply results from SIME/DIME to predict nationwide labor supply effects and costs of six alternative negative income tax (NIT) programs. These six programs are selected for study because they are within the range of alternatives that appear most likely to be considered for implementation. The methodology consists of estimating the parameters of a labor supply response function using SIME/DIME data and then simulating an NIT which replaced current transfer programs using a national data base (the March 1974 Current Population Survey), and applying the experimental response function to the national sample. The simulation is performed with the Micro-Analysis of Transfers to Households (MATH) model.

ESTIMATION OF THE WORK EFFORT RESPONSE TO SIME/DIME

Because of the sample selection procedure used in the experiments, the distribution of families by race, family structure and income differed

from the distribution of families in the U.S. population.¹ As shown in Table 1, greater proportions of SIME/DIME families were in lower income groups, compared with the U.S. population. Moreover, within the SIME/DIME population, experimental families have lower incomes than control families. When analyzing the impact of the experiment on labor supply, and when extrapolating the results to the U.S. population, these distributional differences must be taken into account.

In estimating the work effort response to SIME/DIME, we specify a model that identifies effects on work due to an increase in non-wage income and due to an increase in the cumulative tax rate, which reduced the net wage rate.² These are called the income and substitution effects, and are discussed in greater detail in the paper by Robins, "The Labor Supply Effects of SIME/DIME: An Overview." Identification of these effects enables us to compare the effects of different guarantee levels and tax rates--two parameters of an NIT program that can be set independently by the policy maker.

For the analysis of SIME/DIME data we calculated for each family head the change in disposable income that would occur from an NIT if there were no change in work effort.³ We also calculated the change in the net wage rate. The change in disposable income was calculated on the basis of earnings and nonwage income in the year prior to enrollment in the experiment, and the change in the net wage rate was

TABLE 1
 DISTRIBUTION OF INCOME IN THE SIME/DIME EXPERIMENTAL
 AND CONTROL SAMPLES AND IN THE U.S. POPULATION

Income Category ^a	Percent in Income Category					
	Husband-Wife Families			Female-Headed Families		
	Experimentals ^b	Controls ^b	U.S.	Experimentals ^b	Controls ^b	U.S.
\$1,000	9.8%	6.3%	1.1%	39.9%	36.0%	29.0%
\$1,000-3,000	8.2	6.3	1.8	16.1	11.8	17.5
\$3,000-5,000	9.7	6.9	3.9	13.5	13.1	8.5
\$5,000-7,000	12.8	10.3	5.3	18.8	19.6	12.2
\$7,000-9,000	18.7	19.5	6.1	8.9	12.2	9.4
\$9,000-11,000	17.3	18.5	10.3	1.9	3.8	7.2
\$11,000-13,000	10.3	16.0	11.3	.5	1.7	6.5
\$13,000-15,000	6.9	7.5	10.4	.3	.6	2.5
\$15,000-17,000	3.6	5.2	9.6	.1	.2	1.9
\$17,000-20,000	2.2	2.5	13.4	0	0	1.9
\$20,000	.5	1.1	26.1	0	.2	3.3
Total number of Families	1.023	1.158	39.7	9.68	6.54	5.0
			million			million

^a Income is defined as earnings of all family members plus family nonwage income, excluding taxes and public transfer payments. The incomes in the SIME/DIME sample are for the year before enrollment and are inflated to 1974 dollars. The U.S. incomes are from the March 1975 Current Population Survey and cover the year 1974.

^b Black and white families only. There are approximately 800 Mexican-American families enrolled in the Denver Experiment, but these families are excluded from the tabulations.

calculated as the product of a predicted pre-experimental wage rate and the difference between the experimental and pre-experimental tax rates. The pre-experimental tax rates are derived in accordance with the laws governing the programs in which a family was enrolled prior to the experiments.⁴

Many families enrolled in the experiments are initially above the breakeven level. For these families, the calculated values of the change in disposable income and the change in the net wage rate are zero. Even though the calculated values are zero, some of the families will respond to the experiment. To measure the response of families above the breakeven level, we define three variables that capture the location of the family relative to the breakeven level. These three variables combine the probability of going below the breakeven level with the income and wage changes that would result if such a movement were to occur.

Separate estimates of the above and below breakeven parameters are made for husbands, wives, and female heads of families. The sample consists of a subset of originally enrolled black and white family heads who remained in SIME/DIME for at least two years and for whom accurate data are available. The subgroups for analysis are defined as of the date of enrollment, regardless of changes in marital status that may have occurred subsequent to enrollment. The measure of work effort response is equal to the change in hours of work between the second year of the experiment and the year prior to the experiment.

Results

The estimated labor supply effects are presented in Table 2. For families below the breakeven level the income effects are negative and statistically significant for wives and single female heads of families and the substitution effects are positive and statistically significant for all three groups. Because the NIT experiment raises income and lowers the net wage rate, the income and substitution effects imply an unambiguous negative impact on annual hours of work. For the mean working individual in the SIME/DIME sample below the breakeven level the labor supply effects are -103 hours per year for husbands, -263 hours per year for wives, and -176 hours per year for female heads. The percentage effects for the SIME/DIME sample are -5%, -22%, and -11%, respectively. For families above the breakeven level, only wives appear to be responding to the experiment. All three types of heads exhibit a response above the breakeven level that declines in absolute value with distance from the breakeven level.

THE MICROSIMULATION MODEL OF NIT COST, CASELOAD, AND INDUCED LABOR SUPPLY RESPONSE

The MATH Model and Simulation Data Base

The microsimulation model used in this paper has been introduced above in the paper by Beebout and Maxfield. Generally, individual population data are collected, treated, and used for calculations of individual responses to various new programs, which are then tabulated to produce estimates of a national response.

TABLE 2
ESTIMATED EXPERIMENTAL EFFECTS
ON ANNUAL HOURS OF WORK

	<u>Husbands</u>	<u>Wives</u>	<u>Female Heads</u>
<u>Below Breakeven Level</u>			
Income Effect (per thousand dollars)	-34	-143***	-101**
Substitution Effect (per dollar per hour)	83**	168*	126*
<u>Above Breakeven Level</u>			
Constant Effect	-13	-431*	-345
Effect of breakeven level (per thousand dollars)	- 6	8	73
Effect of Earnings Above Breakeven Level (per thousand dollars)	12	48	35
<u>Estimated Experimental Effects for the Average Working Individual in SIME/DIME below the Breakeven Level</u>			
Hours per year	-103	-263	-176
Percentage effect	- 5%	- 22%	- 11%

- * Indicates significance at 10% level
- ** Indicates significance at 5% level
- *** Indicates significance at 1% level

In this study the Micro Analysis of Transfers to Households (MATH) model was applied to data contained on the March 1975 Current Population Survey (CPS) to estimate the impact of replacing current welfare programs with an NIT. As noted in the preceding paper, MATH applies eligibility requirements and benefit determination schedules. It mimics decision-making behavior of low income families regarding welfare participation and work effort (as measured in SIME/DIME) under an NIT. Then the eligibility rules and benefit determinations for each family are summed, assuming 100% participation of eligible families. The simulation of labor supply response of the head and the spouse involves multiplying the NIT-induced change of filing unit income and the effective wage rate of each by the two response parameters. The two products are summed to produce a predicted change of annual hours of employment under the NIT. (Other wage earners are assumed to maintain a constant level of appointment.) Then, the postresponse earnings, family income, and tax liability are calculated. These variables are used to produce the figures presented in the next section. The tabulated variables include the sum of NIT payments, or NIT program costs, both gross and net of the payments of the abolished AFDC and Food Stamps programs, the number of families participating in the NIT is accumulated, and the numbers of hours of employment before and after the labor supply response are tabulated. The changes of earned income and tax liability resulting from the labor supply response are computed. Finally, the portion of the NIT payments that is composed of tax reimbursement is accumulated.

The data base used in this study is the March 1975 Current Population Survey and Income Supplement. The CPS consists of a stratified random sample of approximately 50,000 households throughout the 50 states and the District of Columbia. The survey contains a series of questions about family demographic characteristics and structure. Respondents are also asked to list the income of each person in the family for the previous calendar year, in this case January through December 1974. The survey asks about each person's labor market experiences, for the previous year and for the survey month.

The nature of the questionnaire imposes some limitations on the simulation model. The incomes recorded are annual, so simulating welfare programs with an accounting period less than a year is difficult. The simulation of many tax and welfare programs requires income, demographic, and family structure information simultaneously; however, the recorded income amounts were accrued in the calendar year prior to the month to which the demographic and family structures pertain. The discrepancy could bias results. Under-reporting of income is a common problem, for welfare payments and other income sources as well. Inadvertent under-reporting occurs both when the respondent inaccurately recalls the receipt of income during the prior year and when a person who contributed to the family income in the year prior to the survey left the family before the interview was administered.

Assumptions

Many assumptions are made in the simulation model, some of which are apparent from the model and data description. The most important assumption is that the labor market behavior of the experimental samples is similar to that of the national population. The experimental tax and income effects which were measured for residents of a few downtown neighborhoods of Denver and Seattle are extended in this study to low-income families throughout the nation. This assumed similarity of behavior may be tested in the future by careful comparison of the results of the several different income maintenance experiments.

A second set of assumptions is made about the participation of eligible families in the NIT. In the regression model, families which were income-eligible to receive an NIT payment during the enrollment survey are assumed to receive a payment throughout the first two years of the experiment. This was not always the case; for example, a family head who suffered a temporary spell of unemployment at the time of the enrollment survey but found a well-paying job before the second year of the experiment would not fit this pattern. The simulation model makes analogous assumptions. The year of the simulation data base is analogous to the beginning of the experiment. Every income-eligible family in the data base is assumed to participate in the NIT throughout the two years represented by the simulated response.

Additional assumptions are made about the annual accounting period of the NIT and the level of income under-reporting. The NIT

payments of most families in the experiments were computed every month based on monthly income. Because of data limitations, the simulated NIT payments were computed only once for the year from annual income. This assumption probably biases the simulation estimates of NIT costs and caseloads downward slightly. The simulated NIT payments are computed from income reported to the CPS interviewer, thus assuming that the amount of misreporting to the NIT agency is the same as that to the CPS.

The estimates are made under 1974 labor market conditions. Some preliminary testing of the model indicates that the response to an NIT may be quite sensitive to the unemployment rate. The interaction of labor supply response with other behavioral responses to the NIT, such as family composition response, is not incorporated in these estimates. The simulation model assumes that the response to changes in effective wages and income caused by other tax and transfer programs is the same as the response to an NIT.

Finally, the estimates are based on average responses. The simulation model does not reproduce exactly the distribution of responses in dimensions other than family-type and income, primarily because of the limited comparability of the experimental with the simulation data base. A more accurate representation of the distribution of responses would be important to any study of the number of people dropping out of the labor force or made better or worse-off by the NIT.

SIMULATION RESULTS

Labor Supply Responses

The predicted labor supply responses to six different NIT plans are presented in Table 3 and are reported in two ways: first, the average predicted response for all participating families; and second, the average predicted response for the U.S. population, including that of certain non-participants, as well as participants. (The non-participants who respond are families that previously received welfare, but are above the breakeven level of the NIT program, and therefore increase their labor supply when the NIT replaces welfare.)

In interpreting the results, it is important to keep in mind that the responses vary not only because of changing guarantee levels and tax rates, but also because of a changing pool of participants. For example, at any given guarantee, as the tax rate increases the pool of participants decreases. The total effect depends on the distribution of income within the relevant population subgroup. For the programs simulated, the number of families receiving benefits would range from 3.3 million to 19.3 million.

For participating husband-wife families, the magnitudes of the average responses are positively associated with both the guarantee and the tax rate. For participating families headed by women, the responses are positively associated with the guarantee, but do not vary with the tax rate. For both groups, the results indicate fairly sizable reductions in labor supply, ranging from between 10 and 21%

TABLE 3

**AVERAGE LABOR-SUPPLY RESPONSES FOR ALL PARTICIPATING FAMILIES
AND FOR ALL FAMILIES IN THE UNITED STATES**

NIT Support Level	NIT Tax Rate 50%					NIT Tax Rate 70%				
	Participating Families		All U.S. Families			Participating Families		All U.S. Families		
	Change in Annual Hours of Work	% Change	No. of Participating Families (millions)	Change in Annual Hours of Work	% Change	Change in Annual Hours of Work	% Change	No. of Participating Families (millions)	Change in Annual Hours of Work	% Change
<i>50% of Poverty Level^a</i>										
Husbands	-104	-7.0%		-4	-0.2%	-136	-10.8%		-2	-0.1%
Wives	-92	-23.3		-2	-0.3	-111	-29.9		0	0.0
Total (H+W)	-196	-10.3	2.4	-6	-0.2	-247	-15.1	1.3	-2	-0.1
Female heads	0	0.0	2.3	+16	+1.6	-10	-2.7	2.0	+20	+2.0
<i>75% of Poverty Level^a</i>										
Husbands	-106	-5.9		-19	-1.0	-157	-11.2		-9	-0.5
Wives	-110	-22.8		-19	-2.4	-126	-32.5		-5	-0.6
Total (H+W)	-216	-9.5	7.6	-38	-1.4	-283	-15.8	2.8	-14	-0.5
Female heads	-47	-6.7	3.0	-23	-2.4	-47	-9.3	2.5	-12	-1.2
<i>100% of Poverty Level^a</i>										
Husbands	-119	-6.2		-47	-2.4	-164	-10.1		-23	-1.2
Wives	-130	-22.7		-50	-6.3	-144	-32.0		-18	-2.3
Total (H+W)	-249	-10.0	15.7	-97	-3.5	-308	-20.6	5.8	-41	-1.5
Female heads	-99	-12.0	3.6	-69	-7.1	-95	-14.9	3.0	-52	-5.3

Note: Average hours of work per year before response, all husbands in the U.S. = 1,999. Average hours of work per year before response, all wives in the U.S. = 793. Total number of husband-wife families in the U.S. = 39.8 million. Average hours of work per year before response, female heads in the U.S. = 974. Total number of female-headed families in the U.S. = 4.9 million.

^a Poverty level is \$5,000 per year for a family of four in 1974.

for husband-wife families and between 0 and 15% for families headed by women.

The responses averaged over the entire U.S. population are quite small relative to the average responses of participating families because most U.S. families do not participate in the program. While the magnitudes of the average responses again increase with the guarantee (as they do for participants), they decrease with the tax rate for both groups. This inverse relationship between the average U.S. response and the tax rate is an interesting and perhaps unexpected result that is a consequence of the fact that the number of participants decreases by an amount large enough to offset the effect of a large response among participants.

Effects on Welfare Participation

Since the simulations assume that several welfare programs (Aid to Families With Dependent Children-AFDC, AFDC-UP and Food Stamps) are replaced by the NIT, some families are made worse off by the NIT (i.e., their disposable income is reduced).

Table 4 presents a tabulation of the number and percentage of welfare families that are made worse off by the NIT. As this table indicates, the percentages are unexpectedly large, particularly for the generous NIT programs. For example, under an NIT program with a support level equal to the poverty line and a tax rate of 50%, one quarter of the

TABLE 4

NUMBER AND PERCENTAGE OF WELFARE*
FAMILIES MADE WORSE OFF BY THE NIT

NIT Support Level	NIT Tax Rate 50%		NIT Tax Rate 70%	
	Number Made Worse Off (Millions)	Percent Made Worse Off	Number Made Worse Off (Millions)	Percent Made Worse Off
50% of poverty level^a				
Husband-Wife Families	1.2	79%	1.4	89%
Female-Headed Families	1.8	93	1.9	95
Total	3.0	87	3.3	92
75% of poverty level^a				
Husband-Wife Families	.7	43	1.2	71
Female Headed Families	1.4	67	1.6	75
Total	2.1	59	2.8	73
100% of poverty level^a				
Husband-Wife Families	.4	23	.7	41
Female Headed Families	.5	25	.7	33
Total	.9	24	1.4	37

* AFDC, AFDC-UP, Food Stamps

^aPoverty level is \$5,000 per year for a family of four in 1974.

families, participating in AFDC or Food Stamps are made worse off.

The reason that so many families are made worse off may be due to the fact that there are provisions in the existing welfare system (especially in allowable work-related expenses as deductions) that enable families to face very low benefit reduction rates. These low benefit reduction rates imply that welfare grants remain high even when family members earn a substantial income. Thus, although the support level of the NIT may be higher than the support level of welfare, the higher NIT tax rate makes many working welfare families worse off.

NIT Program Costs

The total budgetary cost of each NIT plan is the sum of the payments received by all NIT participants.⁵ Two kinds of budgetary costs are presented: the sum of the NIT payments (gross costs) and this gross cost less the current cost of AFDC and Food Stamps for the NIT participants (net costs). Net costs represent the net changes in the federal welfare budget caused by replacing AFDC and Food Stamps with the NIT. Net costs also represent the net change in aggregate transfer payments received by each income category of families. Table 5 portrays the total budgetary cost of the six NIT plans before any labor supply adjustments take place. Increasing the support level increases the program costs more than proportionately. The least expensive plan has a gross cost of about \$7 billion. At a high-benefit reduction rate, increasing the support level to 75% of the poverty line adds \$8 billion to the cost, and increasing it to 100% adds \$17 billion. At the lower

TABLE 5
NIT PROGRAM COSTS⁸

NIT Plan	Number of Participants (in millions)	Gross Costs		Percent Change	Net Costs		Percent Change
		Preresponse (in billions)	Postresponse (in billions)		Preresponse (in billions)	Postresponse (in billions)	
70/50	4.25	6.77	7.44	10	-3.14	-2.48	21
50/50	7.17	8.09	8.83	9	-1.83	-1.09	40
70/75	8.36	15.17	15.17	14	3.44	5.25	52
50/75	15.15	18.89	21.09	12	8.97	11.17	24
70/100	13.17	23.82	27.76	17	13.91	17.85	28
50/100	26.58	40.30	45.39	13	30.38	35.48	17
Current Services	--	9.92	--	--	--	--	--

tax rate, increasing the support level from 50% to 75% increases costs by \$11 billion, and increasing it from 50% to 100% adds \$32 billion. These costs can be compared to the cost of current welfare programs--approximately \$10 billion for the AFDC and Food Stamps programs in 1974.

The fourth and seventh columns of Table 5 indicate the percentage change in the program costs due to the labor supply response to the program. The pre-response program cost is the sum of the payments to participating families computed from the earnings observed in the simulation data base. After the labor supply response to the welfare reform has been simulated, each family has a new level of earned income, resulting in eligibility for a different NIT payment amount. Since most of the labor supply responses of participating families are to work fewer hours, the post-response cost is greater than the pre-response cost. Thus, the third column shows gross costs increasing by roughly 9% in the least generous plan to 17% in the most generous plan, as a result of the later supply response.

The budgetary cost of an NIT is determined both by the size of payments to a given family and by the number of families that participate in the program. The latter determinant is portrayed in the first column of Table 5 which shows the number of families simulated to receive an NIT payment. Such families are called participants for convenience, although the use of the term assumes that every eligible family participates. Approximately 4.2 million families participate in the least generous plan. At the 70% tax rate, increasing the support

level to 75% of the poverty line adds 4.1 million families, and increasing the support to 100% adds another 4.8 million families. The 50% support/50% tax includes 7.2 million participating families. Increasing the support level to 75% at the lower tax rate adds 7.9 million families, and increasing the support level to 100% includes an additional 11 million families.

Changes in the benefit reduction rate have a greater impact on program costs than do tested support level changes, especially among the more generous plans. All but two NIT plans--those with a support at 50% of the poverty line--are more costly than current welfare programs. These two plans reduce the total transfer payment amounts going almost to all participants.

Disaggregation of Results
by Family Type, Family Income and State

An important quality of the microsimulation technique is the ability to disaggregate results by a number of economic and demographic dimensions. This section displays the effects of the NIT on recipient families by family income, type of family head, and by state. Family income disaggregation is used because a primary objective of a transfer program is the modification of the distribution of income. Husband-wife families are isolated because they form the largest group of low-income families currently ineligible to receive a payment from AFDC in many states; and female-headed families are isolated because they form the group currently covered by AFDC. The results by state offer examples of the deviation of individual states from the average effects in the nation. Unlike the results reported previously,

these disaggregated results do not include families that were initially above the NIT breakeven point who then reduced their work effort in order to become eligible for a payment.

The differential aspects of replacing current welfare programs with an NIT among families with different types of heads is shown in Table 6 which displays the distribution of budgetary costs by family type broken into two income classes. The figures are presented for a 75% support, 50% tax plan, chosen for its similarity to reform proposals currently being considered, and for existing welfare programs.

TABLE 6

NIT BUDGETARY COST BY FAMILY TYPE
WITH NIT AT 75% SUPPORT, 50% TAX
(Billions of Dollars in 1974)⁹

	Gross Cost	Net Cost	Current Services Cost
<u>Husband/wife families</u>			
Income < \$3000	2.98	2.19	.79
Income > \$3000	5.12	3.98	1.14
<u>Female headed families</u>			
Income < \$3000	5.73	.40	5.33
Income > \$3000	1.14	.08	1.06
<u>Other families</u>			
Income < \$3000	3.33	3.04	.29
Income > \$3000	.60	.49	.11

In the first column of Table 6, the largest proportion of gross payments at low income levels goes to families with children headed by a woman. This reflects the composition of the low-income population. As income rises, the total payment declines smoothly but the composition shifts from female-headed families with children to families headed by a couple.

Net costs present a different picture. The proportion of net costs allocated to female-headed families with children is small throughout the entire income range. This is due to the fact that the largest portion of current welfare payments is received by female-headed families. Under the current transfer programs, husband-wife families receive only limited assistance, and families headed by a single person without children receive virtually nothing.

Table 7 disaggregates the labor supply response to a 50% tax with a 75% support by family type and non-welfare annual income. Generally, the percentage reduction of labor supply falls as income rises. This pattern is consistent with the assumption that the NIT has a larger impact on the economic circumstances of lower-income families. The average annual hours of employment per person rises with income, showing the association of non-employment with low-income families. The income distribution of total reduction of hours of employment is influenced both by this reduction of the average labor supply response as income increases and by the increased population density as income increases.

TABLE 7

LABOR SUPPLY RESPONSE BY INCOME AND TYPE OF RESPONDENT, 50/75 NIT PLAN ¹⁰

Income	Husbands			Wives			Female Heads		
	Average Preresponse Hours	Average Postresponse Hours	Percent Reduction	Average Preresponse Hours	Average Postresponse Hours	Percent Reduction	Average Preresponse Hours	Average Postresponse Hours	Percent Reduction
less than 0	2362	2048	-13.29	829	575	-37.88	937	937	0.
0-999	676	594	-12.11	215	139	-35.35	106	87	-17.92.
1000-2999	989	874	-11.63	317	217	-31.55	708	654	- 7.63
3000-4999	1443	1317	- 8.77	398	290	-27.14	1258	1194	- 5.09
5000-6999	1833	1715	- 6.45	473	386	-18.39	1596	1499	- 6.08
7000-8999	2041	1934	- 5.23	464	376	-18.97	1613	1524	- 5.52
9000-10999	2180	2077	- 4.72	432	355	-17.82	1616	1566	- 3.09
11000-12999	2196	2103	- 4.21	473	405	-14.38	1549	1549	0.
13000-14999	2169	2077	- 4.25	521	445	-14.59	2143	2134	0.
15000-16999	2177	2063	5.26	329	257	-21.88	1471	1477	0.
17000-19999	2138	2065	- 3.39	1195	1087	- 9.04	0.	0.	0.
20000 and up	2807	2807	0.	702	702	0.	0.	0.	0.
Total	1794	1681	-- 6.30	438	346	-21.00	699	652	- 6.72

TABLE 8

THE EFFECTS OF A NEGATIVE INCOME TAX
ON WASHINGTON AND COLORADO¹¹

	Washington (millions)	Colorado (millions)	United States (billions)
Gross costs (annual 1975 dollars)			
50/75 Plan	\$ 255	\$ 200	\$ 19.1
70/50 Plan	90.4	79.8	----
50/100 Plan	570	502	----
Costs net of current AFDC and Food Stamp expenditures, 50/75 Plan (annual 1975 dollars)	(millions) \$ 111	(millions) \$ 112	(billions) \$ 8.25
Eligible families	(thousands)	(thousands)	(millions)
50/75 Plan	\$ 213	\$ 195	\$ 15.4
70/50 Plan	79.8	54.8	----
50/100 Plan	396	380	----
Labor supply reduction	(percent)	(percent)	(percent)
50/75 Plan			
Husbands	3 %	1%	3%
Wives	20	16	19
Female heads	1	2	2

Table 8 summarizes several simulation results for the states of Washington and Colorado. These results are not comparable to other figures presented in this section because they are derived from a different simulation data base, the 1976 Survey of Income and Education (SIE), rather than the 1975 CPS. The SIE is similar to the CPS except that it contains roughly three times the number of households. This larger size is necessary for making statistical estimates for subsamples as small as a state. Several national estimates are derived from the SIE in order to provide a basis for comparison of the state figures.

An NIT with a tax rate of 50% and a support level of 75% is used as a base program with which to compare the states of Washington and Colorado to each other and to the nation as a whole. This program is estimated to cost, prior to any labor supply response and in 1975 dollars, \$19.1 billion for the total United States, \$225 million for Washington, and \$200 million for Colorado. There are estimated to be 15.4 million families (including families of one) receiving aid in the United States, 213 thousand in Washington, and 195 thousand in Colorado. The net cost (total cost less AFDC and Food Stamps) would be \$8.25 billion for the United States, \$111 million for Washington, and \$112 million for Colorado. The larger caseload of Washington, relative to Colorado, appears to consist mostly of female-headed families with children currently receiving welfare payments. Thus, the gross NIT cost of Washington is larger but the net cost is similar in the two states.

The cost and caseload of a less generous NIT--a tax rate of 70% with a support of 50%--are lower in both states. The gross program cost is \$90.4 million in Washington and \$79.8 million in Colorado. There are 79.8 thousand families receiving payments in Washington and 54.8 thousand in Colorado. The discrepancy between Washington and Colorado costs and caseloads is larger for this less generous plan, due to the large group of female-headed welfare families in Washington who are at very low levels of income and still eligible under the less generous plan.

An NIT with a benefit reduction rate of 50% and a support level of 100% is an example of a relatively generous program. Such an NIT plan would cost \$750 million in Washington and \$502 million in Colorado. There would be 396,000 families eligible for a payment in Washington and 380,000 in Colorado. The costs and caseloads are disproportionately larger than the base plan (which has the same tax, 50%, and a 75% support) because the breakeven level of income is higher, falling at a point in the national income distribution curve which contains a great number of families.

The reduction of the number of annual hours of employment induced by the base plan is 3% for husbands in the United States, 3% for husbands in Washington, and 1% in Colorado. The response in the total United States is smaller in 1975 than in 1974 primarily because the unemployment rate rose from 6.5% to 8% over the year. The increased incidence of unemployment fell disproportionately upon

low-income families who would be eligible for an NIT payment. Little labor supply reduction is possible for a family head who is out of work. The reduction for wives is 19% in the United States, 20% in Washington, and 16% in Colorado. Females heading families with children reduce their labor supply by 2% in the United States, 1% in Washington and 2% in Colorado.

CONCLUSIONS

The regression analysis of the experimental samples and the simulation generalization to the nation have estimated that an NIT brings about a moderate reduction of the average work effort of participants. Furthermore, the reduction is larger for more generous NIT plans. One implication of this result is that accurately measuring the degree to which a program achieves its intended objective requires taking into account the changed work effort induced by the program. The budgetary cost of the program, the impact on the distribution of income, and the amount of additional income accruing to participating families are significantly affected by the induced labor supply response. The tradeoff between budgetary cost and adequacy of support for a proposed transfer program, for example, can only be determined after taking account of the labor supply response.

A second implication is that the behavior of participants, in this case participants' work effort, must be directly weighed by the policy maker against the other objectives of the program. The labor supply response of an NIT must be considered not only in order to measure the program cost or adequacy of support accurately, but also because

the employment, or economic self-sufficiency, of low-income families is itself an important goal of the welfare program. Employment and work experience is a program goal in that it potentially provides a mechanism for escaping from poverty. A welfare program that induces increased employment is likely to result in more families exiting from poverty than would a program that does not.

Viewing participant employment as a program goal is too simple, however, because the policy-maker must also place a value on the uses of participant nonmarket time. The nonmarket activity of a single parent may have a different value than that of a parent who has a working partner. The non-employment of a young person who uses the time for training or education may have a value different from that of a young person who uses the time in some other way. The different uses of time outside the labor force may lead to consideration of categorical program regulations such as the work requirement under discussion in the proposed Better Jobs and Income Program.

Future Directions

The results presented here may be extended in several ways. The NIT examined in this study was not financed with taxes. One extension of this work is to simulate an income tax surcharge to finance the NIT in order to estimate the effect of the NIT on the overall income distribution. Policymakers also need to know what effect on fiscal flow systems from partially financing the NIT with state supplementation or with federal-state matching grants, taking account of labor supply response and state and local tax structures. The model could

be improved to give a clearer picture of labor market behavior of initially ineligible families who reduce employment in order to qualify and families who become ineligible by securing private sector jobs. The potential response of employers to the reduced work effort of the low-income work force should be examined and possibly incorporated into the simulation model. The modeled transfer program could be more completely integrated with other transfer programs such as Medicaid and Unemployment Compensation. Finally, measures of the program's effect on recipients' well-being could be refined to take account of the increased leisure, training, or home production that is made possible by lowered labor supply.

NOTES

1. The SIME/DIME sample is limited to husband-wife families and female-headed families where the head is between the ages of 18 and 58 and there is at least one dependent child.
2. The net wage rate is a measure of the net economic returns from working. If w is the gross (before tax) wage rate, t_e is the NIT tax rate, and t_p is the pre-NIT tax rate, the change in the net wage rate caused by the NIT is given by $-w(t_e - t_p)$.
3. The change in disposable income calculated on the basis of pre-program work effort is equal to the NIT payment received at enrollment.
4. For details of the calculations as well as a more complete description of the model, see Michael C. Keeley, Philip K. Robins, Robert G. Spiegelman, and Richard W. West, "The Estimation of Labor Supply Models Using Experimental Data", American Economic Review (forthcoming, December 1978).
5. From Michael C. Keeley, Philip K. Robins, and Richard W. West, "The Labor Supply Effects and Costs of Alternative Negative Income Tax Programs: Evidence from the Seattle and Denver Income Maintenance Experiments: Part II, National Predictions Using the Labor Supply Response Function", Research Memorandum 39, Center for the Study of Welfare Policy, Stanford Research Institute, Menlo Park, California, May 1977.
6. Source: Simulation runs prepared for SRI International by Mathematica Policy Research and The Hendrickson Corporation.
7. These budgetary costs do not take account of the reduction of tax revenues caused by the labor supply response.
8. Source: Calculated by Mathematica Policy Research by applying the Micro-Analysis of Transfers to Households model to the March 1975 Current Population Survey and Income Supplement.
9. Id.
10. Id.
11. Source: Figures calculated by Mathematica Policy Research by applying the Micro-Analysis of Transfers to Households model to data contained in the 1976 Survey of Income and Education.

DISCUSSION

TED LANE: How did you treat the change in income or the change in their actual earned wage that resulted from their training?

MAXFIELD: In this model we didn't handle the training trainee sample very well. We entered them in a very simply way of having an intercept adjustment for, "Are you on training, or are you not?"

LANE: What about their post-training changes in wages, if there were any?

MAXFIELD: That was also not entered in the model, as we were not measuring the effect of the manpower treatments. This analysis is only of cash payment, and that's the only thing that we purport to be generalizing as well.

LANE: Did your sample exclude the people who were given a manpower treatment?

MAXFIELD: No, we did have an intercept adjustment for them, and it might have been reasonable to leave them out of the sample. We chose not to in order to keep the sample large, and obtain efficient statistics.

MICHAEL STERN: It makes some difference (on the labor supply response) doesn't it, whether you're adding to the first \$1,000 or \$10,000?

MAXFIELD: Yes. The results that we'll talk about today assume that it did not make a difference; but you are correct, and we did estimate a model which allowed these numbers to vary by the level of income, by where you're starting.

LANE: Even this one. If you have a non-zero slope for your regression, then your average and marginal rates should be different. So the effect for a \$1,000 change somewhere further up the income curve is going to be different than for somebody further down.

MAXFIELD: Yes, I think Mr. Stern's point is still there. We did try exactly what he is suggesting, and the paper by David Greenberg will, I think, incorporate that specification. We chose not to in this paper, because the interaction of the wage effect with initial income was not very significant statistically, which is to say that the effects didn't change very much by the initial level of income for our sample. We were somewhat surprised to find that the case.

LANE: Is this right? Husbands work 1700 hours out of about 2000 hours a year. Is this reflecting unemployment?

MAXFIELD: Yes. It's not as though everyone is working part-time, but some of the people are working full-time.

LANE: Well then, when you have "percent unemployed", is that for a point in time, or is it averaged over the year?

MAXFIELD: I'm not sure about that. I think that's at a point in time. I think that's just the answer to the question, "Are you unemployed currently?"

LANE: If the average put in over 1700 in 2,000 hours possible a year, you have about 83% average employment.

MAXFIELD: Yes.

LANE: Some of your sample must have been employed almost for the full year.

MAXFIELD: Oh, definitely. Because the participation rate for men is so very high--around 90%.

LANE: Did you make any adjustment for persons employed full-time but at low-income jobs?

MAXFIELD: Yes.

LANE: For these, did you add income? The only way that they're going to increase their number of hours worked is by moonlighting.

MAXFIELD: That's right.

LANE: That's different than the person who's unemployed for part of the year and decides to work more often. It's a different pattern of behavior. Did you deal with that differently?

MAXFIELD: Not explicitly. First of all, other papers deal with those issues. Second, for almost everybody, the effect of an NIT is to make you work less, not more. So, the more important issue concerns people who don't work, not the people who work full-time. Lastly, the responses that were in the regression model and the simulation are again, averages

LANE: I'm just wondering what went into the regression; I'm wondering about the ability to generalize it. If an employed person came in and said, "I'd like to go back to college"; is that considered a withdrawal from the labor force. Are you attributing that behavior to the training treatment, or are you attributing it to the change in income effect or the substitution effect?

MAXFIELD: That's a difficult question. In general, we are attributing that to the effect of the NIT, not to the manpower program.

LANE: So then, you would have a tendency to overstate the withdrawals from the labor force on account of the training component?

MAXFIELD: There probably is a bias due to not taking account of the manpower treatments in a better way. Whether or not you're in a manpower treatment is in the intercept; so it's not as though they're not there at all. They're partially controlled for, but not completely controlled for.

STERN: (When computing net costs) what did you say your assumption was for AFDC costs?

MAXFIELD: That's a big subject. We don't use the statutory 67% tax. We use a computation where the AFDC payment was regressed against earned income, to account for all of the work-related expenses--such as child care expenses. Doing this, we get a tax of around 33%.

STERN: But is it true that a person working full-time would have twice the expenses of a person working half-time---that the percentage would be about the same?

MAXFIELD: I don't think there's any evidence that work-related deductions taper off as income goes up. I think it's pretty steady. However, many states impose a maximum payment--an arbitrary cut-off. This will make the relationship nonlinear.

STERN: I can see that would be true if you work more days per week, but suppose you just get a raise or a job that pays more.

MAXFIELD: I can only speculate. If you get a raise, you hire a better day-care center; you improve the quality of the work-related services that you're purchasing. This is just an empirical relationship, and empirically I have not observed it to fall off.

LANE: Have you tried looking at the difference between men and women on that? It strikes me a lot of women have half-time jobs, but looking at the other side, you still have to take a bus downtown and back and be there for four hours.

MAXFIELD: I'm sure that's true. Most of the work has been done on women because the number of men in AFDC is so small that it's hard to do anything statistical. I'm sure there are nonlinearities in this work-related expenses relationship for half-time workers, but I haven't looked at that specifically.

STERN: I realize you don't want to draw major conclusions, but this suggests that if you want to preserve the best work incentive, maybe what you ought to do is have a low-guaranteed level with a high tax rate, because in effect, you'd be writing off the people who get benefits. You just say, "Well, they're never going to make it", they're going to be in this trap that you've got for them, but you would want to expose the fewest possible people to that trap.

MAXFIELD: Yes, that's true. If your objective is to minimize the disincentive costs to society of an NIT you impose the low support/high tax rate--the least generous plan. It imposes the least disincentive costs because there are the fewest number of people on the program, right? Am I understanding you?

STERN: Yes, but that's not a prima facie assumption, that's based on some experimental results.

MAXFIELD: Yes.

STERN: Because you might just as well assume, before you put any program into effect, that everybody will just say, "Well, if I can get something for nothing, I'd rather do nothing." Perhaps my point is, that if what you really would like to do is avoid putting people in a position where they have to be weighing work disincentive effect, then it's, of course, better not to affect them. The more conventional wisdom that's been used in designing legislation over the last decade has been different. It's that, "Well, let's have a lower effective tax rate. That's half the greater incentive." This may suggest that once you get above some level, maybe after you get above a 50 or 60% tax, people just react by saying, "That's confiscatory", and after that the difference between 60 and 80% doesn't affect them. I don't know. Their reaction just may be, "Well, I'd rather cheat", "I'd rather just not work", or something. They may not be that sensitive to that kind of difference, but the difference between 40 and 60% might make a big difference. If that is true then you'd be better off forgetting a program with the rate at 50% or so, which wouldn't have the effect you like, and just saying, "Well, let's have a higher rate and just try to affect less people". That might be your program conclusion from this kind of table here.

MAXFIELD: I think that's true, there may be a couple of things to say. We did test wide variations in the tax rate--both in the experiment and in the simulation. Now, the difference between 30% tax for AFDC and 50% in an NIT may not seem large, but male-headed families who had incomes low enough to be free of federal taxes, the tax went from zero to 50%, and those jumps are incorporated into these behaviors.

MAXFIELD: My first reaction to your suggested plan to minimize the disincentive cost is that it happens to be the least generous plan. Presumably, one of the goals of a transfer program is to transfer income, to provide income support or income maintenance. I think the conclusion that I would draw is different; I would conclude that nothing's free, that there's a tradeoff between the disincentive costs and the adequacy of income support. Another response to that tradeoff might be to consider as an option the declining tax rate plan tried in SIME/DIME. It had a very high tax rate on the lowest income people (70%) but it smoothly declined so that people who were working were only faced with a 20% rate.

STERN: I think unfortunately that that can only be done in an experiment

MAXFIELD: Yes, I think that's right.

LANE: If you look at the occupational structure of the employment, and if most of the wives work as domestics, and if in fact then, their income maintenance support goes up, they still have the same job opportunities available to them. The probability is that they'll not take advantage of extra job opportunities if it's at a menial wage or in a boring job. If the husbands are at a higher occupational category of higher average earnings, then they're going to withdraw less.

MAXFIELD: Yes.

LANE: If you want to minimize labor supply withdrawals, under an income maintenance system, you also have to couple it with better jobs. That the biggest impact is going to be on the low-wage menial side of the labor market.

MAXFIELD: I think it's a reasonable thing to look at. This analysis doesn't do it. You're talking about looking at the labor supply of one person in relationship to the labor supply of a spouse. There is other regression work which does deal with that. That was just one of the many complications that we didn't incorporate here.

SUPPORT LEVELS AND TAX RATES:
IMPLICATIONS FOR A NEGATIVE INCOME
TAX PROGRAM DESIGN

by

David Edson
Research Associate
Policy Studies Division, MPR

and

Myles Maxfield, Jr.
Economist, MPR

with the assistance of

Katherine J. Siena

In this paper, we present estimates of the tradeoffs between some important goals of an income maintenance program. The goals we examine include (1) raising the incomes of the poorest families, (2) encouraging private employment of the poor, and (3) reducing the budgetary cost of a nation-wide negative income tax (NIT), relative to Aid to Families with Dependent Children (AFDC) and Food Stamps. As in the preceding paper by Maxfield and Robins, we used a microsimulation model of existing tax and transfer programs, of an NIT, and of the labor supply response of transfer recipients to the new program to produce our estimates. The data base is the March 1975 Current Population Survey. To estimate tradeoffs we simulated the transfer payments, the employment impacts, and budgetary cost of each of six different NIT programs. In this paper, unlike the preceding

paper, we attempted to distinguish between the effects of the support level and the benefit reduction rate (the tax rate).

Whereas a tradeoff between adequacy of support and program cost is to be expected, the tradeoff between support and encouragement of employment is not clear. An NIT may be made more generous by increasing the support level, by decreasing the benefit reduction rate, or by a combination of both. Generally, a more generous support level has been found to induce a larger reduction of work effort, and a lower benefit reduction rate a smaller reduction of work effort.

We found that large increases of the income of poor families, small reductions of the work effort of recipients, and low budgetary cost cannot be achieved simultaneously. Large income increases and low budgetary costs may be achieved at the expense of inducing many poor people to reduce their work effort substantially. Large income increases and small reductions in work effort may be achieved at the expense of large budgetary costs. Small labor supply reductions and low budgetary cost may be achieved only by providing inadequate income floors to poor families.

THE SIMULATION MODEL

The microsimulation model is described in the papers by Beebout and Maxfield, and Maxfield and Robins,¹ above. To review briefly, after the data base is prepared, federal, state, and social security tax liabilities of each family are computed. These operations produce a tax liability and a marginal tax rate for each program. Second, the rules of existing transfer programs are applied and transfer income is imputed to each family found eligible, with adjustments to account for the fact that some of the eligible families will

not apply for assistance. The model of current transfer programs produces the payment and the effective benefit reduction rates of the AFDC and Food Stamps programs.² Then, the hypothetical NIT plan is simulated in a manner similar to that of the existing programs, except that every filing unit eligible to receive a payment is assumed to participate.

For the purposes of the NIT simulation, we treat a filing unit as made up of the primary nuclear family in each household and all other nuclear families in the household whose head is related to the primary family; other nuclear families in the household whose head is not related to the primary family form separate filing units. We also treat filing units whose heads are older than 64, younger than 15, in the military, or in an institution and filing units with no adult head as not eligible for an NIT payment.

The NIT payment is defined as the basic support, less a tax equal to a portion of earned income (and all unearned income). The support--the payment the unit would receive if it had no other income--is expressed as a percentage of the poverty line for the unit. The amount of earned income included in the calculation depends on the benefit reduction rate of the plan.

The third simulation--that of the labor supply response to the NIT--consists of behavioral rules governing the employment decisions of each healthy, non-institutionalized family head and spouse. This simulation assumes that: (1) The welfare reform changes the disposable income of many families and the wage rates, net of tax rates and transfer program benefit reduction rates, of many people; (2) The change in family income and personal effective wage rates change the amount of time that these people want to

be employed; (3) These people change their labor supply, which changes the amount of money they earn; and (4) The change of earnings causes a change in the amount of taxes for which the person is liable and a change in the payment that the person is eligible to receive.

The simulation of the labor supply response uses wage and income parameters derived from the Seattle and Denver income maintenance experiments (SIME/DIME). (See the papers by West and Robins, and Maxfield and Robins, above.)³ Since the wage and income effects reflect the removal of all major public assistance payments, the response in hours of work is the combined response to that removal and to the introduction of the NIT. Note that some families may experience losses in net income, reductions in cumulative tax rates, or both, and for them, the NIT may induce an increased labor supply. The model is also designed to allow for a positive work response by people who do not work; as such people have no observed wage rate, but they may be engaged in productive activities in the home, a wage⁴ is imputed to them.

An annual accounting period is assumed in both the model of the NIT and the model of labor supply response. That is, the annual NIT payment is computed once in the simulated year using annual income and each potential worker is assumed to determine annual employment once per year based upon the annual NIT payment and the NIT benefit reduction rate.

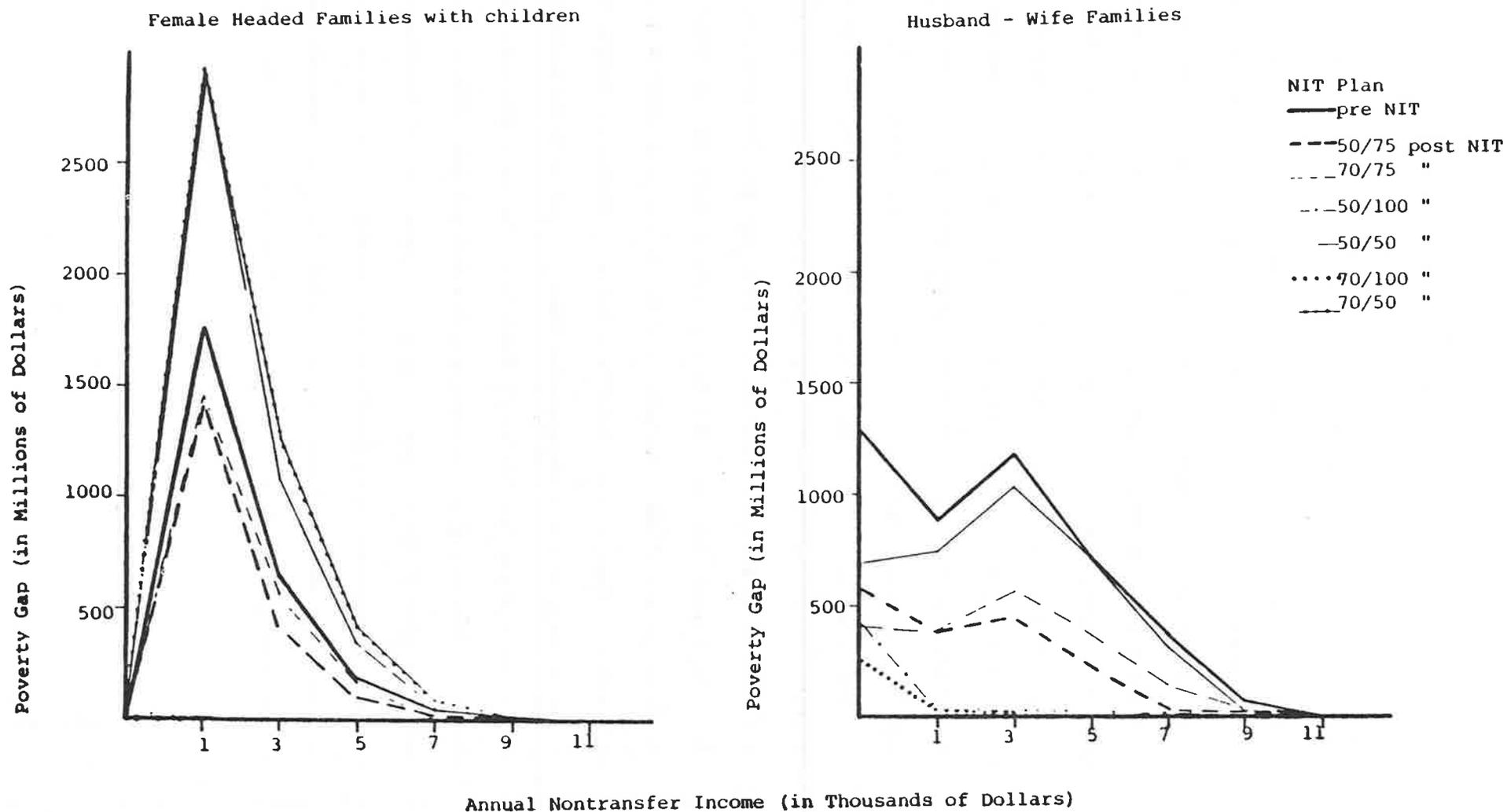
RESULTS

Adequacy of Support

The poverty gap serves as a measure of adequacy of support under the NIT plans tested. Figure 1 shows the poverty gap facing husband-wife and

FIGURE 1

POVERTY GAP BY NIT PLAN
FEMALE HEADED FAMILIES WITH CHILDREN AND
HUSBAND-WIFE FAMILIES



female-headed families and is distributed across pre-transfer income categories. The current poverty gap represents the aggregate income amount needed to boost the current disposable income, including transfer payment, of all units up to the poverty line. The poverty gap under each NIT plan represents the difference between the poverty level for the unit and its post-labor supply response disposable income, including the NIT benefit.

Not surprisingly, the figure shows that the poverty gap decreases as the NIT plans become more generous. This is true at all income levels. In contrast, lowering tax rates at a given support level reduces the gap at upper income levels but has little effect on low income units.

The least generous plans increase the poverty gap for families composed of children and a mother only. For them, these NIT plans are less generous than current programs. Since current programs are directed primarily to these one-parent families, and do not provide much aid to other groups, even the least generous plans register a reduction in the poverty gap for husband-wife units. Plans of intermediate generosity provide a slight improvement in income support for female-headed families and more substantial reductions in the poverty gap for two-adult families. The most generous plans come close to eliminating the poverty gap for all groups. Some units remain below the poverty line under these plans since the NIT reimburses only a portion of negative earnings (business losses). Other units face such a high earnings benefit reduction rate under existing programs that a NIT can not be sensibly simulated. A NIT benefit was not calculated for these units.

The poverty gap under the various NIT plans tested is summarized in Table 1. In general, increases in the support level decrease the poverty gap but at a decreasing rate. At the higher benefit reduction rate, increasing the support from 50% to 75% of the poverty line decreases the poverty gap by six billion dollars. Increasing the support from 75 to 100% decreases the poverty gap by five billion dollars. A similar pattern holds at the lower tax rate.

Changes in the benefit reduction rate, or tax, are less potent modifiers of the poverty gap than are changes in the support level. At either the 50% or the 75% support levels, reducing the tax rate from 70% to 50% decreases the poverty gap by one billion dollars.

NIT Program Costs

Table 1 lists the total program cost, or the sum of the payments to all participants, for the six NIT plans before any labor supply adjustments take place. The least expensive NIT plan costs about seven billion dollars. In comparison, the total cost of the AFDC and Food Stamps programs in 1974 was ten billion dollars. At a 70% benefit reduction rate, increasing the support level to 75% increases the cost by six billion dollars, to 16 billion. Increasing support to 100% of the poverty level adds another ten billion dollars. At the 50% benefit reduction rate, increasing the support level from 50 to 75% increases costs by eleven billion, and increasing it from 75 to 100% adds another 21 billion dollars. Thus, regardless of the benefit reduction rate, increasing the support level increases program costs at an accelerating rate.

TABLE 1

ESTIMATED REFORM AND CURRENT SERVICES PROGRAM COSTS AND CASELOADS
(in billions of dollars, millions of units and millions of hours)

Description	Support (Percentage of Poverty Line)		
	50%	75%	100%
	Amount	Amount	Amount
.50 Benefit Reduction Rate			
Gross Cost	8.1	19.0	40.3
Net Cost	-1.8	9.1	30.4
Post-Reform Poverty Gap	10.8	5.0	.6
Change in Earnings	.001	.006	.015
Change in Total Hours	.7	2.0	4.7
Pre-response Caseload	6.8	14.1	25.1
Change in Cost Due to Response	.7	2.0	4.8
.70 Benefit Reduction Rate:			
Gross Cost	6.6	13.2	23.9
Net Cost	-3.3	3.3	14.0
Post-Reform Poverty Gap	11.6	5.7	.3
Change in Earnings	.001	.003	.006
Change in Total Hours	.5	1.2	2.5
Pre-response Caseload	4.9	7.9	12.3
Change in Cost due to Response	.4	1.5	3.7
Current Services			
Cost (AFDC, Food Stamps)		9.9	
Pre-Reform Poverty Gap		11.2	
Caseload ¹		5.6	

SOURCE: Simulation by Mathematica Policy Research, Inc. and The Hendrickson Corporation.

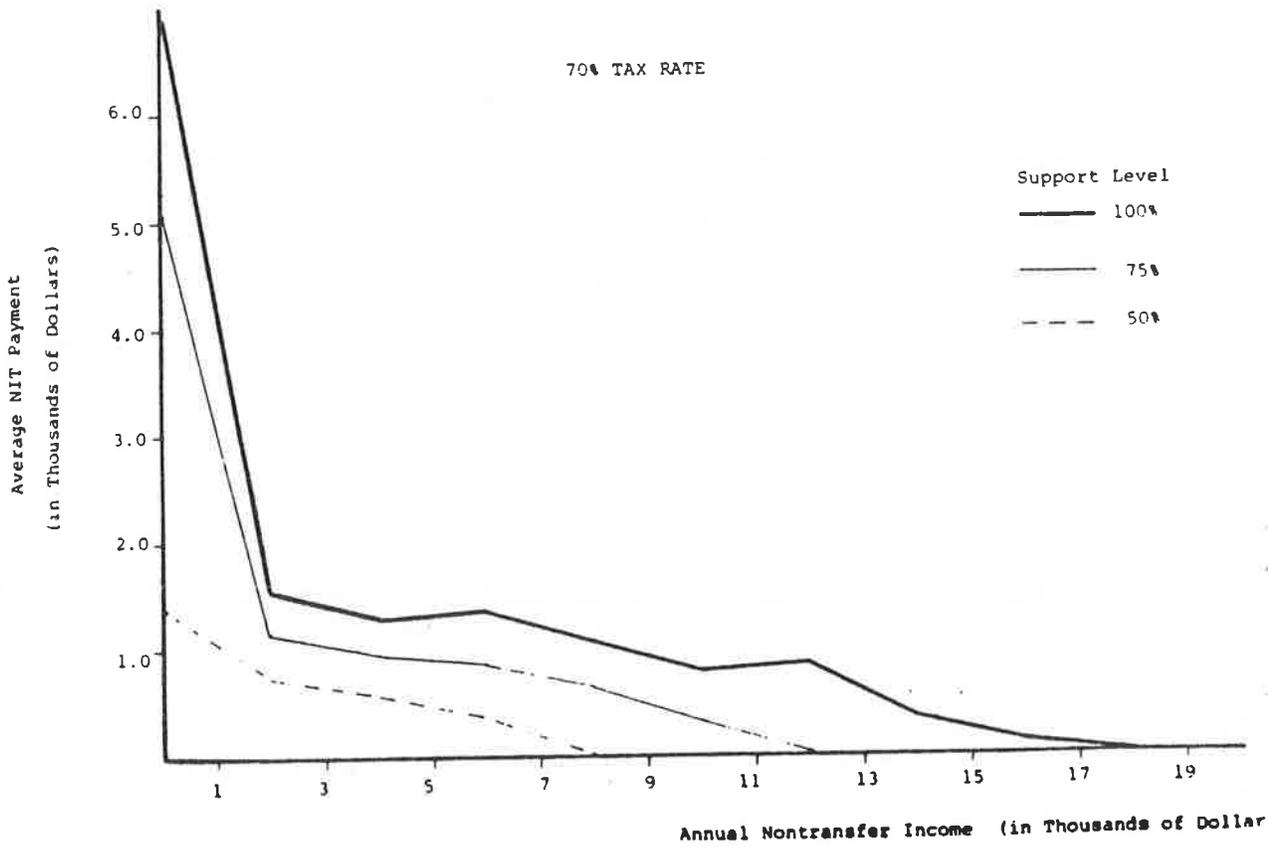
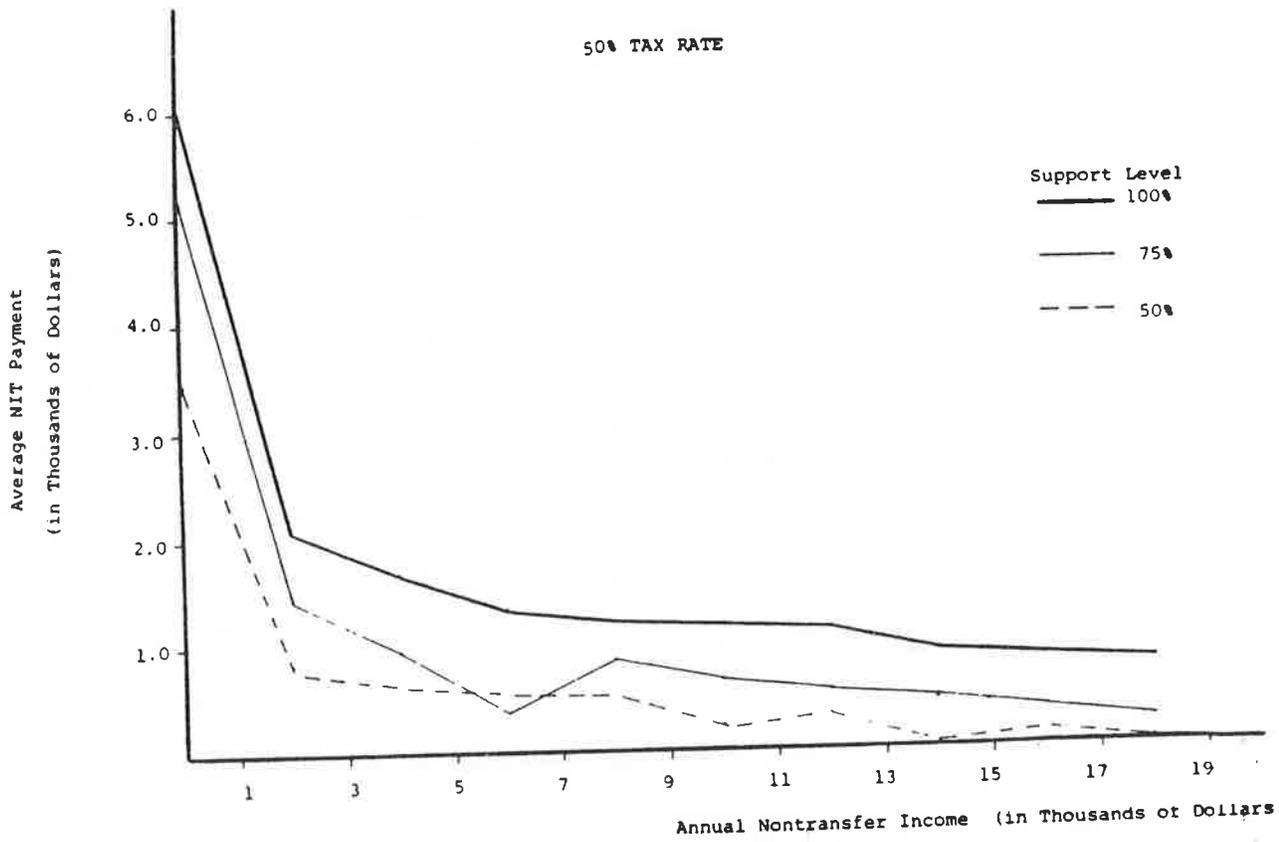
NOTE: 1. For comparison purposes, current services caseload is expressed as the number NIT filing units which received AFDC, AFDU and/or Food Stamp benefits.

This occurs because cost of an NIT is determined both by the size of payments to a given family and by the number of families that participate in the program. Table 1 also shows the number of families simulated to receive an NIT payment. (Such families are called participants for convenience, although the use of the term assumes that every eligible family participates.) Approximately five million families participate in the least costly plan, the plan with the 50% support and 70% tax. At that benefit reduction rate, increasing the support level to 75% adds three million families, and increasing the support to 100% adds another four million families. At the 50% benefit reduction rate, increasing the support level to 75% adds seven million families, and increasing the support to 100% adds another eleven million families. Thus, the number of families eligible for an NIT payment increases disproportionately with the percentage change in support level, regardless of benefit reduction rate. This occurs because with a higher support level, families with higher income become eligible for payments and there are proportionately more families as we move up the national income distribution curve towards the middle income group. At the 50% support level, decreasing the tax to 50% adds two million families. Reducing the benefit reduction to 50% at a 75% support adds about seven million families, and with a 100% support, almost thirteen million families. In sum, reducing the benefit reduction rate disproportionately increases the pool of eligible families at higher support levels, compared to lower support levels.

The relative generosity of the six tested NIT plans is presented in Figure 2, which shows the average annual NIT payments received by participants

FIGURE 2

AVERAGE NIT PAYMENT PER ELIGIBLE FAMILY



in each plan. The nonlinearity of average payments with respect to income is due to the reimbursement of taxes, the distribution of family size over family income categories, and the proportion of earned to unearned income in each total income category. Payments at zero income vary by support level and, as income rises, lower benefit reduction rates cause payments to fall more slowly than do higher rates.

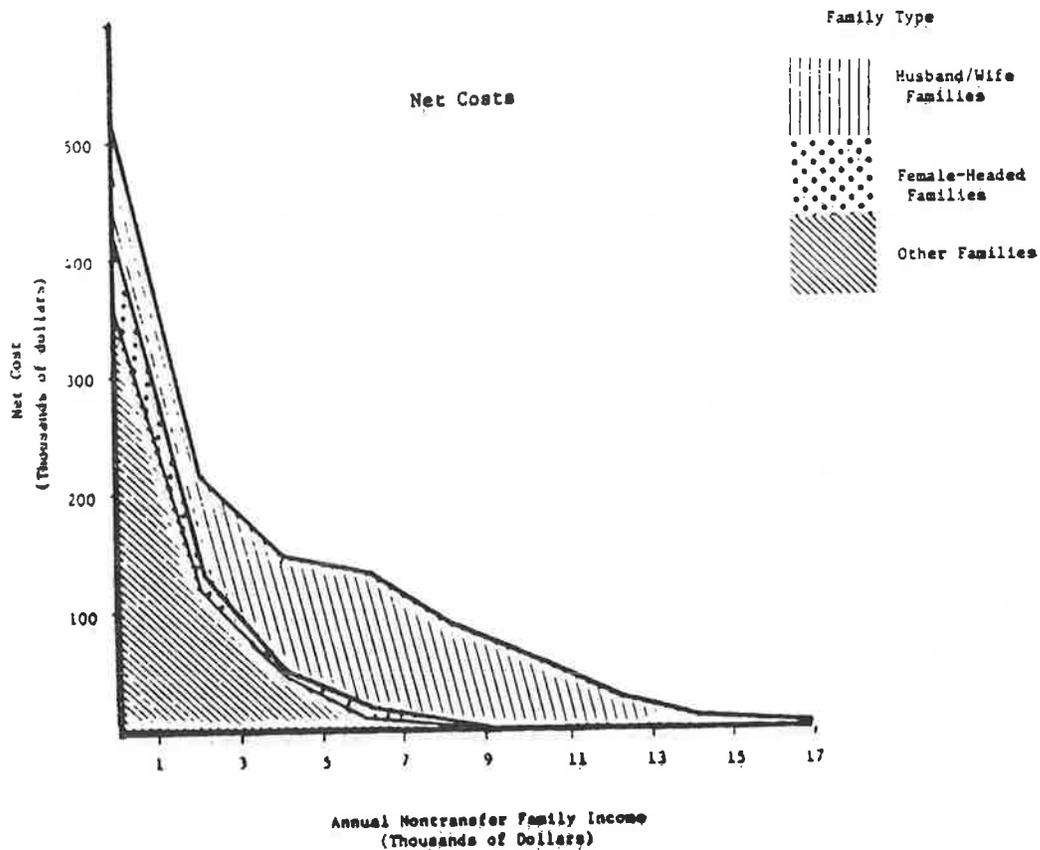
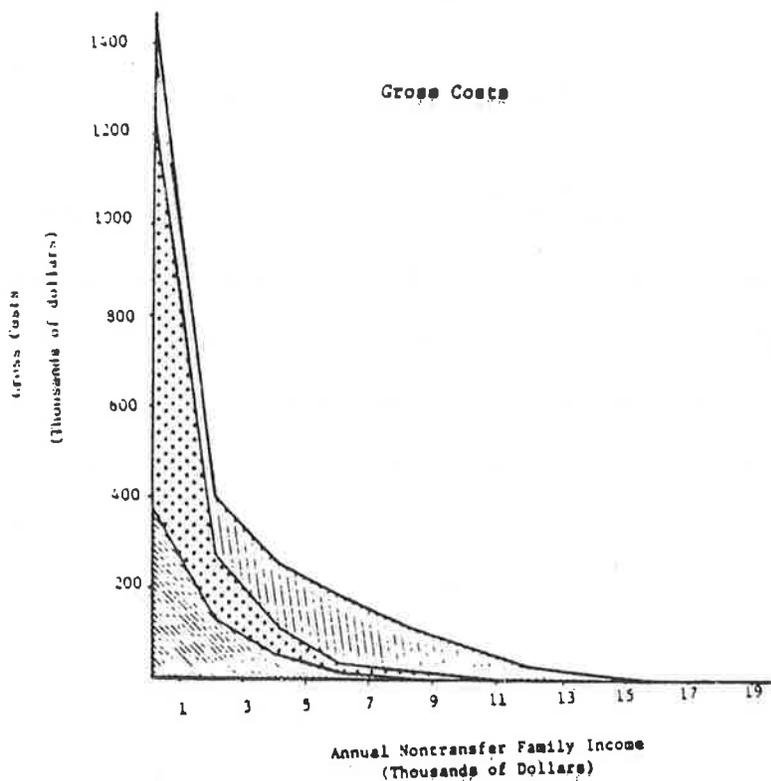
The differential effects of replacing current welfare programs with an NIT among types of families is shown in Figure 3. The figure shows the gross costs (total NIT payments), and net budget costs (gross costs less AFDC and Food Stamps) by family type under a 75% support/50% tax plan. The largest proportion of gross payments at low income levels goes to female-headed families with children. This reflects the composition of the low-income population. As income rises, the payment total (the highest curve) declines smoothly but the composition shifts from female-headed families with children and other families (single male heads and females without children) to families headed by a couple.

Net costs present a different picture. The proportion of total NIT payments exceeding the total AFDC and Food Stamps payments that are received by female-headed families with children is small throughout the whole income range. This occurs because the largest portion of current welfare expenditures is being received by families, with children, headed by a woman.

The net costs of a 75% support/50% tax plan are positive for every income category. In other words, taken as a whole, each income group would be better off if we replaced existing welfare programs with this

FIGURE 3

GROSS AND NET COSTS BY UNIT TYPE - 50/75 NIT PLAN



base NIT. Among individual families within each income category, some families may be made worse off by the new plan.

Figures 4 and 5 provide a comparison of the income distribution of costs and participants among the six NIT plans. The plans are ordered by cost as they are by the number of participants, with the most costly having the largest number of participants. All but two NIT plans--the 50% support plans--are more costly over most of the income range than current welfare programs. The 50% support level reduces the total transfer payment amounts going to most of the NIT participants.

The uppermost curve of figure 5 represents the income distribution of the total number of families, both participating and not participating. All the NIT's include virtually all families with no income. As income increases, the more generous plans include a greater proportion of the population. Many low-income families are ineligible for NIT payments because of small family size or a high proportion of unearned income.

Figure 5 indicates the relative target efficiency, or the ratio of the number of low-income families eligible for a payment to the total number of families eligible for a payment, of the several NIT plans. In general, the less generous plans are seen to be more target efficient. The changes of the benefit reduction rate that are tested affect target efficiency more than the tested changes of the support level. This difference is stronger among the more generous plans, whereas both the benefit reduction rate and the support level are equally influential on the target efficiency of the less generous plans.

FIGURE 4. TOTAL GROSS NIT COSTS BY NIT PLAN

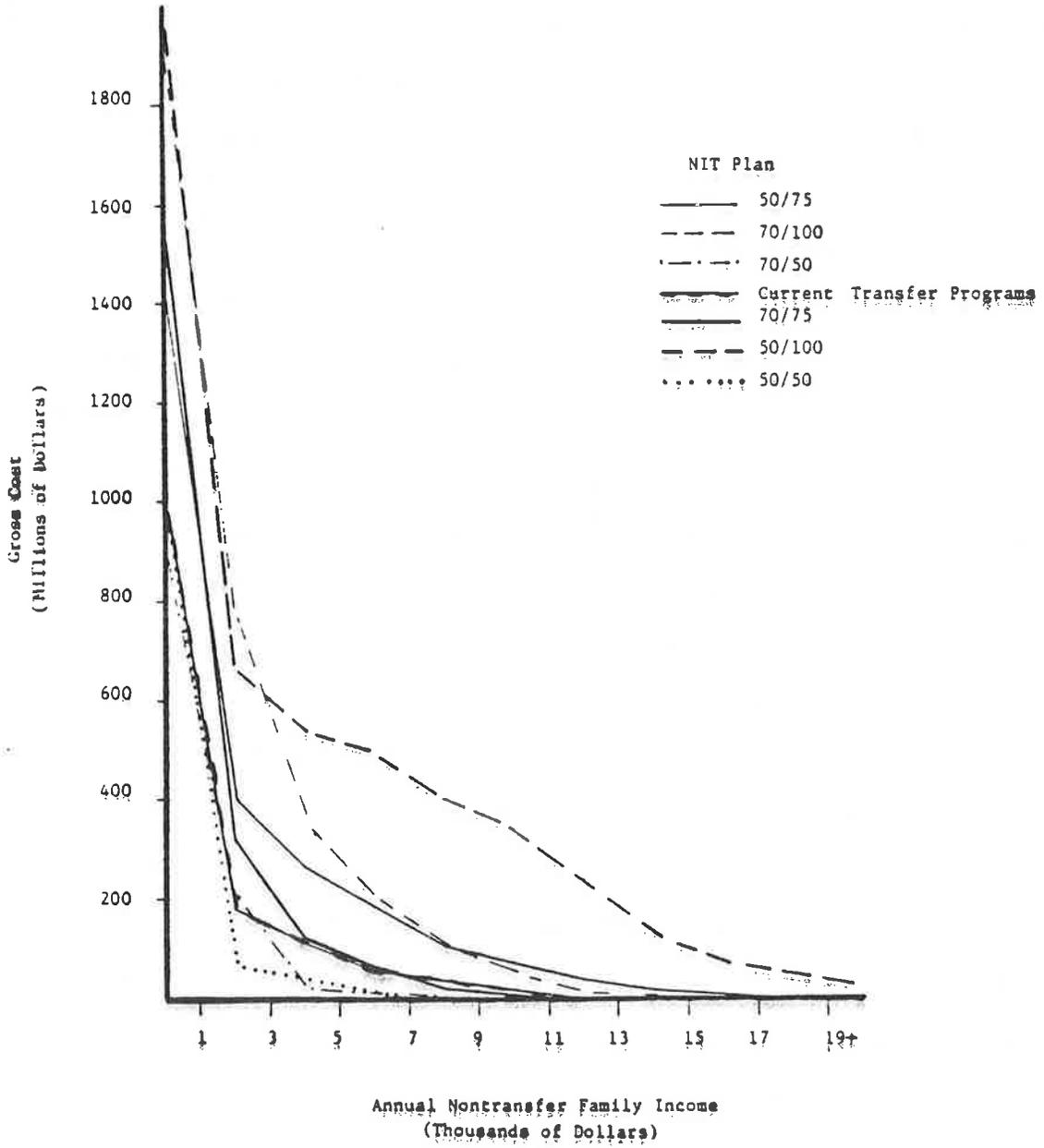
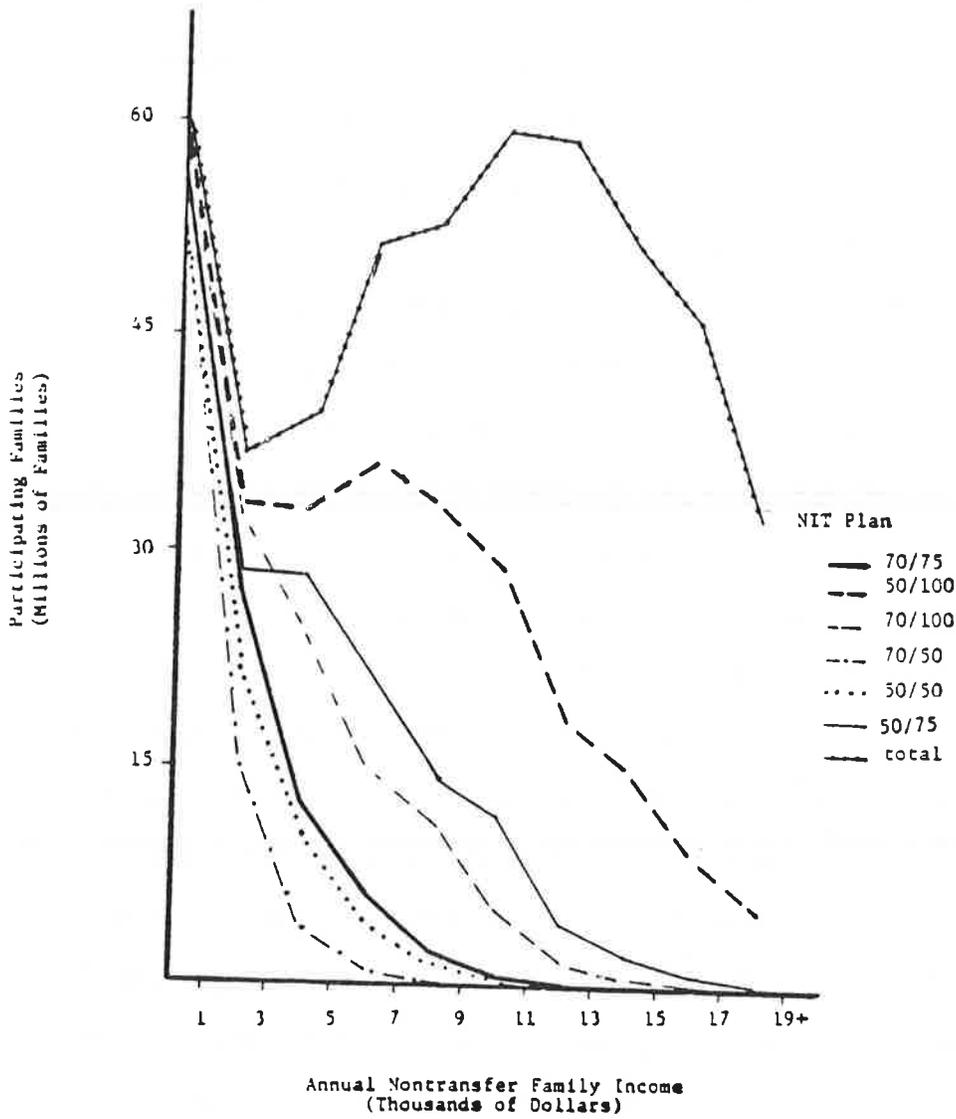


FIGURE 5. TOTAL NUMBER OF PARTICIPATING FAMILIES BY NIT PLAN



Labor Supply Response

It should first be noted that because labor supply and program costs are calculated for those families which received NIT benefits based on income reported on the Current Population Survey, they do not include families above the tax break-even point which may have reduced their labor supply until they were below the tax break-even point. These results ignore those who would become participants in the program due solely to their labor supply response.

Figures 6 and 7 show the effects of the several NIT plans on the work effort of recipients. The labor supply effects are displayed in terms of the total and average number of annual hours of employment per recipient. The response of average hours per recipient measures the effect of the NIT on individual work effort, while the response of total hours (listed in Table 1) is a measure of the amount of goods and services lost to society due to the tested welfare reforms.

Figure 6 presents the change in average annual hours of employment of husbands, wives, and female heads of families with children induced by an NIT with a 75% support and 50% tax. Husbands exhibit a fairly uniform percentage reduction of hours of employment, ranging between 4 and 12%, and averaging 7%. Wives are simulated to have much larger percentage reduction of labor supply, ranging from 10 to 30%, averaging around 20%. The labor supply of single female heads ranges from 0 to 17%, and averages roughly 2%. All of the responses decline proportionately more as pre-transfer income increases because the NIT affects the income of low-income families more than that of high-income families. The response of female family heads declines faster as income increases,

FIGURE 6

AVERAGE LABOR SUPPLY RESPONSE AND PERCENT CHANGE IN HOURS
 FOR HUSBANDS, WIVES, AND FEMALE HEADS WITH CHILDREN
 50/75 NIT PLAN

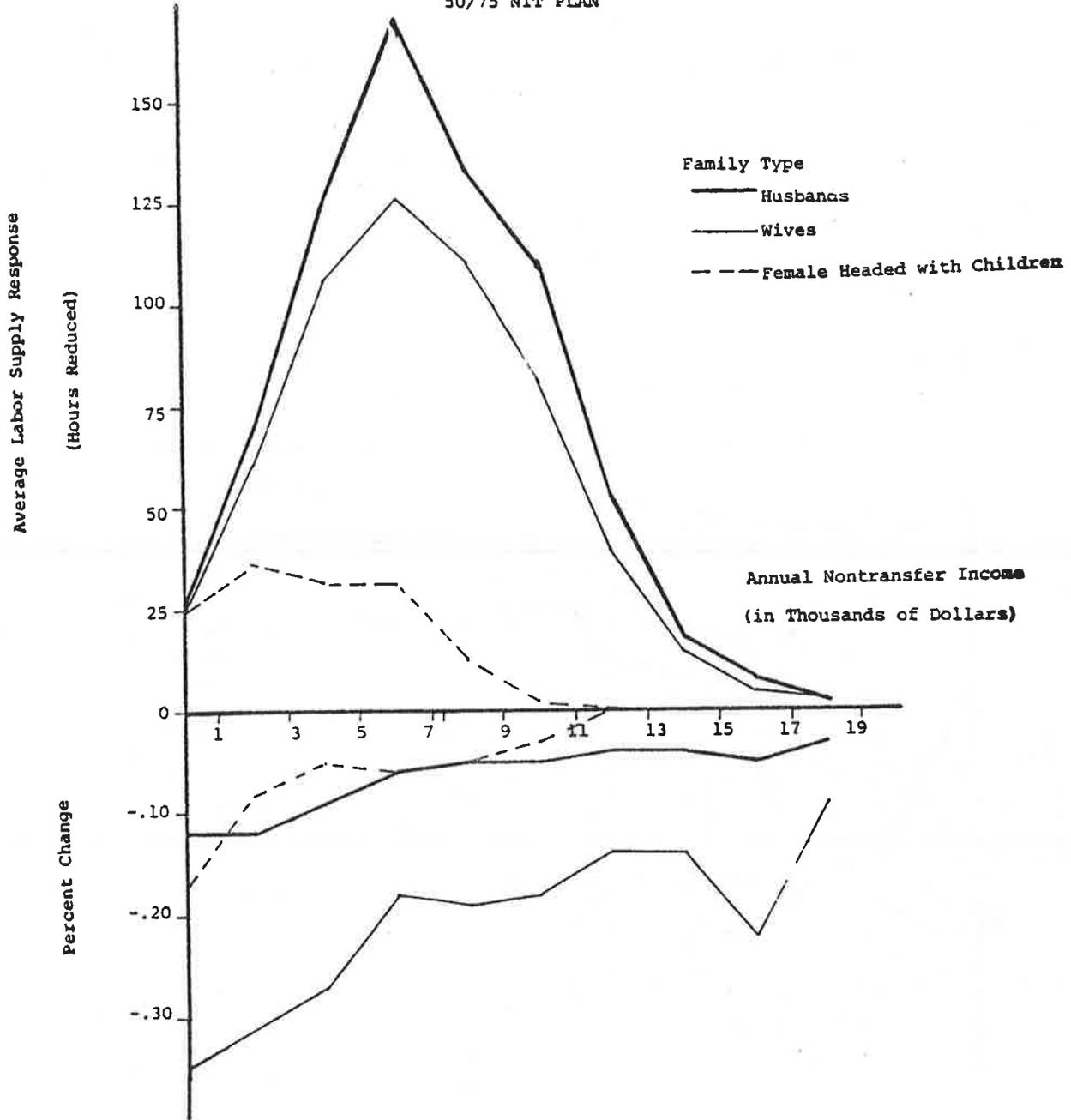
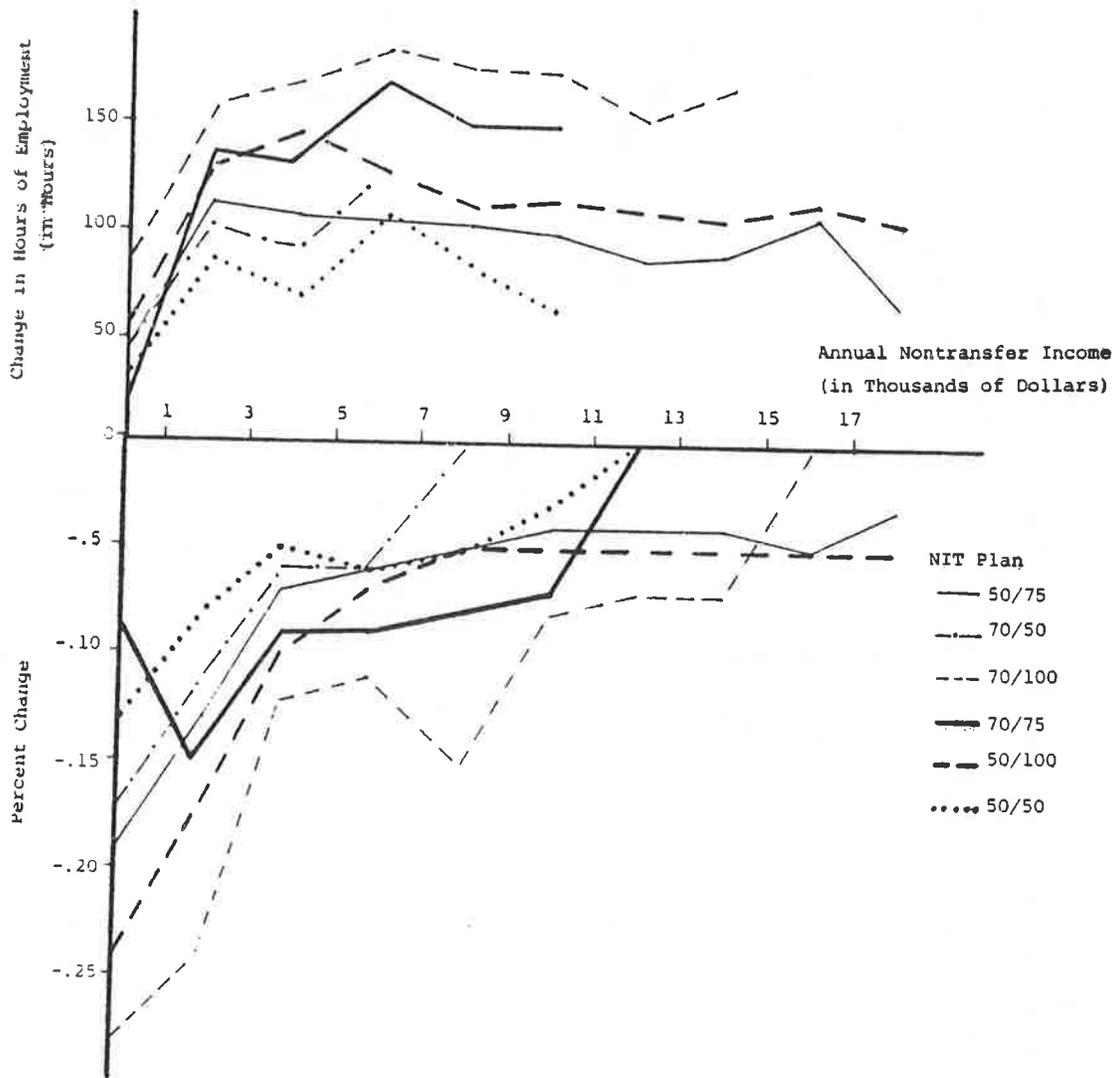


FIGURE 7
 AVERAGE LABOR SUPPLY RESPONSE AND PERCENT CHANGE BY NIT PLAN



compared to the responses of the other groups. Overall, the average response is smallest for female family heads and largest for wives.

Comparisons among the total labor supply responses to the six NIT plans are presented in Figure 7 and in Table 1. These comparisons show that the more generous the support levels, the greater the reduction of employment. Lower, more generous tax rates induce less response than do higher rates. The effect of the support level dominates the effect of the tax rate.

Based on these reductions in hours of employment, multiplied by wage rates, we can calculate a lower level of earnings and higher level of NIT payments for each family. Cost increases among high income filing units is disproportionately large, as a reduction of one hour of work by a high-wage person increases the NIT cost more than it does for a low-wage person. The change in total NIT gross budgetary costs caused by the induced labor supply reduction is shown in Table 1. For all plans, changes are positive and the cost is greatest for those NIT plans which induce the greatest labor supply reduction.

IMPLICATIONS FOR PROGRAM DESIGN

Increasing the level of income support of an NIT produces a more than proportional increase in the budgetary cost of the program, due both to higher payments and to larger numbers of eligible families. Changes in the two tax rates tested resulted in larger cost differences than did the changes in the three support levels. This difference occurs because a reduction of the tax rate raises the NIT break-even level and increases the pool of participants more than a roughly comparable percentage increase in the support level.

More generous NIT plans are seen to induce larger reductions in work effort. Both reduction of the total amount of employment and the reduction of the individual hours of work are greater under more generous plans, with the support level a more potent parameter than the tax rate.

Combining the results on labor supply and cost suggests a low-support-level, high-benefit-reduction-rate plan to minimize both costs and labor supply reduction. This is the least generous plan, however. To achieve the goals of lowest cost and smallest reduction of work effort, it is necessary to sacrifice income support. The trade-off is summarized in Table 1. As NIT plans are made more generous, income support is improved at a decreasing rate while both costs and labor supply reduction rise at an increasing rate.

The policy makers may also want to know whether an NIT will affect different families in different ways, compared to existing transfer programs. Two main differences appear. First, since the AFDC program focuses on families composed of a mother and children, the NIT changes the status of this group less than other units. Thus, we see a smaller reduction of labor supply by female heads compared to either spouse in husband-wife families. Second, several effects of the NIT vary by family income. Because the model used the federal poverty line (which varies by family size) and deducted 100% of unearned income from the NIT payment, a number of low-income families remained ineligible for the NIT. On the other hand, positive taxes were reimbursed, making relatively high-income families eligible for payments under the more generous NIT plans. The number of these eligible high-income families is sensitive to the benefit reduction

rate. The percentage labor supply reduction is greatest among low-income families, although the impact of the reduction on NIT program costs is greatest among middle-income participants. The least generous NIT plans are relatively more effective at targeting benefits to the low-income population.

NOTES

1. See also Harold Beebout, "Microsimulation As a Policy Analysis Tool: MATHLAB," (Mathematica Policy Research, Inc., Washington, D.C., 1977) and Myles Maxfield, "Estimating the Impact of Labor Supply Adjustments on Transfer Program Costs: A Microsimulation Methodology," (Mathematica Policy Research, Washington, D.C. 1977).
2. Effective AFDC benefit reduction rates are those estimated empirically by Douglas Bendt, "The Effects of Changes in the AFDC Program on Effective Benefit Reduction Rates and the Probability of Working," (Mathematica Policy Research, Inc., Washington, D.C. 1975) from the 1973 AFDC survey of case records. For the Food Stamp program, the simulation uses the average benefit reduction rate over a small income range.
3. See also Michael Keeley, et al., "The Labor Supply Effects and Costs of Alternative Negative Income Tax Program: Evidence from the Seattle and Denver Income Maintenance Experiments Part I: Labor Supply Response Function," (Research Memorandum 38, SRI International, Menlo Park, California, 1977).
4. The imputation equation was estimated using data from the Dual Job Supplement to the May 1973 Current Population Survey. For a description of the equations, see Myles Maxfield, "Estimating the Impact of Labor Supply Adjustments on Transfer Program Costs: A Microsimulation Methodology," (Mathematica Policy Research, Inc., Washington, D.C., 1977).

DISCUSSION

Of the two authors, Dr. Maxfield attended a different workshop and Dr. Edson presented this paper.

DAVID PARKER: You said you ignored people who were above the break-even point at the outset. How much of an understatement do you make by ignoring those people?

EDSON: The paper by Maxfield and Robins includes those people. I think that they account for anywhere between \$200 million to two billion dollars, depending on the generosity of the plan. This can be anywhere from 30 to 50%.

HAROLD BEEBOUT: Don't you think that we should view those results with considerable suspicion? It depends greatly on how the program is administered. These people that are sitting up here above the break-even level could be thinking about jumping in.

EDSON: Yes. Because I have reservations about the methodology used in the Maxfield and Robins paper, I didn't include it in this paper.

PARKER: Are you saying that your results are not sensitive to "drop-ins" from above the break-even level?

EDSON: No, because of my reservations concerning the methodology used, I have omitted people above the break-even point from the analysis.

RONALD DEAR: Why did they use the poverty line throughout the program? I have questions about how good a measure that is of anything.

BEEBOUT: The plans were developed in response to two different questions. First, how was the generosity of the plans chosen to be tested? They were set according to an estimate of what could be afforded. The higher the plan, the more it's going to cost. Second, what was the range of policy space that needed to be examined, and how much needed to be expended on the experiment in order to get statistically significant responses? That does not have much to do with the poverty line. It was convenient to display the results in terms of poverty line ratios because they are used so frequently.

Now I have a question, although it may be rhetorical. I'm looking at your Table 1. A program that has a net cost of \$30 billion is not going to be legislated anywhere in the remote future. It will be difficult even at the \$10 billion level. Yet, if you set the support level at anything less than 75% of the poverty line, you are making lots of current recipients worse off. Can you infer from that that a simple NIT has been proved infeasible by SIME/DIME, and that there's no point in this policy space that can achieve your objectives?

EDSON: I think that you can't have both adequate income support and low cost and low work response. This raises some questions about whether the NIT is the proper vehicle for comprehensive welfare reform.

DEAR: It's very interesting. You are going back to some of the thinking developed initially by some of the best-known welfare people in the country-- E. Burns, Mitch Ginsberg, Alvin Schorr.

EDSON: I don't think that this has really surprised anybody in retrospect, but the major thrust of our paper is to provide estimates of what the magnitudes of the trade-offs really are.

BEEBOUT: We have been surprised, not by the direction of things, but that things now look less feasible because of the magnitude of the labor supply response. If you didn't have a labor supply response, you could have a fairly generous guarantee with a high tax rate, and you can eliminate most of your poverty gap, and achieve all your objectives. But SIME/DIME shows the labor supply response to be too high.

PARKER: In a sense you're only talking about net costs of 20%.

BEEBOUT: Roughly five out of \$35 billion.

DEAR: That's five billion dollars lost because of labor response.

PARKER: Out of \$35 billion of your "net profit." I'm not sure that you're saying proportionately that's too large a sacrifice or anybody would like to pay that much for leisure.

BEEBOUT: I guess I was making some value judgment that Congress would not be interested; they may even not like the \$25 billion cost, without the labor supply response.

DEAR: You show the pre-reform poverty gap as \$11.2 billion, and you are talking about a program that's going to be pumping in around \$30 or 35 billion. Why doesn't that more than adequately cover that gap?

EDSON. The more generous programs are not target efficient.

BEEBOUT: You have several surprising results, don't you? As another example, people would not have guessed--prior to seeing these results-- that you would find that a 100% support level would induce a greater labor supply response in the aggregate with a high benefit reduction rate, compared to a low one.

EDSON: Yes.

BEEBOUT: This has really confounded policy makers. They always thought before that you wanted to have a low benefit reduction rate even though it cost a bit more.

EDSON: It wasn't clear which would be more powerful, the benefit reduction rate or the support level. It turns out that the support level is more powerful.

BEEBOUT: The horns of your dilemma are equally severe no matter what point you're at.

DEAR: That's why we've had this welfare dilemma for so long. Anybody who studies it over the years knows it. It's gone on literally for hundreds of years.

THE MACROECONOMIC IMPACT OF A NEGATIVE INCOME TAX

by

Robert A. Moffitt
Assistant Professor of Economics, Rutgers University
and Research Economist, MPR

Understanding the impact of a negative income tax (NIT) on the national economy is extremely important to welfare reform. Unfortunately, however, there has not been a great deal of research on this topic, and we are presently at an infant stage of development in our knowledge of such effects. To some extent, this is a natural result of the way in which research on an NIT has evolved: the impact of the NIT experiments on an individual's work effort was naturally studied first, and generalizations of these impacts from the experimental populations to the national population was a natural second step. However, estimates of the macroeconomic impact of an NIT, such as are presented here, rest on a very different analysis, and also involve considerably more uncertainty simply because our understanding of many macroeconomic forces--inflation, unemployment, income distribution--is so poor.

This paper summarizes the results of two research efforts to estimate selected macroeconomic effects of an NIT. Both efforts focus on a single subset of questions, in order to simplify the problem. The first of these research efforts, performed by the author,¹ involves the impact of an NIT on the labor market. That an NIT will have an impact is unquestionable, but the size of the impact is unknown. As discussed in other papers in this collection, an NIT will lower work effort, or labor supply, in the nation as

a whole. However, the analysis should not end here: the reductions in aggregate work effort are likely to have impacts on the aggregate levels of employment and unemployment, the level of wages, and the rate of inflation. These effects are likely to have secondary effects on the supply of work effort itself, thus changing the original estimates of the change in the labor supply. This research went one step beyond the national work-effort reductions reported above in the papers by MPS and SRI to the effect of an NIT on the labor market as a whole.

The other piece of research, also summarized below, examines the impact of an NIT on the consumption patterns of the population.² If NIT increases the disposable incomes of the low-income population, it will also increase the demand for goods and services. If the NIT is financed through the federal income tax system, the disposable incomes and expenditures for consumption of higher-income groups will decline. This second research effort traces the effects of these changes in consumption on employment and the distribution of earnings.

IMPACT OF NIT ON LABOR MARKET

In a competitive labor market, where there are numerous employers and numerous workers looking for jobs, at the equilibrium wage level the number of workers needed by employers equals the number of workers willing to work. The equilibrium level of employment equals the number of workers willing to work at this wage level. This notion of "equilibrium" is usually adjusted to account for a certain number of people looking for work--that is, for some "normal" level of unemployment. (Normal unemployment is generally

thought of as "frictional," composed of workers routinely searching for new jobs, and quitting old ones. Not too long ago it was thought that the normal level of unemployment was about 4%, although many now think it is higher.)

In this type of labor market, an NIT that causes some workers to reduce their work effort will increase the equilibrium wage level and reduce the level of employment. As soon as workers leave the labor market, employers will be forced to raise wage rates to attract them back. As a general rule, the increase in wage rates forced upon employers will cause them to lower their desired number of workers, either by substituting other elements in production (such as capital) for labor or by requiring higher productivity per employee.³ Consequently, employers will generally raise wages enough to attract some, but not all, of their work force back into the labor market. The new employment level will be lower, and the wage level, higher than before.⁴

Mathematically, for each 1% decrease in the supply of labor, the wage level increases by a percent equal to $\frac{1}{\epsilon_s + \epsilon_d}$ and the employment level decreases by a percent equal to $\frac{1}{\epsilon_s + \epsilon_d}$ where ϵ_s is the "elasticity" of labor supply and ϵ_d is the "elasticity" of labor demand. The elasticity of labor supply is the percentage increase in the supply of workers after a 1% increase in the wage level, and the elasticity of labor demand is the percentage decrease in the demand for workers after a 1% increase in the wage level.

Two special cases are of note here. First, where demand is completely "inelastic" ($\epsilon_d = 0$), the NIT does not produce any decrease in employment. In this case, employers' demands for workers are insensitive to a change in the wage, and they simply keep raising wages until they attract their entire

work force back into the labor market. Another special case, that assumed in the simulations discussed elsewhere in this collection, is that the elasticity of the labor demand is infinite ($\epsilon_d = \infty$), so that even the slightest increase in the wage level causes the demand for labor to fall to zero. In this case, there is no change in the wage level, and the employment multiplier $\frac{\epsilon_d}{\epsilon_s + \epsilon_d}$ equals 1, implying that a 1% reduction in labor supply also reduces employment by 1%.

More realistically, neither of these special cases is likely to hold in the real world. Labor demand would not fall to zero with an increase in the wage level, although it would fall somewhat. Likewise, labor supply is likely to increase, although probably not by much overall. For example, if $\epsilon_d = 1$ and $\epsilon_s = .2$, the wage level would rise and employment would fall by .83% for every 1% decrease in labor supply. Thus, all the employment reductions reported in previous simulations would have to be lowered by .17% per percentage point reduction in labor supply.

These calculations are based upon the assumptions that the labor market is in equilibrium before and after the NIT is introduced, and that the adjustment from one equilibrium to the next is quick. Neither assumption is realistic. At the present time, with the national economy operating below capacity and with high levels of involuntary unemployment, the model seems particularly unsuitable. Even though the economy is recovering, the speed of adjustment is anything but quick.

The effect of an NIT under these circumstances is likely to be quite different than discussed above. For example, if workers who are involuntarily

unemployed leave the labor force (i.e., stop looking for work) because of the NIT, it will have no effect on the employment level. But also, if any workers leave their jobs because of the NIT, they may be replaced by workers who were involuntarily unemployed. (I call this the "replacement effect.") In any case, whether the NIT withdrawals are from the ranks of the unemployed or from the employed, the unemployment rate (defined as the number of unemployed divided by the number employed or unemployed) will fall and any reduction in aggregate employment will be muted.

The rate of wage inflation will likely increase as the unemployment rate falls. Even though there does exist a pool of involuntarily unemployed, employers who wish to hire workers will probably have to increase wages more when that pool is small than when it is large. (The oft-observed inverse relationship between unemployment rates and wage inflation was originally interpreted as a manifestation of just this type of labor market adjustment.⁵) As the rate of wage inflation rises, some workers will tend to return to the labor market and some employers will tend to reduce their desired number of employees; both of these factors tend to slow the decline in the unemployment rate, if any.

Thus, estimation of effects of an NIT is considerably more difficult than the simple calculations reported above. Assumptions must be made regarding the rate of labor-supply adjustment (the rate at which workers enter the labor force as wages rise), the rate of labor demand adjustment (the rate at which employers reduce their desired work force as wages rise), and the rate at which wages rise as unemployment falls. Estimates of these rates have

been taken from the existing empirical literature in order to calculate the short-run effects of an NIT on the labor market.⁶

Table 1 shows the effects of an NIT on the labor market, where it induced a 1.6% reduction in aggregate labor supply.⁷ If the economy is at a high level of unemployment (8%), the unemployment rate would fall by as much as 1.5% and the employment level would fall by no more than .34% after a year. After two years, the employment level would have fallen by no more than .67%. Thus, the NIT would have little effect on the (depressed) level of employment. The rate of wage inflation would increase by no more than .50% after one year and .60% after two years.

However, if the initial unemployment rate is 5% (presumably close to "full employment"), somewhat bigger reductions in employment and bigger increases in the rate of wage inflation would occur. After one year, employment could fall by .45%, and, after two years, from .09 to .64% (higher, on average, than before). The rate of wage inflation could increase by 1.2% after one year and by 1.0% after two years, at maximum.

IMPACT OF AN NIT ON CONSUMPTION AND EARNINGS DISTRIBUTION

The above analysis concentrates exclusively on the labor market, by taking, in economists' words, a very "partial equilibrium" approach. One of the most important omissions in that analysis, from the standpoint of the economy as a whole, is the impact of increase in incomes for low-income families on the market for consumer goods. The increase demand for goods created by an NIT that is more generous than existing welfare will have impacts on output, inflation, unemployment, and income distribution.

TABLE 1

EFFECTS OF A NIT ON THE LABOR MARKET

	After One Year			After Two Years		
	Decrease In Unemployment Rate (%)	Decrease In Employment Level (%)	Increase In Rate Of Wage Inflation (%)	Decrease In Unemployment Rate (%)	Decrease In Employment Level (%)	Increase In Rate Of Wage Inflation (%)
If Initial Unemployment Rate is 8 Percent:						
Low	1.1	0.03	0.30	1.4	0.06	0.30
High	1.5	0.34	0.50	0.8	0.67	0.60
If Initial Unemployment Rate is 5 Percent:						
Low	0.5	0.05	0.70	0.8	0.09	0.40
High	0.9	0.45	1.20	0.3	0.64	1.0

Golladay and Haveman⁸ have built a microsimulation model that traces several types of effects an NIT might have. First, they estimate the effect of the transfer system on family incomes, classifying by race, region, and so on. This effect will of course depend upon the eligibility rules of the program, the level of benefits, and the way the program is financed. Second, they estimate the effects of the change in disposable incomes on consumption expenditure for 79 commodity categories within each region of the country. Third, they translate these changes in consumer demand into change in commodity production and output in each region. Fourth, they translate the change in output into changes in the demand for workers, classifying them into 114 different occupations. Fifth (and last), they estimate the changes in employment by occupation on the earnings distribution within each region and for the nation as a whole. These effects are all simulated for an NIT with a guarantee level of \$2400 (1973 dollars) with tax rates of 67% and 60% on earned and unearned income, respectively, and financed by a surcharge on the federal income tax.

The simulation model is an enormously complex and expensive undertaking, and is by far the most detailed and sophisticated estimate of the full-system impact of an NIT which has been developed. However, as the authors admit, the feasibility of the undertaking was possible only by initially making some simplifying assumptions that do not match reality. For example, they assumed a completely elastic supply of labor for all occupations--that is, they assumed that firms would have no trouble finding all the workers needed to fill newly created jobs. (This could occur only in an economy with sub-

stantial unemployment.) Also, they assumed that all prices and wages were fixed--thus ignoring any inflationary impact either in the commodities market or in the labor market, as discussed in the previous section. In addition, they assumed that the NIT would have no effect on the work effort of recipients.

Despite these limitations, the Haveman-Golladay study illuminates our analysis of the economic impact of an NIT in interesting ways. One rather expected result of their study is that there would be a substantial net increase in aggregate output--to the tune of about three billion dollars (1973 dollars)--resulting from the above-described macroeconomic process. This is an underestimate of the final output effects, for, in classical Keynesian fashion, the increase in earnings generated by the stepped-up production would increase consumption demands again, further increasing production and setting off a chain reaction of income increases. However, the authors do not emphasize this result, for, if the economy really were in a great enough recession to allow unimpeded increases in employment, virtually any government expenditure would have the same effect. The NIT is certainly not unique in this regard.

The uniqueness of the NIT is in its redistributive potential, and it is here that the more interesting finding turns up. Golladay and Haveman find that the secondary impact of the NIT--the impact on the earnings of those workers in industries producing commodities for which the NIT increased demand--works in an opposite direction to the initial impact. Thus, although the initial impact of the NIT is to increase the incomes of low-income families, the secondary impact tends to increase the earnings of higher-income

families, thus partially offsetting the initial redistributive impact of the NIT. In short, low-income consumers tend to buy goods produced by highly skilled, high-wage workers (in the auto, steel, petroleum, and education industries), while high-income consumers (whose consumption will drop) tend to buy relatively more goods produced by unskilled, low-wage workers (in the textile, apparel, hotel, and services industries).

In a more recent paper, Haveman et al.⁹ have used their model to simulate the effects of the Better Jobs and Income Program. They find qualitatively similar results, as one would expect: an increase in aggregate output (four billion dollars in 1975 dollars) and secondary distributional effects offsetting these effects.

CONCLUSION AND SUGGESTIONS FOR FUTURE RESEARCH

The discussion in this short paper has illustrated some of the ways in which an NIT would affect the national economy. The most unexpected result of the first research effort discussed was that in times of high unemployment, an NIT would have little impact on wage inflation or aggregate production (e.g., GNP). The Haveman-Golladay research produced the startling discovery that secondary income-distributional impacts of an NIT work in the opposite direction of the initial effects.

Many other potential macroeconomic effects need study. The two research efforts described here raise more questions than answers. With regard to the impact of an NIT on the labor market, there remain several important questions regarding (1) the impact on job search in the labor market and thus on voluntary unemployment, (2) the impact on different submarkets

rather than the aggregate labor market (sectoral Phillips curves would be needed for this analysis), (3) the degree to which firms would substitute regional and/or high-wage labor for low-wage labor as wages for the latter rise, and (4) the degree to which firms can pass through wage increases to the consumer. With regard to expenditure impacts, there remain important questions regarding (1) the redistributive impact of an NIT with different levels of unemployment and with full employment, (2) the degree to which wages would change, inducing the factor substitutions by firms mentioned above, (3) the impact on commodity price inflation, and (4) the effect of work-effort changes induced by the NIT.

NOTES

1. R. Moffitt, "The Short-Run Labor-Market Replacement Effect of a Negative Income Tax," (Working Paper C-22, Mathematica Policy Research, Princeton, 1977).
2. R. Haveman, K. Hollenbeck, D. Betson and M. Holmer, "The Poverty Institute Regional and Distributional Model: Its Application to the Program for Better Jobs and Income," (Paper presented at the Conference on Microeconomic Simulation Models for the Analysis of Public Policy, Washington, D.C., March 1978); F. Golladay and R. Haveman, "Regional and Distributional Effects of a Negative Income Tax," American Economic Review 66 (September 1976): 629-641.
3. Technically, marginal workers contribute less and are now unprofitable.
4. If the labor-supply curve is backward-bending, i.e., if higher wages call forth fewer workers, the same conclusions apply, as long as employers' desired employment falls at a faster rate than labor supply.
5. R. Lipsey, "The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1867-1957: A Further Analysis," Economica, N.S. 27 (February 1960): 1-31; A. Phillips, "The Relation Between Unemployment and the Rate of Change of Money Wages in the United Kingdom, 1867-1957," Economica, N.S. 25 (November 1958): 283-299.
6. See note 1, above.
7. This estimate is for husband-wife families facing an NIT with a 50% tax rate and a guarantee at the poverty line. It is similar to that of the MPR-SRI simulation except for the assumption of a 50% participation rate in the eligible population.
8. F. Golladay and R. Haveman, note 2, above.
9. R. Haveman, K. Hollenbeck, D. Betson and M. Holmer, note 2, above.

DISCUSSION

JOHN McCOY: In your estimates of the impact on inflation, did you take into account the fact that as inflation increases the cost of living, there would be an automatic escalation in the NIT payment?

MOFFITT: No, I didn't.

McCOY: You assumed that the NIT would stay fixed?

MOFFITT: Yes, I did -- fixed in real terms, assuming that people would react to the real adjusted amount. I think if the benefits were not adjusted, then the real monetary equivalent of those dollars would increase, and that may cause some adjustment.

McCOY: But if it is adjusted periodically so that it maintains real dollars, would that have a compounding effect on inflation overall?

MOFFITT: No, it wouldn't -- not in the simplest case -- unless people were failing to realize what was going on. If they were reacting in unexpected ways to the adjustment, then it might. But if they were adjusting for inflation so that the real value stayed the same, it wouldn't compound effects.

PHILIP ROBINS: How did you calculate the rate of wage inflation? Did you use a Phillip's curve?

MOFFITT: Yes; this is a curve which describes the trade-off between unemployment and inflation. Generally, it has a negative slope, which means that the lower the rate of unemployment, the higher the rate of inflation. As the labor market tightens up, it's harder to find workers, the unemployment rate is very low, employers have to increase wages to draw workers, and inflation tends to increase. The negative relationship has very clear important policy implications. That is, if you try to reduce the unemployment rate in the labor market -- by whatever public policy measure you choose -- then that's going to have an effect on the economy by increasing the wage inflation rate.

JODIE ALLEN: You show a big drop in the labor force.

ROBINS: That is a very curious result, because you have an economy that is very slack to begin with. You introduce a program which calls for people to work even less, and you end up with more inflation. The critical element here is how the labor force changes.

MOFFITT: The labor-force reduction is simply based upon the MPR-SRI simulations. Yes, the economy is slack to begin with, but fewer people in the labor force means fewer people banging on employers' doors for jobs, and that causes some increase in wage inflation.

ALLEN: Did you compare these results with David Greenberg's first version of his model? He did try to simulate some sort of wage response.

MOFFITT: Yes. He did something even simpler than this. He didn't actually estimate the effect on the inflation rate, but he assumed an "equilibrium" level of wages within a simple supply and demand framework.

ALLEN: Did it turn out anything near the same?

MYLES MAXFIELD: I think generally he found that the effects were large.

MOFFITT: That's because he calculated the long-run effect. Basically, what I'm presenting here is the short-run effect, which should be smaller.

ROBINS: Suppose the labor force consists of ten people, five of which are unemployed, and the reduction in labor-supply is the unemployed person dropping out of the labor force. Then, does unemployment rate go down instead of up?

ALLEN: The unemployment rate went down. Unemployment decreases because a lot of people dropped out of the labor force.

MOFFITT: And that makes the labor market tighter. The other important result is that the decrease in the employment level is also fairly small in a high unemployment economy. The GNP is not going to fall, but the economy is not going to recover as rapidly as it would have in the absence of a welfare program.

GLADYS McCORKHILL: Does that take into account the NIT-induced increase in demand for goods and services?

MOFFITT: No, it doesn't: that could have some very stimulating effects.

McCOY: Do you have any plans to study the combined impact?

MOFFITT: Yes, we do. We hope to get a larger picture by combining these assumptions.

A SIMULATION OF THE PROGRAM FOR BETTER JOBS AND INCOME

BY

David Greenberg, David Betson,
Richard Kasten

(Staff Economists, Office of Income Security Policy, Office
of the Assistant Secretary for Planning and Evaluation,
Department of Health, Education, and Welfare)

On August 6, 1977, President Carter announced a proposal for welfare reform--the Program for Better Jobs and Income (PBJI). This paper describes in general terms the model developed to analyze welfare programs of this type and, then, the use of the model to simulate the national response to the proposal. The model, which has become known as the KGB model, after the initials of the author's names, was developed in the Office of Income Security Policy, within the Department of Health, Education, and Welfare.¹ It can be used to simulate welfare reform proposals with varying support levels, tax rates, including positive tax rates, and job guarantees for certain subsets of the population.

Prior to the development of the KGB model, the available models in existence were unable to simulate a reform package that combined a cash component with a jobs component. Because information about the costs and effects of various combinations of cash and jobs were urgently needed

so that major decisions about the Administration's welfare reform plan could be made, thus, the model was developed.

STRUCTURE OF THE MODEL

The model is exceedingly complex, and contains over 8,000 individual steps. Speaking generally, however, the simulation of proposed alternatives to the existing tax and transfer system can be grouped into four stages. First, the economic characteristics of a representative sample of the nation's families is developed. Some characteristics, such as hours, earnings, and unearned income, are obtained directly from data in the Survey of Income and Education (SIE) by the Census Bureau. Others, such as taxes, tax rates, transfer income, benefit reduction, and so forth are calculated from the income and other data, using current schedules or rules of eligibility. Adjustments are made for inflation from the date of the SIE survey to 1978.

Second, the values of net wage rates (wage rate less tax rate) and disposable income are adjusted to what they would be were the simulated reform measure implemented, but work effort and earnings remained unchanged. For workers who are eligible for a public service employment (PSE) job, it is necessary to compute what the values of the variables would be if: (1) the worker leaves the conventional labor market to take a public employment job, (2) he remains attached to the conventional job sector and takes a public employment job only when he is unemployed, or (3) he does not participate in public employment at all.

The third step consists of accounting for labor supply responses to changes in wage rates and disposable income under each public employment strategy. First, the values obtained from the first two steps are used to calculate the changes in net wage rates and disposable incomes. Predictions of the effects on work hours of the reform are then derived by multiplying these calculated changes by appropriate labor supply parameters that were estimated from the Seattle-Denver income maintenance experiment (SIME/DIME). These labor supply adjustments are then used to determine the number of hours individuals would work during the post-reform period. Then household earnings, transfer payment receipts, and tax payments are all recomputed.

The final step is the simulation comparison of the value to the worker under all three assumptions, to determine whether an individual who is eligible for public employment will take such a job, and whether he or she will do so even when in the labor force, only when he is unemployed, or not at all. This third step is quite unique, and rests upon an assumption that the individual will follow an economically rational course and seek to maximize disposable income. We also assume that the number of hours worked remains the same, or, if a legal maximum is imposed in a PSE job, that hours do not exceed the maximum. Then, along with comparisons of the computed work and transfer incomes available, we determined the costs associated with a change of status--such as the relinquishment of unemployment compensation, and waiting periods required prior to entry into a PSE job.

This methodology assumes that this choice is made on the basis of perfect information about the alternatives, an assumption that probably results in an overstatement of the population likely to seek PSE jobs. Unless persons are aware of public employment and know something about it, the program is not a viable alternative for them. Knowledge of the program is likely to be positively related to the publicity given the program, the length of time the program has been in operation, and the size of the program. But, assuming perfect knowledge, individuals are assigned to the choices that make them better off.

SIMULATION OF THE CARTER ADMINISTRATION'S WELFARE REFORM PROGRAM

Major Features of PBJI²

The Carter Administrations PBJI welfare reform proposal is designed to provide low-income people with access to jobs and to a consolidated cash assistance program that replaces the federal share of Aid to Families with Dependent Children (AFDC), Social Security Income (SSI) and food stamps. To be eligible, adults who were expected to work would have to register with an employment and training agency that would seek to locate private sector employment for them. But, if a private sector job could not be found the program would attempt to place the principal earner in families with children in PSE. These would pay at least the minimum wage and would be full-time for all recipients except for those single parents with young children who preferred a part-time PSE job.

Under the cash assistance component of PBJI, the federal government would provide a basic minimum benefit to all those who were eligible and encourage the states to supplement this basic benefit. (For example, the federal government would pay 75% of a state supplementary payment of the first 12.32% above the federal support and then 25% up to the poverty line.)

A separate benefit schedule would apply to families containing a member who was expected to work and recipients who do not. Those not expected to work include the aged, the blind, the disabled, and single-parent families in which the youngest child is under 7. In 1978 dollars the basic federal cash benefit paid to a single parent with one child with no other income is set at \$3000; for each additional child this basic benefit would be raised by \$600. An aged, blind or disabled individual with no other income would receive a basic benefit of \$2500, while an aged, blind or disabled couple would receive \$3700. These basic benefits would be reduced by 50¢ for each dollar of earnings in states where supplements are not received and by not more than 70¢ in states that supplement federal benefit levels.

Families with a member expected to work include two-parent families with children, single-parent families with children age 7 or older (when the children are 7 through 13 only part-time work is expected), childless couples and individuals. The basic benefit for the spouse and each child in two-parent families would be equal to \$1100 and \$600 respectively,

in 1978 dollars, while the adult expected to work would receive nothing during an initial eight-week period of job search and nothing thereafter if he or she refused work. If a regular job could not be found, the government would attempt to create a public service job. The worker in such a family could keep the basic benefit while earning the first \$3800. After that, cash payments would be reduced by 50¢ for each dollar of earnings under the federal program and by not more than 52¢ in states providing supplementary benefits. If after eight weeks no job could be found or created, the family would receive the same benefit as a family of four with no member expected to work.

Single-parent families with all children over 14 would be treated much the same, except that a parent of a child aged 7 through 13 would only be required to work part-time, and would receive full benefits during the eight-week search period, although benefits would be reduced after the search period if the parent refused a job that paid the minimum wage or better. Benefits would be reduced by 50¢ per dollar earned under the federal program and by not more than 70¢ in states that provide supplements.

Public service jobs would not be provided to childless couples and single individuals. These persons would be eligible for a basic benefit of \$1100 per adult, paid during the job search period and until a job is found. Benefits terminate, however, if an individual refuses a job

paying at least the minimum wage. A reduction of benefits by 50¢ for each dollar of earnings would begin with the first dollar of earnings.

In addition to the jobs and cash components just described, the PBJI would also amend the existing Earned Income Tax Credit (EITC). At present, low income families with children may receive a refundable tax credit of 16¢ for each dollar of earnings up to \$4000. The benefit falls by 10¢ for each dollar of gross income beyond \$4000. Thus, the benefit is zero for those with no earnings, reaches a maximum of \$400 for those who earn \$4000, and falls to zero when income equals \$8000. Under the PBJI, families would receive 10% of their first \$4000 of private sector earnings plus 5% of any private sector earnings between \$4000 and 216% above their federal basic benefit level. The tax credit would be reduced by 10¢ for each dollar of income above this amount until it phases out. To make private sector employment relatively more attractive, the tax credit would not apply to earnings from PSE.

To avoid the cumulation of tax rates due to any overlap between the cash assistance system and the federal income tax, a scheme to reimburse taxes is included in the proposal. PBJI would provide an additional payment to families who are eligible for cash assistance, but who are also required to pay federal income taxes. This payment would be equal to 20% of earned income in excess of the point at which earnings became subject to taxation, but less than 216% of the unit's federal basic benefit level. Income in excess of this latter point would be used to phase out the reimbursement at a rate of 20%.

Specific Assumptions Underlying the Simulation

The simulation results that are reported below state the effects of PBJI for 1975, assuming the program had been operative in that year, with all behavioral adjustments to the program made. We have assumed that (1) demographic composition and economic conditions for 1975 are accurately represented by the SIE, in particular, an 8.5% unemployment rate for the overall economy; (2) the 1975 tax and transfer system (with adjustments for later amendments contained in the Food and Agriculture Act of 1977) represents pre-reform conditions facing individuals; (3) states would supplement up to the state's 1975 AFDC benefit levels plus the bonus value of food stamps for families with children; their 1975 General Assistance amounts for single individuals and childless couples; and their 1975 SSI benefit levels plus the bonus value of Food Stamps for Aged, Blind and Disabled filing units (except the supplement would not exceed the amount matched by the federal government); (4) states would not grandfather current AFDC, SSI or GA recipients; (5) the work test would function perfectly for families with children; (6) sufficient public employment jobs would be created so that all eligible workers who want one can obtain one; (7) the program would use an annual accounting period; (8) a self-employed worker with more than a grade school education would choose a PSE; and (9) single adults who live with a relative will file as part of the relative's family if it maximizes total transfer benefits.

Results of the Simulation

Simulation results fall into three areas: First, eligibility for and participation in the cash component of the program; second, the

characteristics of the participants in the jobs component; and third, effects of PBJI on disposable income and work effort.

We found that had the PBJI existed in 1975, 34.3 million people or 16% of the total U. S. population would have been eligible for cash payments totalling \$21.5 billion. However, experience with welfare systems indicate that more persons would be entitled to benefits than would participate in the program. Simulations with the KGB model attempt to account for this and Table 1 gives separate estimates of the population that is eligible and that portion predicted to participate. Thus 81% of those eligible, or 27.8 million people would have received a cash payment under PBJI in 1975. This level of participation reduces the annual entitlement by 13%, to \$18.7 billion. Participation among different recipient groups varies considerably, ranging from 71% for childless couples (who are not aged, blind or disabled) to 89% for families with children where parents are not expected to work. Benefits range from 79% to 94% of actual entitlements, reflecting the expectation that families become less likely to participate in the programs as the size of the potential benefit diminishes.

The second main inquiry focuses on characteristics of persons who would enter PSE jobs. Table 2 shows the economic and demographic groups likely to participate in the jobs component of PBJI. The simulation indicates that a total of 2.8 million individuals would have participated in the PBJI jobs program sometime during 1975, that they would require 1.29 million slots, and that their average length of stay in the program during the year would have been 24 weeks.

TABLE 1

ELIGIBILITY FOR AND PARTICIPATION IN THE CASH ASSISTANCE COMPONENTS
OF THE PROGRAM FOR BETTER JOBS AND INCOME

	U.S. POPULATION		ELIGIBLE FOR CASH ASSISTANCE			PARTICIPANTS IN CASH ASSISTANCE		
	PERSONS (millions)	FILING UNITS (millions)	PERSONS (millions)	FILING UNITS (millions)	CASH ENTITLE- MENTS (\$billions)	PERSONS (millions)	FILING UNITS (millions)	CASH PAYMENTS (\$billions)
EXPECTED TO WORK FAMILIES WITH CHILDREN	103.3	25.1	12.3	2.6	3.7	9.5	2.0	3.0
SINGLE INDIVIDUALS (NON-ABD)	26.7	26.7	2.0	2.0	1.4	1.4	1.4	1.1
CHILDLESS COUPLES (NON-ABD)	33.3	16.7	1.3	.7	.6	.9	.5	.5
NOT EXPECTED TO WORK FAMILIES WITH CHILDREN	20.5	6.7	12.7	3.8	10.0	11.4	3.4	9.4
SINGLE INDIVIDUALS (ABD)	11.7	11.7	4.4	4.4	4.8	3.4	3.4	4.0
CHILDLESS COUPLES (ABD)	15.3	7.7	1.6	.8	1.0	1.2	.6	.8
TOTAL	210.9	94.1	34.3	14.1	21.5	27.8	11.2	18.7

Note: Estimates pertain to persons receiving benefits at any point during the year, and assume an annual accounting period.

TABLE 2

 CHARACTERISTICS OF THE PUBLIC SERVICE EMPLOYMENT PARTICIPANTS
 (in 1000's)

	PURE			MIXED			TOTAL		
	PERSONS	SLOTS	AVERAGE WEEKS	PERSONS	SLOTS	AVERAGE WEEKS	PERSONS	SLOTS	AVERAGE WEEKS
TOTAL	1117	669	31	1679	623	19	2796	1292	24
SEX:									
MALE	317	226	31	988	377	20	1305	603	24
FEMALE	800	443	29	691	246	19	1491	689	24
EDUCATION:									
0 - 8	269	153	30	328	123	20	597	276	24
9 - 11	324	189	30	490	193	21	814	382	24
12 AND OVER	524	327	33	861	347	21	1385	647	24
AGE:									
18 - 25	305	176	30	462	160	18	767	336	23
26 - 35	281	162	30	597	228	20	878	390	23
36 - 45	249	155	32	401	157	20	650	312	25
46 - 55	181	116	33	176	62	18	357	178	26
55 AND OVER	100	60	31	44	16	19	144	76	27
PRIVATE SECTOR WAGE RATE:									
LESS THAN \$2.10	1106	664	31	204	68	17	1310	732	29
\$2.11 - \$3.00	8	3	20	410	143	18	418	146	18
\$3.01 - \$5.00	3	2	33	591	214	19	594	216	19
OVER \$5.00	0	0	0	473	196	22	473	196	22
REGION:									
NORTH EAST	213	123	30	384	171	23	597	294	26
NORTH CENTRAL	253	147	30	444	165	19	697	312	23
SOUTH	485	294	32	533	180	18	1018	474	24
WEST	165	105	33	318	107	18	484	212	23

The KGB model distinguishes between public employment participants who follow a "pure strategy" (seeking only PSE jobs) and those who follow a "mixed strategy" (seeking PSE jobs only when unemployed). The simulation results indicate that 60% of the participants in the public employment component of PBJI would be workers who would follow a "mixed strategy", but that these persons would occupy only 48% of the total slots and would remain in the program for 19 weeks. While 19 weeks is considerably higher than the average duration of unemployment in 1975 (14.1 weeks), 30.7% of the unemployed remained so for more than 15 weeks³ and are more likely to seek a PSE job. Other information on Table 2 is quite interesting. One provision of the jobs program to which there has been considerable objection, especially among feminist groups is the rule limiting these jobs to primary earners. Even though it is expected that this screens out many wives, our projections indicate that over half the PSE participants would be women. PSE participants would apparently be about equally divided between high school graduates and non-graduates. A majority would be under 35 years of age. (However, there appears to be a modest increase in weeks of stay on the program during the year as age increases.) Almost all "pure strategy" participants have a market wage rate that is below the 1975 minimum wage of \$2.10.⁴ "Mixed strategy" participants, on the other hand, tend to be drawn from relatively higher wage groups. Finally, the southern part of the country, where there are a large number of people working at less than a minimum wage, would have more participants--36% of the participants are located here.

The third area of inquiry focuses on various disposable income and work effort. Table 3 indicates the net increase in total disposable income would be almost \$4 billion. This increase results from some fairly dramatic changes in certain components of income, particularly earnings. Earnings from the PSE jobs would equal almost \$6 billion, and are partially offset by a \$1.5 billion decrease in private sector earnings due mainly to a shift into PSE jobs and partly to a shift to cash assistance: net gain in earned income would be \$4.49 billion.

Direct cash assistance payments (including food stamp benefits) to low-income families would decrease by \$2.25 billion. This is due, in part, to changes in eligibility requirements and allowable work-related deductions; it is also partly attributable to the increase in income of low-income families derived from PSE jobs. The decrease in cash assistance payments would be partly offset by an expansion of the earned income tax credit. However, the \$2.8 billion increase in tax credits would not only be paid to cash assistance recipients, but also to those with considerably higher income.

The net increase in earnings that would result from PSE would also produce certain additional effects. For example, unemployment benefits would be reduced by \$570 million, and state and federal tax revenues would expand by a total of \$500 million.

To sum it up, the PBJI would result in an increase in disposable income of \$3.97 billion, and would incur a net cost to state and federal governments of \$5.47 billion--an amount equal to the change in disposable income

TABLE 3

DIRECT EFFECTS OF THE PROGRAM FOR BETTER JOBS AND INCOME ON DISPOSABLE INCOME
(in \$ Billions)

	PRE-REFORM	POST-REFORM	CHANGE
PRIVATE SECTOR EARNINGS	861.13	859.63	- 1.50
PUBLIC SECTOR EARNINGS	0.0	5.99	+ 5.99
OTHER INCOME*	224.70	224.70	0.0
CASH ASSISTANCE	20.98	18.73	- 2.25
AFDC (FEDERAL & STATE PAYMENTS)	9.04	0.0	
SSI (FEDERAL & STATE PAYMENTS)	5.41	0.0	
GENERAL ASSISTANCE	1.38	0.0	
BONUS VALUE OF FOOD STAMPS	5.15	0.0	
FEDERAL CASH BENEFIT AND TAX REIMBURSEMENT PAYMENT	0.0	15.01	
STATE SUPPLEMENT PAYMENT	0.0	3.95	
UNEMPLOYMENT COMPENSATION PAYMENTS	13.30	12.73	- .57
EARNED INCOME TAX CREDIT	1.15	3.95	+ 2.80
TAXES	- 198.23	- 198.73	- .50
FEDERAL INCOME TAXES	- 128.89	- 129.06	
STATE INCOME TAXES	- 26.04	- 26.10	
SOCIAL SECURITY PAYROLL TAXES (EMPLOYEES SHARE)	- 43.30	- 43.57	
TOTAL DISPOSABLE INCOME	923.03	927.06	+ 3.97

*Other income includes Social Security benefits, Railroad Retirement, interest, dividends, government and private pensions, Veteran's Pensions, Workman's Compensation, alimony and child support payments.

Note: Financing the program would, of course, also affect disposable incomes. However, these considerations are ignored for purposes of the presentation in this section.

plus the \$1.5 billion reduction in private sector earnings.

Obviously, these aggregate amounts do not reveal the extent to which the PBJI shifts resources among population strata. For example, the net decrease of \$1.5 billion in private sector earnings is attributable to a \$1.43 billion reduction in the private sector earnings of workers who leave conventional jobs to work only in the public jobs, and a \$70 million reduction in the earnings of workers who would not enter public employment, but would reduce their private sector hours in response to changes in work incentives.

Adjustments in hours and private sector earnings are reflected in Table 4. Table 4 suggests that the PBJI would cause an increase in the hours and earnings of non-PSE participants who currently receive cash benefits, but would result in a decrease for those who do not. This implies, in turn, that current recipients among non-PSE participants would be eligible for fewer cash benefits under PBJI than they presently are, and as a consequence would increase their work effort. While the opposite is true for those who would become recipients of cash benefits for the first time under PBJI.

Persons who take PSE jobs would increase their overall hours of work. The increased earnings that results from these additional hours have a diminishing effect on the amount of cash benefits that would be received by these groups. The reduction in the private sector earnings of "pure" PSE participants is a consequence of substitution of public for private work by these persons. "Mixed" participants would apparently change their private sector hours and earnings by very little, but their total earnings would increase substantially as they took advantage of opportunities to work at PSE jobs during periods of unemployment.

TABLE 4

PROGRAM EFFECTS ON THE DISPOSABLE INCOMES OF THREE CATEGORIES OF RECIPIENTS

CHANGE IN :
(in \$Billions)

	NUMBER OF PERSONS	DISPOSABLE INCOME	PRIVATE EARNINGS	PUBLIC EARNINGS	CASH PAYMENTS	E.I.T.C.
CURRENT RECIPIENTS WHO WOULD CONTINUE TO RECEIVE CASH AFTER THE REFORM:						
NON-PARTICIPANTS IN PSE	5.3	.14	.15	-----	- .21	.21
PURE PARTICIPANTS	1.6	.34	- .50	1.28	- .32	- .04
MIXED PARTICIPANTS	2.0	.71	.05	1.53	- .75	.04
TOTAL	8.9	1.19	- .30	2.81	-1.28	.21
CURRENT RECIPIENTS WHO WOULD NOT RECEIVE CASH BENEFITS AFTER THE REFORM:						
NON-PARTICIPANTS IN PSE	16.5	-1.95	1.05	-----	-2.97	.20
PURE PARTICIPANTS	1.8	.00	- .19	.42	- .17	- .02
MIXED PARTICIPANTS	3.3	.10	.03	.46	- .32	.04
TOTAL	21.6	-1.85	.89	.88	-3.46	.22
FAMILIES WHO DID NOT RECEIVE CASH BENEFITS BEFORE THE REFORM:						
NON-PARTICIPANTS IN PSE	9.1	3.44	-1.3.	-----	2.16	2.39
PURE PARTICIPANTS	.5	.56	- .74	1.39	.16	- .05
MIXED PARTICIPANTS	1.1	.63	- .04	.91	.16	.03
TOTAL	10.7	4.63	-2.09	2.30	2.48	2.37

Note: "current recipient" refers to families receiving AFDC, SSSI, general assistance and/or food stamps.

"Cash benefits after reform" refers to both cash assistance or an earned income tax credit.

NOTES

1. A more detailed and technical discussion of the model and its applications and an analysis of its limitations may be found in David Betson, David **Greenberg**, and Richard Kasten, "A Micro-Simulation Model for Analyzing Alternative Welfare Reform Proposals: An Application to the Program for Better Jobs and Income," Robert Haveman and Kevin Hollenbeck eds., Microeconomic Simulation Models for the Analysis of Public Policy, (Institute for Research on Poverty, Madison, Wisconsin) forthcoming.
2. A fuller description of the program is contained in Department of Health, Education, and Welfare, "Better Jobs and Income Act, H.R. 9030: A Summary and Sectional Explanation", September 13, 1977. Simulations reported on in this paper are limited to the PBJI, as originally proposed to Congress. No account is taken of any congressional changes to that program.
3. Handbook of Labor Statistics, U. S. Bureau of Labor Statistics, Department of Labor, 1977.
4. The 11,000 pure participants who have wage rates that exceed the minimum can be accounted for by two factors. The first is that some States are required to supplement the PSE wage by as much as 10%. The second factor is that certain individuals with very high levels of expected future unemployment will prefer the pure strategy to the mixed, since under the mixed strategy they are forced to incur the public employment waiting period each time they become unemployed.

DISCUSSION

ROBERT HOREL: There is a possible problem in your data base. In California, we decided the SIE would be inadequate, and used Social Security Administration files for the current SSI people and used that as representation of SSI under welfare reform. We also had to go to our own state files for AFDC to really simulate that part.

GREENBERG: Yes. I wouldn't argue with you. You're right. We actually started out using the CPS, Current Population Survey. Then the SIE became available. The basic reasons we used it is because of the richness of the information, although it does not have everything that you need, and it has to be augmented in various ways for use in the simulation model. The other reason we used it is because it was about the largest data file available that provided the information that we needed, and it supposedly has statistically significant information for each individual state so that you can do predictions for individual states. I would certainly admit that the information for individual states is somewhat soft. I just simply think that it provides better inferences than one can provide elsewhere.

DICK RITCHIE: We do the same thing in Washington State, and we ended up using the '77 AFDC characteristic study for the current population and augmenting that with the SIE. Another problem we had with the SIE data was in converting annual figures and to an April '76 situation. We really couldn't use the annual income figures as such, because it was peculiar to the BJ and I Program. You have to know the turnover. Using annual figures really distorts the situation, especially where you have a family breakup during the course of that year.

GREENBERG: That's one of the points I was going to get to, particularly in connection with the accounting periods. In the model we assume an annual accounting period, but Congress is talking about shorter accounting periods. That will mean that a lot of people whose annual income is above program breakeven levels, but whose income for a short period of time is below program breakeven levels will be eligible to participate. We're trying to modify the SIME/DIME data so that we can run it through our model. There are problems with that. For instance, the SIME/DIME data is not representative of the nation, but, at least, we can simulate different accounting period lengths and get some grasp of the accounting-period issue.

ROBERT HOREL: You're working with 1975 data and we're going to be implementing a program in '80 or '81. There are perhaps two ways you can address that problem: to try to simulate the model as though the program were in place in 1975, and then get costs and then try to update those costs to '80 - '81; or to try to go in and update the data in the model to '80 - '81, and run the program on that. Do you want to address the relative value of those approaches?

GREENBERG: I think the latter one is the better one, and that's basically what's being done. As I said, the results I'm going to report on are for 1975, and in many ways I feel more comfortable about that; but for policy purposes we attempt to predict what would actually happen in the year of implementation. That's basically done by updating an existing data-set like the SIE.

MICHELLE BELL: What wage rate was used for the public sector's employment; were you using the minimum wage rates?

GREENBERG: We attempted to use the program wage rates; and in fact, there have been suggestions about higher wage rates and so forth through the program, so we've tried alternative wage rates. But the tricky part is that you have to deflate wage rates somehow to 1975. I think what we ended up using was the actual minimum wage in 1975, which was, I think, \$2.10 an hour. We have also generally allowed for higher wages for people in supervisory positions (10% of the PSE jobs).

STEVEN DIRECTOR: Referring to Table 3, if public sector earnings go up more than private sector earnings go down, does this imply a net increase in work effort?

GREENBERG: Yes. I think that interpretation's correct.

STEVEN DIRECTOR: That would be a very desirable result if I'm interpreting it correctly. You are laying out more assistance and getting a net increase in work effort under this program.

GREENBERG: That's the idea. Basically, we are filling in a lot of time that people would have been unemployed. But your interpretation is not quite correct because the earnings increase for a lot of people simply reflects a higher wage. But I think your general point's correct.

V

MAJOR IMPACTS ON THE FAMILY

INCOME MAINTENANCE AND MARRIAGE

by

Michael T. Hannan
Fellow, Center for Advanced Study
in the Behavioral Sciences

Nancy Brandon Tuma
Assistant Professor, Stanford University

and

Lyle P. Groeneveld
Sociologist, SRI International

Proponents of welfare reform argue that the current welfare system encourages marital dissolutions by prohibiting or limiting assistance when the father of the family remains in the household. In his August 6, 1977 message to Congress, President Carter cited an example of the anti-marriage incentives in the existing system: "in Michigan a two-parent family with the father working at the minimum wage has a total income, including tax credits and food stamps, of \$5,922. But if the father leaves, the family will be eligible for benefits totalling \$7,076." Similar examples of marriage disrupting incentives may be found for other states.

However, there is little empirical evidence to support this view. The several studies that have addressed this issue have failed to demonstrate a causal relationship from these marriage-disrupting incentives to marital dissolution. In contrast, the NIT experiments were designed to test the effects of an NIT on marital stability. In this paper we summarize our findings from the first two years of the Seattle and Denver Income Maintenance Experiments (SIME/DIME).

THEORETICAL FRAMEWORK

There are many reasons for suspecting that an NIT would affect the rate of marital breakups. First, an NIT removes financial incentives to marital dissolution inherent in the present system. Aid to Families With Dependent Children (AFDC) prohibits benefits to families where the father is present. The Aid to Families With Dependent Children/Unemployed Parent (AFDC-UP) program eases this restriction somewhat but only where the father experiences a certain amount of unemployment in a given period. The number of days employed (rather than the adequacy of the family income) determines eligibility. As an NIT determines benefits by family income, it should remove this incentive towards dissolution.

Second, since the lowest income families have the highest marital dissolution rates, an NIT may lower dissolution rates by raising family income. This will not be true if the relationship between low income and marital instability reflects cultural differences between the poor and the rest of society. That is, if the poor have high rates of marital dissolution because they lack appropriate values and personality traits, raising income levels will not improve marital stability.

But marital stability could be affected by an NIT, if the changes in income affect the strains on marriages and the ability of families to cope with a variety of problems and dissatisfactions. An NIT may also affect marital stability through changes in consumption activities. Personal and social worth in our society are evaluated primarily through consumption activities; heads of families who cannot provide certain consumption standards for their families are viewed as failures by themselves and others. One

response to such failure is flight from the marriage. Income supplement programs that substantially improve living standards might reduce the pressures toward dissolution. We refer to effects through changes in family income as income effects. We expect that the income effects of an NIT would lower the rate of marital dissolution.

But there is another effect of an NIT that is often overlooked in policy discussions of welfare and marriage. An NIT guarantees support to unmarried as well as married people. As a result, an NIT will alter the level of resources available outside of marriage and thereby alter the dependence of the members on marriage. We refer to this as the independence effect. If the NIT increases the level of resources outside of marriage, the independence effect will raise the probability of marital dissolution.

There are also non-pecuniary differences between welfare and an NIT that might affect marriage. For example, it is often argued that the current welfare system stigmatizes recipients and that an NIT would reduce stigma. If participation in the current system is degrading, both its income and independence effects are muted. Families receiving payments would not experience the full income effect due to the strain induced by stigma. Likewise, dependent spouses would not experience the full independence effect of the welfare system. In particular, women who believe that receiving welfare is degrading may choose to remain in unsatisfying marriages rather than go on welfare. This suggests that a payment from an NIT program will have stronger income and independence effects than a payment of the same amount from welfare, i.e. welfare is "discounted" in its effects on marriage relative to an NIT.

Another nonpecuniary feature of welfare that may result in its being discounted is the effort needed to establish and maintain eligibility. Our experimental NIT program has a simpler and presumably less alienating bureaucracy that requires minimal effort from participants. Also, the rules of the NIT are carefully explained to the participants. Welfare recipients probably do not have the same knowledge of the eligibility rules and support levels of existing programs. Any of these--stigma, effort, or lack of information--suggest that the effects of welfare may be discounted.

What, then, can be said about the expected impact of an NIT on marital dissolution rates? For an NIT that is more generous than the present welfare system, as is the case with SIME/DIME, we cannot predict the direction of its impact a priori. If the income effects dominate, the NIT will lower the dissolution rate. If the independence effects are stronger, the reverse will hold. Even a less generous program may have both income and independence effects if the changes in the program affect the rate at which welfare is discounted.

THE SAMPLE OF MARRIED COUPLES

The general development and structuring of the sample and the data collection are described above in Part II of these papers. Of particular interest to the present paper is the definition of marriage presented there. We did not require that couples be legally married. For families with children, we considered two persons to be married if they lived together, pooled resources and claimed the relationship was permanent. For couples who did not have children we required an affidavit stating that the relationship was

a legal union or a common-law marriage. With this definition of marriage we obtained mostly long-term unions. Between 85% and 90% of the marriages at the beginning of the experiment had existed for more than two years.

Also important is our definition of "dissolution." We treat the marriage as having ended or dissolved when one partner moves out of the household and the members report that the split is permanent. So the outcomes we study are not legal divorces; they are the endings of marital unions.

There were 2771 experimental and control families who met this definition of "married." The average age of husbands at enrollment was 33.5 years, and wives, 31.3 years. Husbands averaged 10.9 years and the wife 11.0 years of schooling. On the average they had been married for 9.4 years, had 2.3 children and \$6,208 annual income. Most (75%) of the couples had one or more children under ten. About half the wives were employed; 16% of families had been on AFDC in the prior year. There were 1302 white couples, 942 black couples and 527 Chicano couples. The distribution among manpower treatments was counseling only, 19%; counseling with a 50% subsidy, 24%; counseling with a 100% subsidy, 14%; and manpower controls, 42%. The proportion receiving financial treatment at the \$3,800 support level was 18%; at \$4,800 support level, 23%; at \$5,600 support level, 14%; 18% of the financials were enrolled for five years. Financial controls were 45% of the 2771 families. Over two years of the study, 9% of the families dropped out of the experiment.

BASIC EXPERIMENTAL FINDINGS

Dissolutions

Overall, the NIT plans destabilized marriages for whites and blacks, and had little effect on Chicanos. The following gives the percentage

increase in dissolution rates due to the experiment. Controlling for the variables that may affect dissolution¹ we found that the experiment increased dissolution rates 61% for blacks and 58% for whites. There was a 4% decrease for Chicanos. For blacks and whites, this difference cannot be attributed to chance; it is statistically significant at the 5% level. These percentages are for families assigned to the five-year treatment. The effects for three-year families are approximately 80% of the five-year effects.

This finding is consistent with a model in which the independence effects dominate the income effects for the programs tested. But does it imply that all NIT schemes will increase dissolution rates in populations like those we studied? To answer this question we must consider some more complex analyses. Our most provocative findings concern the pattern of impacts by level of income support. The impacts for the three support levels are presented in Table 1.

TABLE 1
PERCENT CHANGE IN MARITAL DISSOLUTION RATE
BY LEVEL OF INCOME GUARANTEE

<u>Guarantee Level</u>	<u>Race-Ethnic Group</u>		
	<u>Black</u>	<u>White</u>	<u>Chicano</u>
90% of poverty line	67%**	96%***	60%
125% of poverty line	93%***	55%**	-28%
140% of poverty line	21%	12%	-.35%
Total Families	939	1297	518

**Significant at the .05 level
***Significant at the .01 level

The lowest support level is of particular interest since it differs little in financial terms from the existing level of support available from AFDC and Food Stamps. If welfare is not discounted, this program should have approximately the same independence effect as welfare. But the dissolution rate for families on this treatment greatly exceeds that of the control groups--by 96% for whites, by 67% for blacks, and by 60% for Chicanos. So we must conclude that the independence effects of welfare are discounted relative to those of an NIT. For each race-ethnic group, the plan that has the highest guarantee, 140% of the poverty line, has the smallest impact, and for all three ethnic groups yields no statistically significant differences between experimentals and controls. It appears from Table 1 that at the highest support level the income and independence effects are approximately equal, while the independence effects dominate at the lowest support levels. The basic results of this analysis are robust. As discussed in the next paper, we found no technical problem that explains away these findings.

Remarriage

Before discussing refinements to our analysis of marital dissolution we briefly discuss our findings on remarriage. The increased dissolution rates we observed do not necessarily imply increased program costs under an NIT. What matters is the number of female-headed families. The effects of an NIT on the proportion of families headed by females depend on its effects on both dissolution and remarriage rates. Thus we also study possible effects of an NIT on the rate at which women remarry.

For whites and blacks the NIT treatment has weak and insignificant total impacts on rates of remarriage. Blacks on an NIT were 7% more likely, whites 21% less likely to remarry. However, more refined analyses² reveal an experimental response that depends on the duration of the unmarried status. The NIT lowered the remarriage rate of women who recently became single relative to comparable controls. For women who had remained single for four years or longer, those on the NIT experienced a higher remarriage rate than controls. These experimental-control differences are statistically significant only for the highest-level treatment (140% of poverty line) for whites and blacks, and for the mid-level support (125%) for blacks. Unlike the findings on dissolution, programs that differ most from the non-experimental environment induce the largest response.

The NIT treatment lowers remarriage rates of Chicanos. The rate of remarriage for those on NIT treatments is approximately 20% that of the controls. This difference is large, and has only a 1% probability that it was due to chance. We have puzzled over the strength of the experimental response for this group. It appears to reflect differences between groups in the non-experimental environment. The remarriage rates of Chicano controls are much higher than those of comparable white and black controls, while the NIT lowers Chicano remarriage rates close to those for black and white controls. We suspect that the key to understanding the strong response of Chicanos lies in the discovery of the reasons behind the much higher remarriage rates among controls.

The models we use permit calculations of the total effect of various NIT programs on the equilibrium proportion of families headed by females by

combining the various impacts on dissolution and remarriage discussed separately above. Our calculations indicate that at the support levels used in the experiment, an NIT would yield only a slightly higher proportion of female-headed black families, since their remarriage response partially offsets their dissolution response. But for both whites and Chicanos, the fraction of families headed by females would increase substantially.³

Refinements

We have studied variations in the effects of the experiment over time, and analyzed the income and independence effects. Responses to the experiment may vary over time for several reasons. There may be a delay in response due to learning and adjustment to a new situation; or there may be a "burst" because the experiment accelerates the dissolution of already unstable marriages. The response of white families in the first six months is significantly higher than in the remaining 18. Thus, the average two-year response overstates the long run response. For black and Spanish-speaking samples the pattern is less clear, though it appears that the response is delayed for six months. This delay may indicate a lack of trust in the experiment and a period of testing prior to taking action. In any case, for the black and Chicano samples the two-year average response may underestimate the full effect.

Even assuming that an NIT alters rates of dissolution and remarriage for only two years, the marital response has serious implications. Our calculations show that it would take a long time--on the order of ten years--for

the proportion of female-headed households to return to its pre-NIT level. Impacts on marriage differ fundamentally from those on labor supply and other outcomes that adjust quickly to new equilibriums. We suggest that it will take much longer than two years for the impacts of an NIT on marriage to manifest themselves fully; and it will certainly take a long time for transitory impacts to recede.

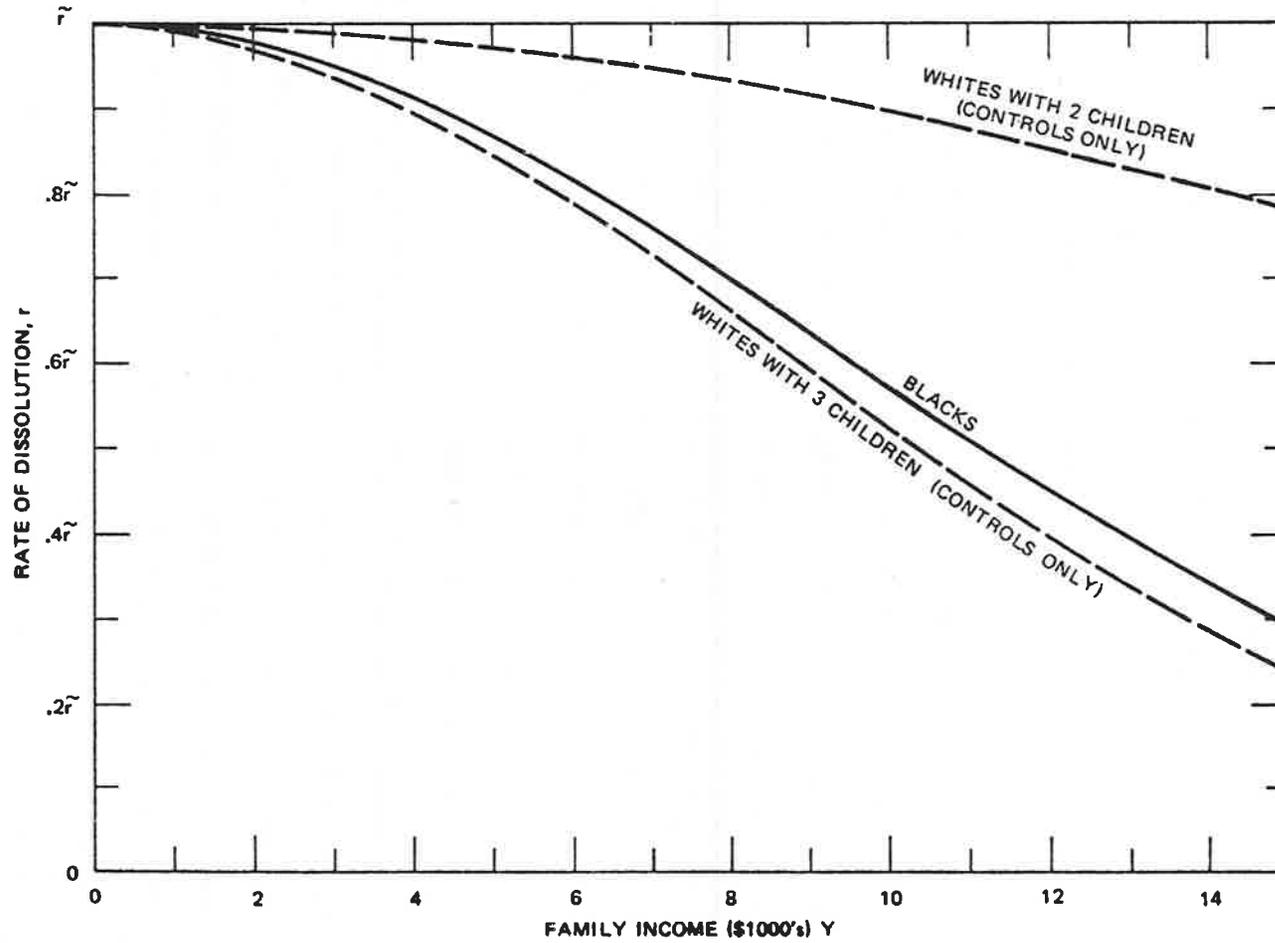
We also studied the paradox of the support levels. Recall that at a support equal to 90% of the poverty level, we found a statistically significant, large impact on dissolution, and found only a slight, insignificant impact at the 140% level. This is puzzling, but the explanation may lie in our hypothesis about income and independence effects. Possibly each of the treatments studied substantially increases independence but only the highest support tested strongly reduced strains within the family, i.e., has an income effect. Thus, the least generous plan may have only or predominantly an independence effect, while the most generous plan may have offsetting income and independence effects.

We sought to explain this pattern of experimental-control difference with a model of the income and independence effects of the NIT plans. Our model and the evidence supporting it are discussed at length elsewhere.⁴

Figure 1 graphs the estimated income effect for black couples with children and white control couples with two and three children. Throughout our analysis we have found that the income effects differ by race and family size. Recall that we estimate multiplicative rather than the more common additive models. In our models the effect of a variable such as family

FIGURE 1

ESTIMATED INCOME EFFECTS. SEE TEXT FOR DEFINITION OF \tilde{r}



income is represented as a multiplier of the base rate predicted by all other variables in the model, which we denote by \tilde{r} . Thus, the vertical axis of Figure 1 gives the dissolution rate in units of \tilde{r} , which depends on other characteristics of the family. Do not be misled by the fact that the multipliers are less than one. Our specification imposes this constraint; it has no substantive significance.

In Figure 1 we see a backward-S shaped relationship between family income and the dissolution rate. All four curves are flat for very low family incomes and steeper for incomes in the range of \$5,600 to \$15,000. For family incomes above \$15,000, which are beyond the range of most of our sample and not pictured in Figure 1, the curve is flat again, indicating that changes in income will have little effect on the dissolution rate for high income families. The curves for three children are steeper than for two children and the curves are steeper for whites than for blacks. A steeper curve means a larger percentage change in the dissolution rate for a given change in family income.

Most of the families in our sample have disposable incomes between \$4,000 and \$7,500 for the year before the experiment. The average increase in their income from the NIT was \$1,056 for the low support level, \$1,608 for the medium support level, and \$2,016 for the high support level. The response of black families to NIT may be seen directly in Figure 1. A family moves along the curve the horizontal distance of the change in their income due to the NIT. For example, for a black family with an initial disposable income of \$6,000, a payment of \$1,500 decreases the rate from $.81\tilde{r}$ to $.71\tilde{r}$, a 12% decrease.

The response for white experimental families is more complex and cannot be pictured in Figure 1. For this group we found that the effect of the change in income from the NIT was larger than the effect of the initial income on the dissolution rate. Furthermore, the change in the dissolution rate depended not only on the magnitude of the change but also on the initial level. In a graph such as Figure 1, we found that the response of whites to the NIT could not be represented in terms of movement along a line for controls but involved a change in the shape of the line as well as vertical movement. The lines retained the basic backward-S shape but the curvature changed (see Hannan, Tuma, and Groeneveld, 1978, for a fuller discussion of the white response). For a white family with three children and \$6,000 annual disposable income, a \$1,500 change in income decreases the rate from $.81\tilde{r}$ to $.74\tilde{r}$, a 9% decrease. These decreases in the dissolution rates would, of course, be offset by increases due to the independence effect.

The estimated independence effect has an S shape. When the income available to the wife outside her marriage ("independence income" which includes welfare and earnings) is low, increases in independence income have little effect on the dissolution rate. When the independence income is somewhat higher, increases in the independence income can increase the dissolution rate substantially. For women with high independence income, changes of the size expected in an NIT program have little effect on the dissolution rate.

Figure 2 shows the shape of the independence effect estimated for whites and blacks. As with the income effect, the independence effect differs by race. For whites the effect differs by family size and we present the effects to families with two or three children. For blacks we found an independence effect only for families with at least one child under ten years old (72% of our sample). When we controlled for the presence of young children we found that the independence effect did not depend on the number of children in the family.

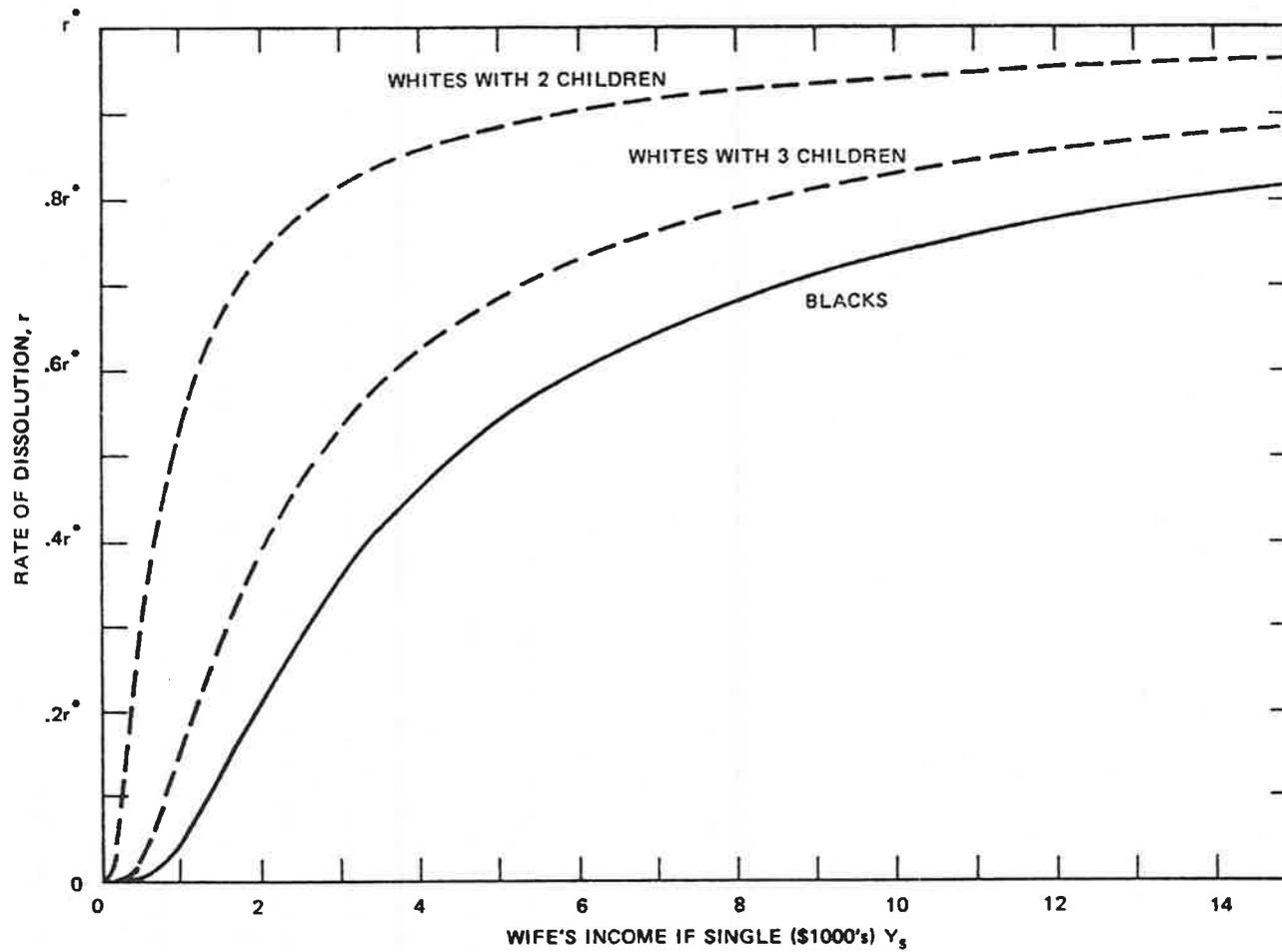
As in Figure 1 the vertical axis of Figure 2 expresses the rate as a multiplier of the portion of the rate determined by other variables which here include the income effects. We denote the portion of the rate determined by these other variables as r^* . Again, the fact that the multipliers of r^* are less than one is a constraint imposed by our model and has no substantive importance.

As discussed above, it is necessary to assume that current welfare benefits are discounted in order to explain the observed pattern of support level effects. We experimented with several discounts and found that the exact discount made little difference as long as some discount was made. Figure 2 presents the independence effects with a 50% welfare discount for all women. That is, we assume that for a married woman considering her economic position outside her marriage, two dollars of welfare (AFDC and Food Stamps) are worth one dollar of earnings or NIT payment.

Most of the women in our sample have preexperimental independence incomes between \$1,000 and \$4,000. The mean increase in annual independence

FIGURE 2

ESTIMATED INDEPENDENCE EFFECTS. SEE TEXT FOR DEFINITION OF r^*



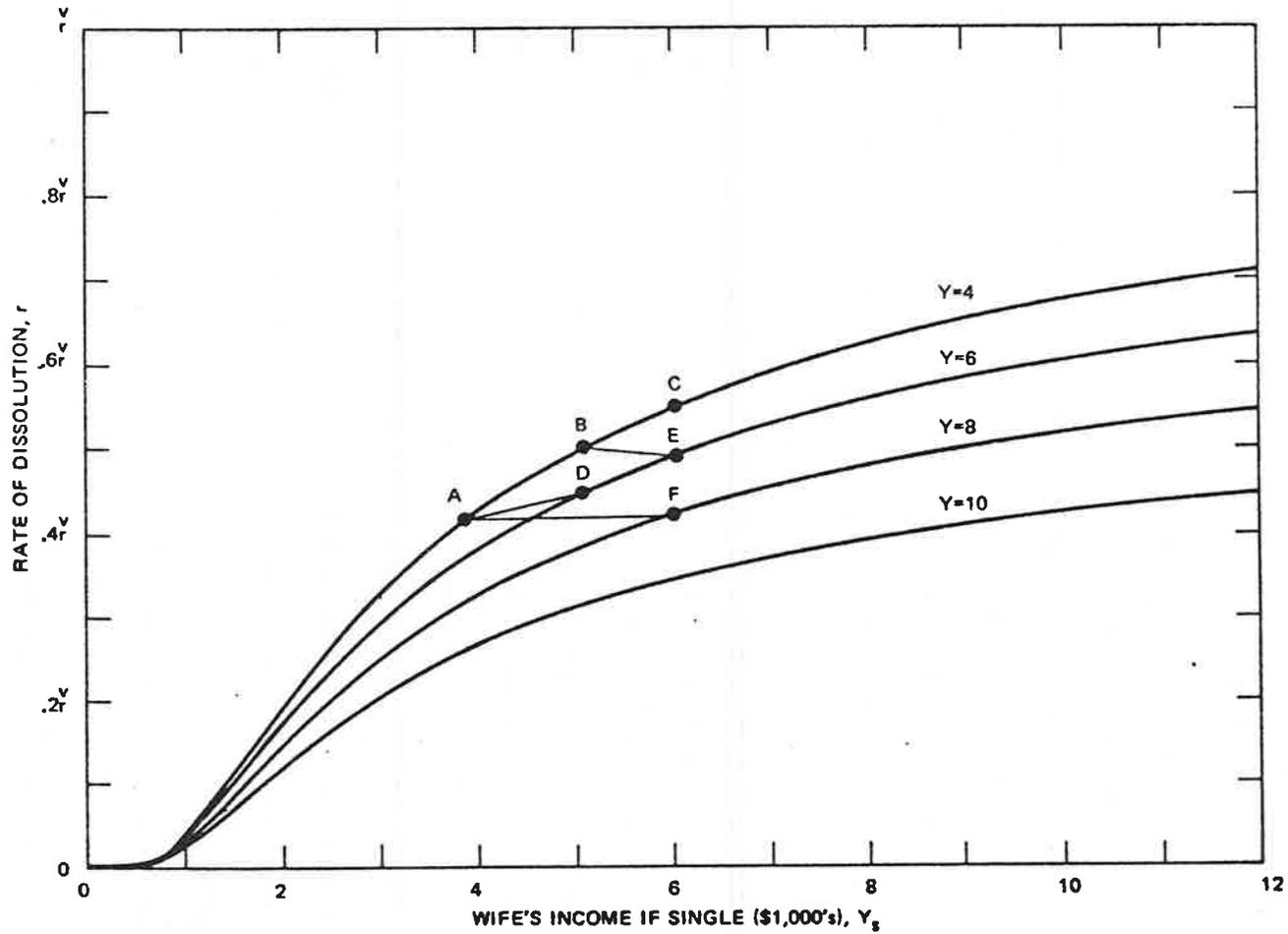
income from the NIT with welfare discounted by 50% is \$1,603 for the \$3,800 support level, \$2,258 for women on the \$4,800 support level, and \$2,872 for women on the \$5,600 support level. Because most of the women are in the range of independence income where the independence effect is steepest, changes of this magnitude have large effects on the dissolution rate. Consider, for example, a black woman with children whose initial independence income is \$2,000. A \$2,000 increase in her independence income from an NIT would change her dissolution rate from .23r* to about .48r*, an increase of over 100%. For a white woman with three children and initial independence income of \$2,000, a \$2,000 increase from an NIT increases her dissolution rate from .42r* to .65r*, a 55% increase. In contrast, a \$2,000 increase in independence income for a black woman with children and with initial independence income of \$4,000 would increase her dissolution rate from .48r* to .61r*, a 27% increase. For a white woman with three children and with initial independence income of \$4,000, a \$2,000 increase in independence income would increase her dissolution rate from .65r* to .75r*, a 15% increase. Of course, both of these examples ignore the offsetting income effect, which would depend on the level of family income before an NIT and the change in family income due to an NIT.

In order to understand the overall effect of an NIT program it is necessary to examine the pattern of combined income and independence effects.

Figure 3 illustrates the results of the combined income and independence effects. As in Figures 1 and 2, the effect is given as a multiplier of the portion of the rate determined by other variables, denoted \tilde{r} . The

FIGURE 3

ESTIMATED INCOME AND INDEPENDENCE EFFECTS:
BLACKS WITH CHILDREN Y DENOTES FAMILY INCOME.
SEE TEXT FOR DEFINITION OF r



horizontal axis is the wife's independence income. Plots for four different levels of family income are drawn. We choose the income and independence effects for blacks with children because the results are clearer graphically. The same patterns of effects can also be demonstrated for whites.

Consider a couple in Figure 3 whose family income is \$4,000 and the wife's independence income is \$4,000. They would be represented by point A in Figure 3. If an NIT plan increased the wife's independence income by \$1,000, the couple would move along the curve to point B, increasing the dissolution rate. If the NIT plan increased the family income by \$2,000 they would move down to point D, still with a net increase in the dissolution rate. On a more generous NIT plan that increased the wife's independence income \$2,000 and family income \$4,000, the couple would move from point A to point F. This would still be a slight net increase in the dissolution rate, but a much smaller increase than in the first case. Next consider a couple with the same family income but with the wife's independence income at \$5,000. If an NIT plan increased the wife's independence income \$1,000 and family income \$2,000, the couple would move from point B to point E, a net decrease. These examples illustrate how a negative income tax program can increase or decrease the dissolution rate through the combined effects of income and independence.

Effects of Various NIT Plans

Our model of the income and independence effects implies that projections of the impact of an NIT for some populations are even more complex than our earlier discussion indicates. Because income and independence effects are

nonlinear, a family's response depends on its income and the wife's independence when the NIT starts, as well as on the generosity of the program. Our estimates show that typical NIT programs increase the dissolution rate for some families, and reduce it for others.

We can address this problem by using our model for the income and independence effects of NIT payments. We found that the impact of any NIT program differs according to the race-ethnicity of the family, the number of children, and a variety of other demographic and background characteristics. We must calculate impacts separately for each combination of characteristics. We cannot be exhaustive here, but instead will illustrate the impacts of various NIT programs on rates of marital dissolution for white families with two children in which each spouse is aged 25, has 11 years of education, and the couple has been married for five years. We vary both family income prior to the NIT, and wife's pre-NIT independence income (i.e., her expected income from earnings and welfare if her marriage is dissolved). As will become clear, the latter plays a crucial role in determining the NIT impact on dissolution rates. We consider two cases typical of those we studied: (1) wives who would not be employed upon becoming single; (2) wives who would earn \$3,000 per year as single women. In each case we assume, in line with the discussion earlier and our empirical findings, that welfare is "discounted." In particular, we assume that each dollar of welfare guarantee has an independence effect half that of a dollar of earnings or a dollar from the NIT program.

We consider plots of the dissolution rate under various programs by levels of pre-NIT family disposable income. In Figures 4 and 5, the curve for controls shows how the dissolution rate declines with family income. To the extent that an NIT plan increases family income, it shifts families to the right on such curves and thereby tends to lower the dissolution rate. The income effect of any given program depends on family income in two ways. First, the payment a family receives depends on the level of their income. The lower their family income, the larger the change in their income due to an NIT. Second, the income effect is strongest for families with moderate incomes and weaker for families with either very low or high incomes. The independence effect of any particular NIT is constant across levels of pre-NIT family income for women with a given level of pre-NIT independence. If there were no income effects, the NIT plans would simply shift the control curve upwards. Because the NIT plans have both income and independence effects, their curves generally have different shapes than the control curve. Only at family income levels above the breakeven level (the region in which families receive no NIT payments or tax reimbursements) do the NIT curves parallel the control curve.

Figure 4 contains the predicted curves for families in which the wives would have no earnings after leaving the marriage. All of the NIT plans depicted in Figure 4 increase the rate of dissolution. For most levels of family income the 90% of poverty level programs increase the rate more than either the 70% of poverty level program or the 125% of poverty level program. For the 90% of poverty level program, increasing the tax rate from .5 to .7 increases the dissolution rate.

FIGURE 4

PREDICTED MARITAL DISSOLUTION RATE FOR WHITE FAMILIES WITH 2 CHILDREN,
WIFE'S GROSS EARNINGS = 0

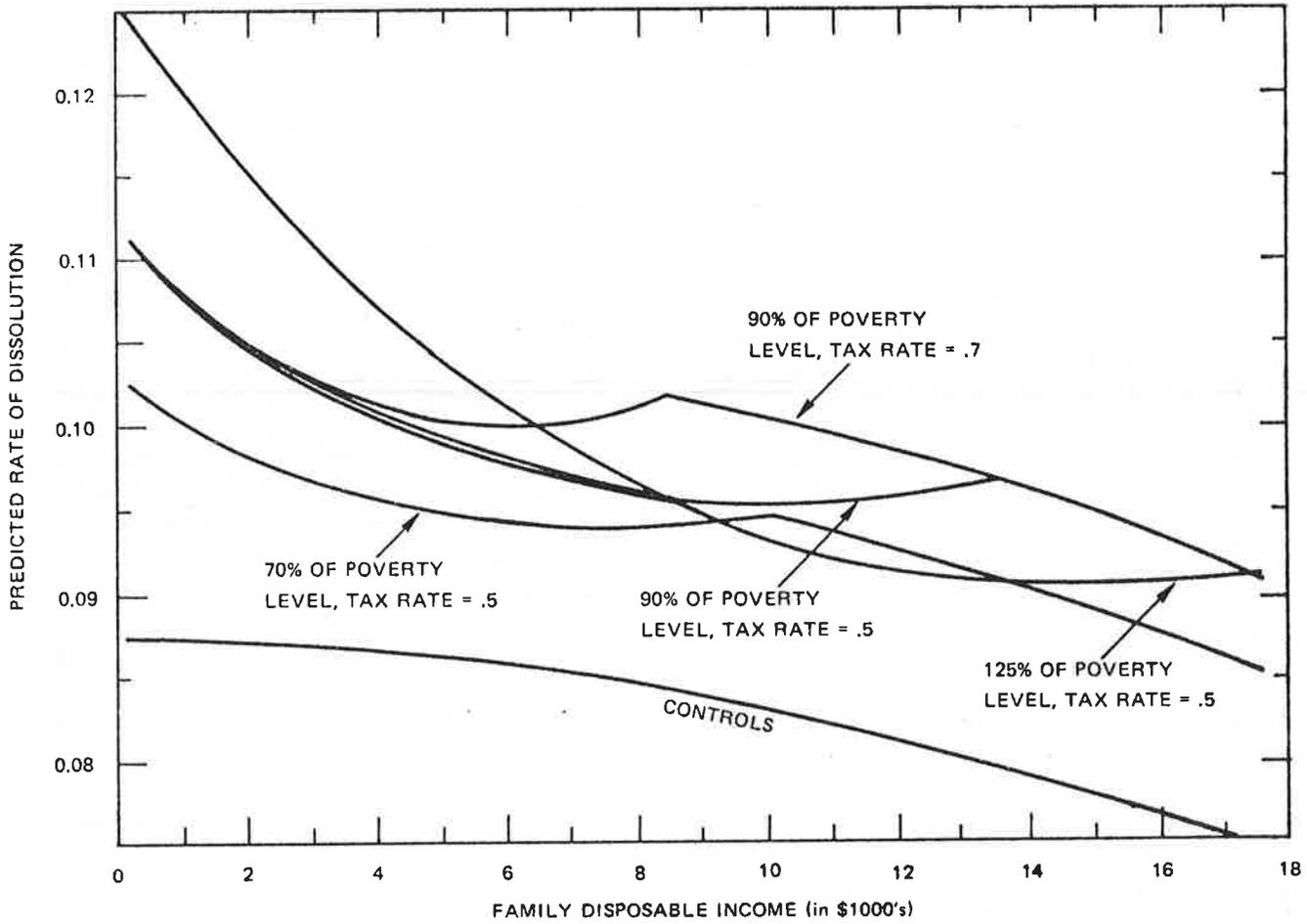
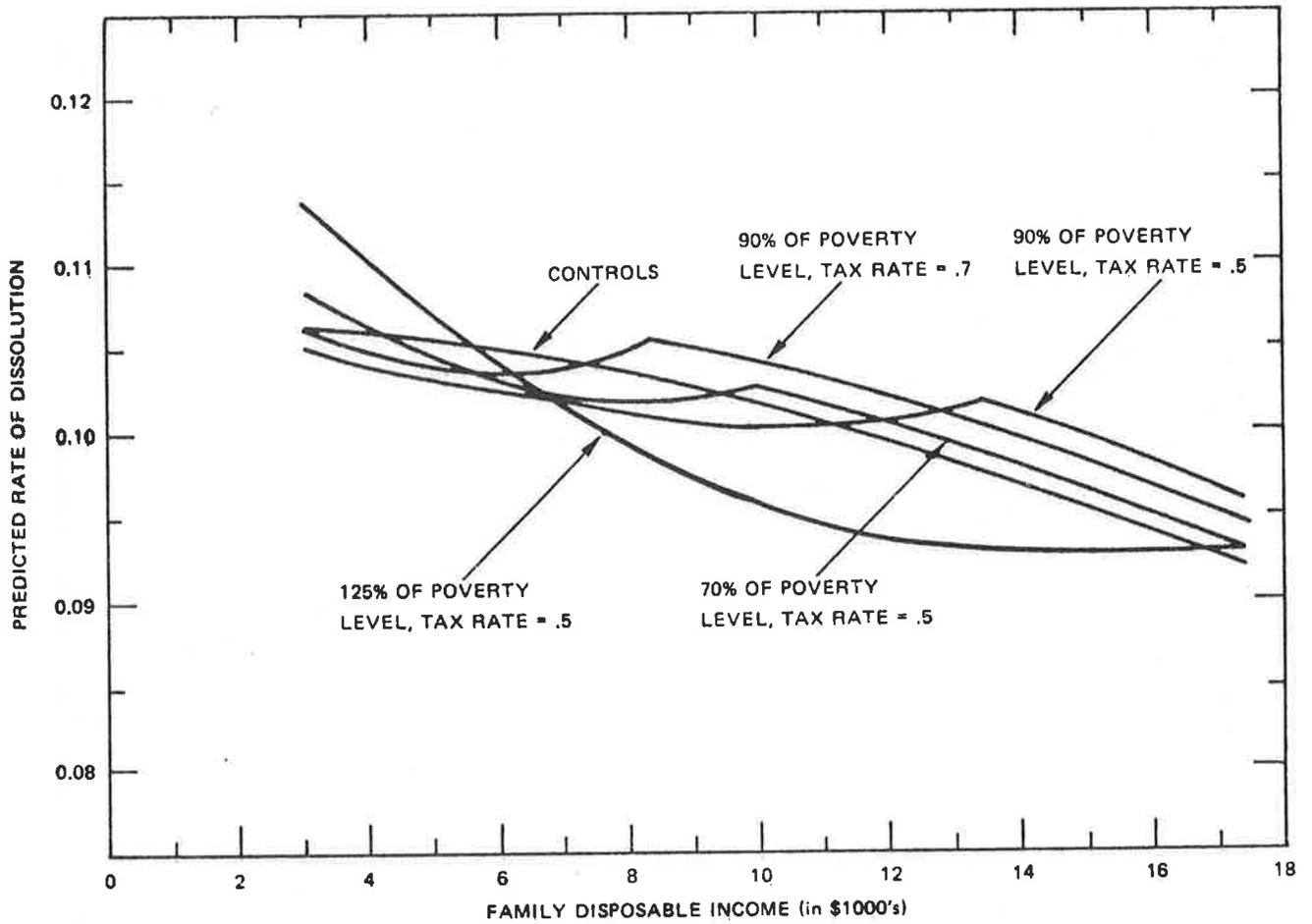


FIGURE 5

PREDICTED MARITAL DISSOLUTION RATE FOR WHITE FAMILIES WITH 2 CHILDREN,
WIFE'S GROSS EARNINGS = \$3000



It is easiest to understand the shape of the curves in Figures 4 and 5 by working from right to left. Consider the curve for 90% of poverty level and a tax rate of .5 in Figure 4. Above \$13,000 of disposable family income, this curve coincides with the curve for the 90% of poverty level and .7 tax rate program, and parallels the control curve. The distance between the NIT plan curve and the control curve is due to the independence effect. Because the family is above the breakeven level in this region and receives no NIT payments, there is no income effect. The breakeven level of the 90% of poverty level and .5 tax rate program is represented by the cusp at approximately \$13,000. To the left of the cusp there is an income effect offsetting the independence effect. Still proceeding right to left we observe that the dissolution rate declines as family income declines, reaching a minimum at approximately \$9,500. Between \$9,500 and \$13,000 we observe that as payments decrease with increasing income, the income effect decreases and is dominated by the independence effect. To the left of \$9,500 the dissolution rate increases monotonically as family income decreases. The shape of the curve in this region reflects the nonlinearity of the income and independence effects and the dependence of the income effect on the level of family income as well as on the NIT-induced change in family income. The other NIT plan curves in Figures 4 and 5 may be interpreted in the same way. The 125% of poverty level plans have a breakeven level that is greater than \$15,000 and therefore, the cusp does not appear in the figures.

Notice that in Figure 5 the control rate is higher than in Figure 4. This reflects the higher pre-NIT independence effect for women who would earn \$3,000 if single compared with women who have no single earnings. In Figure 5 all four of the NIT plans shown are below the control curve when family income is below \$7,000. Each NIT curve crosses the control curve before the breakeven cusp (this is also true of the 125% of poverty level plan, although it is not shown in Figure 5). For lower income families, all four NIT plans decrease the marital dissolution rate for families where the wife would earn \$3,000 if single and raise the marital dissolution rate for higher income families. The more generous plans lower the rate for a greater range of family incomes than the less generous plans.

Several generalizations can be drawn from these figures (and others not reported here). First, the NIT impact on marital dissolution is mainly concentrated in those families with the most dependent wives. For working wives (who we assume would remain working upon becoming single), the introduction of an NIT changes the financial alternatives to an existing marriage only slightly and thereby has less impact on decisions to end a marriage. Second, the high support and low tax programs yield the lowest dissolution rates, due to the offsetting income and independence effects.

The role of the welfare discount is important in the discussion of these figures. As we mentioned earlier, the exact discount used makes comparatively little difference when we are estimating the income and independence effects. As long as some discount is used our results account for the pattern of support level effects. However, the size of the discount plays

an important role in generating the curves in Figures 4 and 5. The discount used affects the vertical distance between the control curves and the NIT curves. (It also affects the shape of the control curve and the NIT curves above the tax breakeven, but this is of less importance.) A larger discount implies that the NIT has a larger independence effect, and consequently a greater distance between the control and the NIT curves. Had we used a 25% discount in Figures 4 and 5 rather than 50%, the control curve and the NIT curves would be closer together. The relative position of the curves, however, would remain the same. Because the independence effect is nonlinear there is no easy way to project the results from one discount to another. However, the general implications of Figures 4 and 5 remain the same when the discount is changed.

The welfare discount has important policy implications. We have seen that the effect of an NIT program on marital stability depends on the generosity of the program, and on the characteristics of the families affected by the program. As we have discussed, the change also depends on the size of the welfare discount. We include the welfare discount in our analysis to capture the effects of stigma, lack of information, and transaction costs that are part of the present welfare system. The necessity of the welfare discount in our analysis suggests that efforts to improve the current welfare system by reducing the stigma and/or transaction costs of receiving aid will affect marital stability even if benefit levels are not changed. For example, our analyses indicate that a change in AFDC administration that reduced the stigma (such as operating the program

through the IRS or the Social Security Administration) would increase the rate of marital dissolution.

CONCLUSION

We conclude with a brief summary of what we have learned about the effects of an NIT on marital stability from SIME/DIME. First, we have learned that the experimental implementation of income maintenance programs increased the rate of marital dissolution. Further, we know that the size of the increase depends on the generosity of the programs and the characteristics of the families affected. The hypothesized income and independence effects appear to be operating in the experiment. The effects are nonlinear and depend on the level of incomes as well as the NIT-induced changes in incomes.

Second, we are aware that projecting the SIME/DIME results to a national NIT program is complicated. The complications arise not only in the complexity of our findings but also from the problems inherent in projecting a short-term experiment in two cities to a permanent national program.

We are presently working on a methodology to make such projections. Even without such a methodology the findings from SIME/DIME can be of use to policymakers. It appears unlikely that a national NIT program would be neutral with respect to marital stability; thus the financial and social costs of changes in marital dissolution rates must be considered in evaluating any proposed welfare reform. In addition, our analyses have shown that the nonpecuniary aspect of welfare reform may have important effects on marital stability that also need to be considered.

NOTES

1. M. T. Hannan, N. B. Tuma and L. P. Groeneveld, "The Impact of Income Maintenance on the Making and Breaking of Marital Unions: Interim Report" (Research Memorandum No. 28, Center for the Study of Welfare Policy, Stanford Research Institute, Menlo Park, California, June 1976), and M. T. Hannan, N. B. Tuma and L. P. Groeneveld, "A Model of the Effect of Income Maintenance on Rates of Marital Dissolution: Evidence from the Seattle and Denver Income Maintenance Experiments" (Research Memorandum No. 44, Center for the Study of Welfare Policy, Stanford Research Institute, Menlo Park, California, February 1977).
2. N. B. Tuma, L. P. Groeneveld and M. T. Hannan, "First Dissolutions and Marriages: Impacts in 24 Months of the Seattle and Denver Income Maintenance Experiments" (Research Memorandum No. 35, Center for the Study of Welfare Policy, Stanford Research Institute, Menlo Park, California, August 1976).
3. Hannan, Tuma and Groeneveld (1977).
4. Id., see also M. T. Hannan, N. B. Tuma and L. P. Groeneveld, "Income and Marital Events: Test of a Model" (read at the Annual Meetings of the American Sociological Association, Chicago, September 1977).

DISCUSSION

(Lyle P. Groeneveld presented the above paper and responded to questions.)

DAVID SWOAP: As I understand it, childless couples were also included in the study. Did you look at their marital dissolution rates separately?

GROENEVELD: We looked at childless couples. We find they behave differently from couples with children, but we still find basically the same pattern of income and independence effects. The coefficients are different, and depending on the racial-ethnic group, the income effect or independence effect would dominate. What's going on isn't clear. We don't have that many childless couples; only 10% of the married couples do not have children, at least at enrollment.

SWOAP: Did you account for the fact that a mother could not leave her children under current welfare rules and still qualify, but she could do that under an NIT program?

GROENEVELD: In our modeling, we always assume that the children stay with the mother. We haven't yet studied how the family actually splits up in a dissolution.

BETH HARRIS: Do you think this discovery of an independence effect has implications for a national welfare policy--that you should figure out how to best limit that effect?

GROENEVELD: This sounds like the kind of a question where one must claim, "I'm a scientist, I don't make policy decisions." When we began our analysis we were faced with the problem not of comparing low-income families and high-income families, but of looking at the same family over time as its income changes; and we had to hypothesize what the income effect would be; and then we considered that the environment outside the marriage is also changing.

HARRIS: It seems to me if you're going to do this kind of study, it's important that you go a step further and see how the independence effect is affecting the whole family relationship, not just the dissolution rate. You seem to be assuming that the maintenance of that relationship is the most important thing for the family.

GROENEVELD: I agree. We're in the process of looking at other outcomes. There's a paper in these proceedings on psychological distress, which is one measure of these events within the family. We'll also be studying about family conflicts in the next few years.

RIKKI BAUM: What percent of the dissolutions were isolated as being initiated by the woman in the unit?

GROENEVELD: We did not ask people why they split.

BAUM: Would that not be relevant to your conclusions on the independence effect?

GROENEVELD: Well, the independence effect is really a theoretical concept. We're not suggesting that people are sitting there figuring out what their independence income is. They may be; maybe if we could find out how they do it, we could do an even better job.

(The following discussion was excerpted from a workshop session focusing on the same paper, rather than the plenary session.)

UNIDENTIFIED PARTICIPANT: You have focused on the wife's decision and her independence income. Are you ignoring the independence income of the husband?

GROENEVELD: We do consider the husband's side of it, but there are a couple of problems. Most of the husbands in the SIME/DIME are employed full time. So the majority of them would be above breakeven point when they left the marriage. It doesn't take much. On a thousand dollar guarantee, he only has to earn two thousand dollars a year before he is ineligible.

UNIDENTIFIED PARTICIPANT: Your model doesn't have anything to do with the decision-making in any way. All that you are doing is getting at the relative independence of the wife in this situation. It may be that given that kind of independence a man might figure, "All right. I can get out of here." I don't see where it makes any difference on the independence level whether we are dealing with a male or female. If the male keeps the kids when the split is occurring, then the male gets the AFDC, if he is not working. You could play games and say whichever (male or female) is earning the most money, that is the most independent person--the one losing the least.

GROENEVELD: Our paper has a more abstract and theoretical discussion, and is more detailed for the more dependent partner. But in most cases that would be the wife; in most cases it is the wife who has no wages, or a small wage, and she who takes the children.

UNIDENTIFIED PARTICIPANT: I heard recently that there are now more women than men who leave their family. If this is true, your theory would have to reflect it.

GROENEVELD: What is interesting to me is that the independence can be perceived by either party. The independence of the wife or the husband can be perceived by either party and it doesn't make any difference whether it is the man or the woman who makes the choice to go, if they go.

UNIDENTIFIED PARTICIPANT: But it does make a difference whether the husband's independent income or the wife's independent income is causing a break-up, because the designer of the program will react differently according to which one it is.

UNIDENTIFIED PARTICIPANT: What is the proportion of families in the experimental group and in the control group that were dissolved?

GROENEVELD: Ten percent of the white control group and 17% of financials dissolved their marriages in the first two years. For blacks, this was 16% for controls and 17% for financials. For Chicanos it was 14% for controls and 18% for financials.

UNIDENTIFIED PARTICIPANT: The higher the grant level that you provide to a family, the greater the likelihood that you are going to reduce the divorce rate?

GROENEVELD: That is what our findings suggest.

UNIDENTIFIED PARTICIPANT: So if it costs too much, what happens?

GROENEVELD: It is not clear that higher grant levels would cost more. We are trying to simulate SIME/DIME type programs on a national sample. We must calculate what the change in the number of families headed by single females would be and figure out what the total cost would be. If you can lower the rate of marital dissolution, even though you are supporting families at a higher rate, you may be supporting fewer people.

UNIDENTIFIED PARTICIPANT: Is your model suggesting that with an increase in income, there is an increase in independence of one family member?

GROENEVELD: The model does not require that. With the model you could increase the family's income and not increase the wife's independence income. For example, it's been suggested to me that the solution is to give money to a married family--to pay them to stay married. Other policy considerations might argue against that, but if promoting marital stability were the paramount goal, that should have an effect.

METHODOLOGICAL ISSUES IN ANALYSIS OF
MARITAL STABILITY

by

Michael T. Hannan
Associate Professor of Sociology
Stanford University

Policy analyses of proposals to replace the current welfare system with a negative income tax (NIT) usually focus on two issues. How does the proposed system alter the incentives to work? And, how does it alter the incentives to marry and stay married? In the case of labor supply responses, it is widely agreed both that changes in welfare systems alter levels of work effort and that such changes constitute a major component of the net increase in costs under a negative income tax scheme. Until recently the social science community did not evidence similar consensus that alternative welfare arrangements might substantially alter marital behavior. Policy analysts have also tended to understate the importance of marital stability to the long run cost of an NIT.

This situation appears to be changing, perhaps in part to the release of findings from the Seattle and Denver Income Maintenance Experiments (SIME/DIME) and other experiments. The claim in the President's message to Congress on welfare reform that his Better Jobs and Income Program (BJIP) would stabilize marriage has presumably also heightened interest in these issues.

All four income maintenance experiments have reported analyses of impacts on rates of marital dissolution--loosely termed "marital stability". Not surprisingly, they do not speak with one voice on this issue as there are some important differences among the experiments. For example, large samples are essential for the study of impacts on rare events like marital dissolutions. The other three experiments are hampered by relatively small samples. In fact, SIME/DIME contains more families than the other three experiments combined. Moreover, the control group in the other two urban experiments (New Jersey and Gary) encountered major changes during the experiment in rules regarding marital status and program eligibility for the regular welfare system--Aid to Families With Dependent Children (AFDC). In New Jersey, the welfare system available to control families changed, permitting families with unemployed fathers to receive AFDC. The levels of AFDC benefits also changed. In Gary, families on an NIT plan were initially told that they would lose eligibility if they dissolved their marriages. Though this rule was changed at the end of the first year, the damage to a useful study of marital stability impacts had been done. For these reasons, SIME/DIME offers the best opportunity to evaluate the effects of NIT programs on marital stability.

This paper discusses some of the main issues in the analysis of SIME/DIME findings on marital stability and their policy implications. The findings are given in the preceding paper. For blacks, whites and Chicanos, we find that an NIT substantially increases rates of marital dissolution. In more refined analyses, we find substantially lower rates of dissolution

for the higher income guarantee. Finally, we find no evidence of any systematic effect on the rate at which single black and white women (with children) remarry. Chicanas on NIT programs remarry at a substantially lower rate compared to controls. So, overall the NIT increases the proportion of female-headed families in the experimental groups.

DEFINITION OF EVENTS

The definition of the marital events studied is tied inextricably to the definition of marriage. And it is no simple matter to define a marital union. Marriage is an institution and a social tie with cultural, legal, social, political, economic and psychological dimensions. How one chooses to define a marriage depends on one's focus. The Census Bureau begs the question and simply asks people if they are "married". Researchers who use Vital Statistics that record legal divorces and separations tend to use a legal definition of marriage. Practice in the current welfare system tends to focus on parent-child relations in defining family units. That is, the AFDC procedure begins with children and then includes their mothers and fathers.

In SIME/DIME we took another approach. We regarded the family as a socio-economic unit. This view led us to define marriage as union, intended to be permanent, of a man and woman involving co-residence and pooling of resources. All couples that met these conditions were considered married, whether or not they had entered a legal marriage contract.

We record the endings of such unions. Whenever a couple ends the co-residence and informs our interviewers that the break is intended to be

permanent, we consider the marriage to be ended. Note that these are not legal divorces, though many may eventuate in divorces. What are the policy implications of these definitions?

Such findings bear only weakly on those welfare proposals that would exclude common law marriages (or long-term cohabitations) from coverage. However, non-recognition of these relationships would increase the costs of an NIT. Women would be supported as though they were single, even when their "husbands" were earning a considerable amount. Thus we suspect such proposals will not be supported widely.

Cost is not the only concern, however. Many policy makers place a high value on marriage. Presumably the concern with the stability of relations does not extend to non-legal unions. From this perspective it might matter greatly whether a marital dissolution occurs to a legal marriage or a common law marriage. Unfortunately, we cannot distinguish families on this dimension. Primarily to maintain good relations with families in the study, it was decided not to ask family heads if they were legally married.

We tend not to place great emphasis on this limitation in the data coverage. Our sampling yielded predominantly marriages of reasonably long duration--about 9 1/2 years, average, at the beginning of the experiment. At this point only 3 to 4% of couples had been married for 12 or fewer months and roughly 75 to 80% were married for three years or more.¹ Moreover, we find that, if anything, the increase in the rate of dissolution is larger for longer-duration marriages.² For both these

reasons, we conclude that our findings apply reasonably well to stable marriages--perhaps even to most legal marriages.

ATTRITION AND FRAUD

Perhaps the major challenge to correct inference in any field experiment is the problem of sample attrition. We have paid particular attention to this problem because marital dissolutions could possibly trigger events such as residential moves, job changes, and so forth, that make it likely that persons will be lost to our interviewing system. As only those receiving NIT guarantees have a strong financial incentive to stay in the study, control families are more likely to be lost. If they are lost because of events related to dissolution, we will undercount dissolutions in the control group and mistakenly infer an experimental effect on marital stability.

Luckily, we have kept attrition rates at levels that are very low for field experiments--at about 9% overall for the first two years of the experiment. Our findings are robust with respect to possible differential attrition. Even if we assume that every missing control family experienced an unreported marital dissolution, the rate for those on the NIT is still above the control rate.³ This robustness is quite important to assessments about the dependability of our findings.

Some critics have raised the question of fraud. Many families could increase NIT payments by fraudulently reporting a marital dissolution. And we have no sure way to detect this. We will, however, follow families for two years after the experiment. If there has been massive fraud, we should

see large-scale "remarriage" of former spouses. At present we are not convinced that our findings reflect fraud in large measure, despite the seeming plausibility of this hypothesis. The incentive to commit fraud should be greater for the most generous NIT programs, but in fact the data reveal an opposite pattern. Secondly, if experimental subjects are inclined towards fraud, they would not report marriages during the experiment. But for blacks and whites we find no systematic differences in reported remarriage rates between controls and NIT subjects. This is not to say there was no fraud, but that it is unlikely that the broad marital patterns are due to fraud.

IMPLICATIONS FOR POLICY

Assuming that these findings accurately reflect the impact of the NIT on marital stability in the experimental population, do they have any larger significance? We believe that they do, but that drawing such implications is an exceedingly delicate and complex task.

An obvious complication concerns geographic specialization. The states of Washington and Colorado historically have had high (legal) divorce rates. The reasons for the higher divorce rates in the West are not well understood. In any case, we cannot be confident that the magnitudes of the impacts we see in SIME/DIME are comparable to those that would obtain elsewhere. We would be especially wary of generalizing to rural populations (even in the West).

The special features of the SIME/DIME NIT plans raise additional complications for any generalization of results. We used levels of income

support considerably above those proposed in the BJIP and we did not include any work requirements for eligibility. It is not clear that we can generalize SIME/DIME even for Seattle and Denver to a plan with work requirements. But, subject to other limitations mentioned, we believe we can speak to other levels of support and other tax rates. This generalization is achieved through mathematical models of the processes by which an NIT affects dissolution rates.⁴ The model allows for two conflicting effects on rates of marital dissolution. On one hand an "income effect" tends to reduce friction in the marriage and thereby stabilize it, while an "independence effect" tends to improve alternatives outside of marriage and thus destabilize it. And the magnitude of these effects in a given NIT depends on a family's initial level of income and independence. Thus, the same program may stabilize some marriages and destabilize others. Microsimulation of results is now in progress using this model, but the presence of nonlinear opposing effects greatly complicates the extension of these results to larger populations.

Generalization is also complicated by the fact that the NIT plan with an income support and tax rate very similar to the existing welfare system should have such a different impact on marriages. This paradox may be due to nonpecuniary dimensions of the programs and we have identified a number of ways in which SIME/DIME and welfare differ.⁵ These focus on stigma, information about programs, transaction costs, etc. These differences--which relate to stigma--suggest that a dollar of welfare will have less impact on behavior than a dollar from earnings or an NIT. That is, welfare is discounted in its income and independence effects.

The welfare discount raises further problems of generalization. Our findings may in some degree be site-specific in that levels of information about benefits on public welfare and the stigma associated with participating in welfare may be different in Seattle and Denver compared to other areas. These factors may also vary among subpopulations in the same urban area.

Moreover, our experiment was a rather special situation in which recipients had relatively high levels of information and incurred low transaction costs in participating. Moreover, the program had no larger social definition such that participation was likely to be stigmatizing. But would this be true of a national NIT? The President's message to Congress regarding the BJIP argues that the proposal would reduce the stigma of welfare participation. If so, our results may bear closely on the income support portions of the proposal. My main point is that the SIME/DIME findings suggest that much more attention be paid to the nonpecuniary characteristics of alternative welfare plans and to the manner in which these affect marital stability.

The last major complication in generalizing our results concerns separating long-term from short-term effects. Social experiments probably induce initial responses that differ from long-run response. And it is always difficult to separate the two in a fixed-length experiment, even when there are two or three fixed experimental periods as in SIME/DIME. The problem is particularly acute in the study of rare events like changes in marital status. We have carefully studied variation over time in marital

stability over the first 24 months of experimental time. We find some variations over time in response to the NIT. For whites the peak response occurs in months 6 through 12 while for blacks and Chicanos, it is in months 7 through 18. Thus, our findings may overstate the long term effects of an NIT. Adjusting for time variation only slightly lowers our estimates of overall impacts, however.

Important consideration must also be given to long-term impacts on the stability of marriages formed under an NIT plan. Unfortunately we have too few observations to conduct reliable statistical analysis of stability of such new marriages. Perhaps when all of the SIME/DIME data are available for analysis we will have sufficient observations on new marriages to address this problem.

CONCLUSIONS

Despite the limitations on these findings and the complexity of the issues, the SIME/DIME results are the best guide available to the likely impact of NIT plans similar to those used in SIME/DIME. The results are extremely dependable by statistical standards and appear robust with respect to model specification, choice of estimator, and attrition bias. So my view on the matter of applications boils down to the following. The SIME/DIME results do bear on issues of policy choice; but drawing precise implications from these results demands very intense and careful analysis.

NOTES

1. Michael T. Hannan, Nancy Brandon Tuma and Lyle P. Groeneveld, "The Impact of Income Maintenance on the Making and Breaking of Marital Unions: An Interim Report," (Research Memorandum No. 28, Center for the Study of Welfare Policy, Stanford Research Institute, Menlo Park, California, 1976).
2. Id.
3. Id. at Chapter 8.
4. See Michael T. Hannan, Nancy Brandon Tuma and Lyle P. Groeneveld, "A Model for the Effect of Income Maintenance on Rates of Marital Dissolution: Evidence from the Seattle and Denver Income Maintenance Experiments," (Research Memorandum No. 44, Center for the Study of Welfare Policy, Stanford Research Institute, Menlo Park, California, 1977); and Michael T. Hannan, Nancy Brandon Tuma, and Lyle P. Groeneveld, "Income and Independence Effects on Marital Dissolution: Results from the Seattle and Denver Income Maintenance Experiment," American Journal of Sociology 83 (November 1978).
5. Michael T. Hannan, Nancy Brandon Tuma and Lyle P. Groeneveld, "The Impact of Income Maintenance on the Making and Breaking of Marital Unions: An Interim Report," (Research Memorandum No. 28, Center for the Study of Welfare Policy, Stanford Research Institute, Menlo Park, California, 1976), and Michael T. Hannan, Nancy Brandon Tuma and Lyle P. Groeneveld, "Income and Marital Events: Evidence from an Income Maintenance Experiment," American Journal of Sociology 82 (May 1977): pp. 1186-1211.
6. Id. See also Michael T. Hannan, et al., "A Model for the Effect of Income Maintenance on Rates of Marital Dissolution."

DISCUSSION

KAY THODE: If I have understood you, the higher support levels had a lower rate of dissolution and families assigned to controls were higher-income families. Would that affect your findings?

HANNAN: I hope not. We've tried our best to control for the assignment where the model allows the effect--the percentage increase in the rate of dissolution--to vary by two pre-experimental characteristics like family income. We believe we're adjusting for that.

THODE: Did the age of the couples in the controls differ?

HANNAN: The control families have a similar distribution by age and marriage.

DAVID SWOPE: You indicated that it was difficult to generalize across the nation because of the higher divorce rate in the north and the west; but wouldn't the presence of the controls solve that problem?

HANNAN: It might. It depends upon the nature of the problem. What I'm worried about is that there is something cultural or something in the local norms about marriage that leads people to be especially sensitive to small changes in economic circumstances in making these marital decisions.

LEE BAWDEN: I'm curious why you didn't use the three-year sample; but if you did, does the same pattern come through with respect to support levels?

HANNAN: All the analysis is done on the entire sample, but contains a multiplier for the three-year and five-year difference.

TED LANE: In the state of Alaska I recently read that the divorce rate among pipeline workers--who had a sudden upward shift in income--increased dramatically. It just may be that marriage formed under one set of economic rules suddenly starts to dissolve when you're suddenly faced with a change in these rules. And after those new rules are in place for any long period of time, the transitory phenomenon should dissipate.

HANNAN: That may well be true, but those who had the largest change in their economic circumstances didn't necessarily display the largest experimental response. Further, even if you affect only existing marriages, the economic effects would last a very long time--something like 20 years.

CRONIN: Just another possible theory about the soaring divorce rate in the West--that it relates to stability of residence. Is that something that you looked at?

HANNAN: We have that information and I'm not sure that it's been analyzed.

ROBERT LERMAN: I notice you didn't pay much attention to--at least in the text of the papers--the tax break. If one conjectures that what's relevant is the relative contribution of the two parties to the marriage, the higher the tax rate the lower the net contribution by a worker to the family's income.

HANNAN: In the paper we wrote we did not discuss in detail the tax rates. That does not mean that we think that tax rates are not important. But if we looked at differences among treatments, we find significant differences among support levels, but not among tax rates. That could be due to technical problems. I think it reflects the fact that the families who respond most to the NIT are families in which the woman was not employed prior to the experiment. If we conjecture that when she leaves the marriage she would stay unemployed and go on welfare or an NIT plan, then the tax rate is not too relevant, since there are no earnings involved.

BETH HARRIS: Did you ask any of the people who were interviewed if they thought that the NIT program had any effect on their marriages?

HANNAN: No.

PROGRESS REPORT ON STUDIES OF THE EFFECTS
OF INCOME MAINTENANCE ON INFANT HEALTH STATUS

by

Barbara H. Kehrer
Senior Economist, MPR

and

Charles M. Wolin
Economist, MPR

Poor health is a well-known correlate of poverty. The question arises, therefore, whether income transfers by means of a negative income tax (NIT) can have a beneficial impact on the health status of the poor. This paper reports on efforts to investigate the effects of an NIT on two objective indices of infant health: the rate of infant mortality (death within the first year of life) and the incidence of low birth weight (less than 5.5 pounds). Both measures are aspects of what the public health literature refers to as "adverse pregnancy outcomes."

We first present information on the problems of infant mortality and low birth weight and indicate that, given the current state of medical knowledge, improvements in these rates are attainable. Next, we set forth the rationale for expecting that an NIT program might influence infant health. We then describe a planned study to explore these issues using data gathered on participants in the Seattle and Denver income maintenance experiments (SIME/DIME). Finally, we present findings from a recently completed study on the Gary, Indiana, income maintenance experiment, which investigated the impact of the NIT on low birth weight.¹

INFANT HEALTH PROBLEMS IN THE U.S.

The infant mortality rate in the United States was 17.7 per 1000 live births in 1973.² This rate is considerably higher than the rates in other nations. In 1971, for instance, the rates in Sweden, the Netherlands, Finland, Japan, and Norway all were below 13 per 1000 live births.³ Within the U.S., there are substantial differences in infant mortality rates among demographic groups. According to data for 1973, the rate for nonwhites is about 60% higher than the rate for whites.⁴ These differences in infant mortality rates among nations and between whites and nonwhites in the U.S. mirror similar differences in the incidence of low birth weight, and the differential mortality rates have been attributed to the differential incidence of low birth weight.⁵

Birth weight is one of the most reliable indicators of an infant's chances of survival. According to U.S. data for 1960, almost all infants weighing 2.25 pounds or less will die in the first year. As birth weight increases, the probability of survival increases systematically until very high birth weights are reached and survival probability declines. The optimal birth weight from the standpoint of survival is within the range of 6.5 to 8.5 pounds.⁶ Finally, children whose birth weights are 5.5 pounds or less tend to experience a wide variety of health problems compared with children who weighed more than 5.5 pounds at birth.⁷

In view of these disparities in infant health status among nations and socioeconomic groups with the United States, it appears that present medical knowledge is such that the health status of disadvantaged infants

in the United States could be improved. The question that motivates this study is whether such improvement might result from the implementation of an NIT.

FACTORS ASSOCIATED WITH ADVERSE PREGNANCY OUTCOMES

Two initial tasks of the study were: (1) to identify mechanisms through which an NIT might intervene in the process that results in these adverse pregnancy outcomes, and (2) to identify factors which, though not subject to change, should be taken account of in the analysis. Factors associated with high risk of adverse pregnancy outcomes may be classified into three general categories: (1) demographic characteristics of the mother and infant, (2) physiological characteristics of the mother, and (3) behavioral practices of the mother. A summary list of such factors is provided in Table 1.

Demographic Characteristics of Mother and Infant

Most of the demographic characteristics of the mother and child that influence the likelihood of an adverse pregnancy outcome are not subject to change. One of these is the mother's age. Children born to adolescent mothers and women 35 and older suffer significantly higher rates of both low birth weight⁸ and infant mortality.⁹ Birth order also is important. Birth weight appears to rise with birth order for mothers over 24 years of age.¹⁰ It also has been reported, without controlling for maternal age, that the incidence of low birth weight is lowest for second births and higher for the first, third, and subsequent births.¹¹ Again without controlling for maternal age, infant mortality rates are associated positively with birth order.¹²

TABLE 1

SOME FACTORS ASSOCIATED WITH HIGH RISK OF ADVERSE PREGNANCY OUTCOMES

Category	Factor
Demographic	Maternal age under 18 or over 35 ^a Birth order: first births and higher orders greater than two ^a Infant's sex female ^a Plurality (i.e., multiple births) ^a Nonwhite race ^a Maternal residence at high altitude
Physiological	Low maternal prepregnancy weight ^a Intrauterine infections, such as rubella, cytomegalic inclusion disease, toxoplasmosis, ^b and syphilis ^b Chromosomal abnormalities Maternal cyanotic heart disease Placental insufficiency Asymptomatic bacteriuria ^b Maternal pre-eclampsia ^b Maternal eclampsia, or toxemia ^b Maternal chronic hypertension ^b
Behavioral	Short interval between pregnancies ^a Maternal cigarette smoking Maternal drug addiction Limited maternal weight gain during pregnancy Maternal malnutrition Work at paid employment during pregnancy

^a Not subject to change once conception has taken place.

^b Medical or obstetric conditions that are diagnosable and treatable.

The infant's sex also influences birth weight and the likelihood of infant mortality. Female infants weigh an average of one half pound less at birth than male infants, and the incidence of low birth weight is higher among females than males.¹³ On the other hand, holding birth weight constant, males experience higher infant mortality rates than females.

Another factor is plurality. Plural births (e.g., twins or triplets) have significantly higher infant mortality rates and are more likely to have low birth weight than single births.¹⁴

As mentioned above, race also is associated with pregnancy outcomes. Infant mortality rates for nonwhite infants are almost twice as high as those for white infants. In addition, nonwhite women give birth, on average, to lower weight children than white women.¹⁵ However, when controlling for birth weight nonwhite infants with low birth weight experience lower mortality rates than do whites.¹⁶

A final demographic characteristic that is important for the planned SIME/DIME analysis is residence at high altitude. This factor is associated with intrauterine growth retardation and higher probability that an infant will have low birth weight.¹⁷

Physiological Characteristics of the Mother

The mother's prepregnancy weight has been shown to exhibit a statistically significant positive association with birth weight.¹⁸ Moreover, the incidence of low birth weight (and, hence, the probability of infant mortality) declines with increases in prepregnancy weight. Prepregnancy weight is not, by definition, subject to change once conception occurs.

However, improvements in nutrition prior to conception might be reflected in higher prepregnancy weight.

The remaining physiological factors listed in Table 1 are medical and obstetric problems. Some of them are treatable and therefore subject to change. Rubella, cytomegalic inclusion disease, and chromosomal abnormalities are, at present, not treatable. In addition, treatment is problematic for expectant mothers suffering cyanotic heart disease or placental insufficiency. The remaining medical problems (asymptomatic bacteriuria, toxoplasmosis, syphilis, pre-eclampsia, eclampsia, and chronic hypertension) are both diagnosable and treatable.¹⁹

Behavioral Determinants

In addition to the demographic and physiological characteristics cited above, certain behavioral practices of the mother may increase the likelihood of an adverse pregnancy outcome. One that is not subject to change once conception occurs (assuming a negligible voluntary abortion rate) is a short time interval between pregnancies: rates of adverse pregnancy outcomes are particularly high for intervals less than 12 months.²⁰

The relationship between maternal smoking and increased risk of an adverse pregnancy outcome is well established.²¹ Effective "treatment" of this problem is for the mother to stop smoking. Hence, the earlier in pregnancy a woman who smokes begins to receive prenatal care, the sooner and more frequently is she likely to be urged to break the smoking habit.

Weight gain during pregnancy is another behavioral practice that may influence birth weight and is subject to intervention during the course of prenatal care. It has been demonstrated that, controlling for prepregnancy

weight, weight gain during pregnancy has a positive association with birth weight and a negative association with the incidence of low birth weight.²²

Another factor listed in Table 1, whose alleviation need not depend on medical intervention, is maternal malnutrition. It is generally accepted that severe nutritional deprivation in an expectant mother increases the likelihood of her bearing an infant of low birth weight, and recent research suggests that more moderate malnutrition also will affect birth weight.²³

Finally, some limited evidence suggests that the risk of an adverse pregnancy outcome is higher for women who work at paid employment during their pregnancies than for other women.²⁴

It is clear that some of the causes of adverse pregnancy outcomes are amenable to remedy. Many require medical intervention. Thus, both the timing and the quantity of prenatal care received are likely to be important in reducing the probability of low birth weight and infant mortality in low-income populations.²⁵

THE PLANNED SIME/DIME STUDY

We view the impact of income maintenance on pregnancy outcomes as the result of a complex set of family responses to the NIT. The structural model underlying our analytical strategy is a simultaneous-equations system in which a pregnancy outcomes response reflects NIT effects on family income and on hours worked, and is developed at length elsewhere.²⁶ In this section, we summarize the hypotheses concerning the impact of a negative income tax (NIT) on adverse pregnancy outcomes. We then turn to the data requirements for the planned SIME/DIME study.

Hypotheses

We expect to observe lower rates of adverse pregnancy outcomes among SIME/DIME experimental families compared with controls, due to greater and earlier use of prenatal care, improved nutrition, and reduction in hours of paid employment by pregnant women in the experimental group.

There are several reasons to expect experimental families to make greater and earlier use of prenatal care. First, by increasing income, the NIT plan increases ability to pay for medical care. Second, the program effectively subsidizes the time cost of obtaining medical services for an individual who takes unpaid time off from work to visit the doctor. (The "tax rate" feature of the plan provides partial compensation for the loss in earnings resulting from the loss in work time.) Finally, SIME/DIME also specifically subsidized the money cost of medical services received,²⁷ which is likely to further encourage use of prenatal care.

Expectant mothers in poor families in the U.S. may be sufficiently malnourished to influence their infants' birth weights. Consequently, improvements in maternal nutrition as a result of the NIT may have a beneficial effect on pregnancy outcomes.

Finally, if paid employment is associated with a higher risk of an adverse pregnancy outcome, the reduced labor force activity permitted by the NIT may have a positive impact on fetal and infant health.

Data Requirements

Two separate data sources will be drawn upon for the SIME/DIME study: SIME/DIME interview data and vital records maintained by the state governments of Washington and Colorado. The experiment's data files will provide information on such variables as family size, wage rates, and nonwage income.

These files also will be the major source of information on the identities of the children born to experiment participants and their dates of birth. Birth records will provide information on birth weight, sex, plurality, mother's age at the time of birth, birth order, and the time interval between pregnancies. Death records will be used to identify deaths that occurred among experiment participants and will provide data on age at death and the cause of death. In addition, death records will be used to identify children who died too soon after birth to have "entered" the experiment's files.

RESULTS FROM THE GARY INCOME MAINTENANCE EXPERIMENT

An analysis of birth weight response to an NIT has already been performed using data on 404 single live births to participants in the Gary income maintenance experiment.²⁸ The results suggest that an NIT will have a beneficial impact on low-birth-weight infants.

Table 2 shows the estimated effects of the NIT for 12 cohorts of mothers. These 12 cohorts were defined by varying the combinations of the following risk factors: mother's age (under 18, between 18 and 34, or over 34), whether she ever smoked, and whether there was a short time interval (less than 16 months) between pregnancies. In nine of the 12 cohorts, there is a beneficial treatment effect, and in four of the nine the effect is statistically significant at the 5% level. Moreover, these four are very likely the highest-risk cohorts of the 12: mothers in all three age groups who both smoked and had a short interval between pregnancies, and mothers under 18 who smoked. The size of the beneficial effect for the four cohorts ranges from .38 pound to 1.17 pounds.

TABLE 2

GARY STUDY RESULTS: TREATMENT EFFECTS BY MOTHER'S AGE,
INTERVAL BETWEEN PREGNANCIES, AND SMOKING STATUS

No.	Cohort			Average weight gain (or loss) associated with NIT (in pounds)
	Age	Short Interval?	Smoker?	
1	under 18	Yes	Yes	1.17*
2	under 18	No	Yes	0.38*
3	under 18	Yes	No	0.42
4	under 18	No	No	(0.03)
5	18-34	Yes	Yes	0.62*
6	18-34	No	Yes	0.01
7	18-34	Yes	No	0.02
8	18-34	No	No	(0.26)*
9	over 34	Yes	Yes	1.04*
10	over 34	No	Yes	0.26
11	over 34	Yes	No	0.29
12	over 34	No	No	(0.14)

*Estimated effect is statistically significant at the .05 level.

Three cohorts show an adverse response to NIT treatment. In only one of these is the adverse effect statistically distinguishable from "no effect:" this is the healthiest cohort, mothers between 18 and 34 who neither smoked nor had a short interval between pregnancies. The apparently perverse effect of an NIT on the healthiest cohort is troublesome. However, this finding might be the result of the Gary sample selection method: birth records were collected only for infants recorded in the experiment's files as entering families. Because infants who died within approximately a month of birth probably were never listed as having entered the family, they were excluded from the Gary study. Most of these children very likely had low birth weights. If the low-weight babies whose births were not recorded were concentrated among mothers in control families, the sample would understate the effect of income supplements on the reduction in the incidence of low birth weight.

CONCLUDING REMARKS

An NIT plan is hypothesized to have a beneficial impact on the outcome of pregnancy among the poor through at least three mechanisms: increased and earlier use of prenatal care, improved nutrition, and the opportunity to withdraw from paid employment. Evidence from the Gary experiment suggests that an NIT will have such an effect for children born to women at high risk of an adverse pregnancy outcome. Anomalous results were found for infants of low-risk women. It is possible, however, that these results are a reflection of bias in the sample selection strategy, which omitted infants who died too soon after birth to have entered the experiment's records. The study proposed for SIME/DIME will avoid this problem of missing observations

by locating the birth records of such infants through a search strategy using death records.

The SIME/DIME study will also augment the Gary study by examination of several racial and ethnic groups (all the participants in the Gary experiment were black). In addition, the SIME/DIME study will add to the Gary findings the analysis of the influence of an NIT on infant death, permitting inferences about another adverse pregnancy outcome in which public health policy makers have shown strong interest. Nonetheless, the finding to date, that an NIT does seem to generate gains in birth weight among infants born to high-risk mothers, is dramatic evidence of the broad benefits of a more adequate income support system.

NOTES

1. The 1,799 families who participated in the Gary experiment (1,028 experimentals and 771 controls) represented a black urban sample with a high proportion of single parent families headed by women.
2. Vital Statistics of the United States, 1973. Volume II--Mortality. Part A. DHEW Publication No. (HRA) 77-1101, National Center for Health Statistics, Rockville, Md., 1977.
3. United Nations Statistical Yearbook: 1972, New York, 1973, Table 21, pp. 89-94. From Helen M. Wallace and Hyman Goldstein, "The Status of Infant Mortality in Sweden and the United States," Journal of Pediatrics 87 (December 1975): pp. 995-1000.
4. Vital Statistics of the United States, 1973. Volume I--Nativity. DHEW Publication No. (HRA) 77-1113. National Center for Health Statistics, Rockville, Md., 1977. The difference in rates across races and socio-economic groups is well known; see, for example: Victor R. Fuchs, Who Shall Live? Health, Economics and Social Choice (New York: Basic Books, 1974) and Diana Hunt, "Health Services to the Poor: A Survey of the Prevalence and Causes of Infant Mortality in the U.S.," Inquiry 6 (December 1969): pp. 27-38.

5. For example, Fuchs, op. cit., pp. 34-35; R. L. Naeye et al., "Urban Poverty: Effects of Prenatal Nutrition," Science 166 (November 1969): pp. 1026, and "International Comparison of Perinatal and Infant Mortality: the United States and Six West European Countries." Vital and Health Statistics Series 3, No. 6, Public Health Service Publication No. 1000, National Center for Health Statistics, Washington, D.C., March 1967.

6. "A Study of Infant Mortality from Linked Records: By Birth Weight, Period of Gestation, and Other Variables, United States." Vital and Health Statistics, Series 20, No. 12, DHEW Publication No. (HSM) 72-1055, Rockville, Md.: National Center for Health Statistics, May 1972.

7. See, for example, "Trends in Prematurity: United States: 1950-67," Vital and Health Statistics Series 3, No. 15, DHEW Publication No. (HSM) 72-1030, Rockville, Md., National Center for Health Statistics, January 1972; J. W. B. Douglas and C. Mogford, "Health of Premature Children from Birth to Four Years," British Medical Journal 1 (April 4, 1953): pp. 748-54; Hilda Knobloch, Benjamin Pasamanick, Paul A. Harper and Rowland V. Rider. "The Effect of Prematurity on Health and Growth," American Journal of Public Health 49 (September 1959): pp. 1164-73; Cecil Mary Drillien, "School Disposal and Performance for Children of Different Birthweight Born 1953-1960," Archives of Diseases of Childhood 44 (1969): pp. 562-70, and Sandra Scarr, "Effects of Birth Weight on Later Intelligence," Social Biology 16 (December 1969): pp. 249-56.

8. See: James F. Donnelly, "Etiology of Prematurity," Clinical Obstetrics and Gynecology 7 (1964): pp. 647-57; Victor B. Penschaszadeh, Janet B. Hardy, E. David Mellits, Bernice H. Cohen and Victor A. McKusick, "The Effect of Certain Maternal Characteristics on Birth Weight, Gestational Age and Intra-Uterine Growth," Hopkins Medical Journal 131 (July 1972): pp. 11-23; and Arnold J. Rudolph, "Anticipation, Recognition and Transitional Care of the High-Risk Infant," Care of the High Risk Neonate, Marshall H. Klaus and Avroy A. Fanaroff, eds., W. B. Saunders, Philadelphia, 1973, pp. 23-24.

9. "A Study of Infant Mortality from Linked Records," May 1972, op. cit.

10. Steve Selvin and Dwight T. Janerich, "Four Factors Influencing Birth Weight," British Journal of Preventive Social Medicine 25 (1971): pp. 12-16.

11. "A Study of Infant Mortality from Linked Records," May 1972, op. cit.

12. Ibid.

13. Alexander J. Schaffer and Mary Ellen Avery. Diseases of the Newborn, 3rd ed., W. B. Saunders, Philadelphia, 1971.

14. "A Study of Infant Mortality from Linked Records," May 1972, op. cit.

15. Vital Statistics of the United States, 1973. Volume I--Natality, op. cit.
16. "A Study of Infant Mortality from Linked Records," May 1972, op. cit.
17. See, for example, S. Gorham Babson and Ralph C. Benson, Management of High-Risk Pregnancy and Intensive Care of the Neonate, 2nd ed., C. V. Mosby, St. Louis, 1971; and Avron Y. Sweet, "Classification of the Low-Birth-Weight Infant," Care of the High-Risk Neonate, Marshall H. Klaus and Avroy A. Fanaroff, eds., W. B. Saunders, Philadelphia, 1973, pp. 36-57.
18. Nicholson J. Eastman and Esther Jackson, "The Bearing of Maternal Weight Gain and Pre-Pregnancy Weight on Birth Weight in Full-Term Pregnancies," Obstetrical and Gynecological Survey 23 (1968): pp. 1003-1025; and Kenneth R. Niswander, Judith Singer, Milton Westphal, and William Weiss, "Weight Gain during Pregnancy and Prepregnancy Weight," Obstetrics and Gynecology 33 (April 1969): pp. 482-91.
19. Babson and Benson, op. cit.; Ian A. McDonald, A Method of Obstetrics and Gynaecology, Pergamon Press, Potts Point, NSW, Australia, 1971; and Robert E. L. Nesbitt, Edward R. Schlesinger, and Sam Shapiro, "Role of Preventative Medicine in Reduction of Infant and Perinatal Mortality," Public Health Reports 81 (August 1966): pp. 691-702.
20. David M. Kessner, James Singer, Caroline E. Kalk, and Edward R. Schlesinger, Infant Death: An Analysis by Maternal Risk and Health Care, National Academy of Sciences, Washington, D.C., 1973; and Joe D. Wray, "Population Pressure on Families: Family Size and Child Spacing," Rapid Population Growth, Study Committee, Office of the Foreign Secretary, National Academy of Sciences, Johns Hopkins Press, Baltimore, 1971, pp. 403-61.
21. See, for example, Babson and Benson, op. cit.; Penchaszadeh et al., op. cit.; and Sweet, op. cit.
22. Eastman and Jackson, op. cit., and Niswander et al., op. cit.
23. See the literature review in Lawrence Bergner and Mervyn Sessner, "Low Birth Weight and Prenatal Nutrition: An Interpretative Review," Pediatrics 46 (December 1970): pp. 946-65; Naeye et al., op. cit., and David Rush, Hillard Davis, and Mervyn Susser, "Antecedents of Low Birthweight in Harlem, New York City," International Journal of Epidemiology 1 (1972): pp. 375-87.
24. Alice Stewart, "A Note on the Obstetric Effects of Work During Pregnancy," British Journal of Preventive Social Medicine 9 (1955): pp. 159-61; and Martti A. Kauppinen, "The Correlation of Maternal Heart Volume with the Birth Weight of the Infant and Prematurity," Acta Obstetrica et Gynecologica Scandinavica 66, supplement 6 (1967).

25. A number of studies described in the public health literature point to the importance of the timing and quantity of prenatal care received in reducing the probability of adverse pregnancy outcomes. See, for example, Kessner et al., op. cit., B. Y. Iba, J. D. Niswander, and L. Woodville, "Relation of Prenatal Care to Birth Weights, Major Malformations, and New-born Deaths of American Indians," Health Service Reports 88 (October 1973): pp. 697-701; Gerald Wiener and Toby Milton, "Demographic Correlates of Low Birth Weight," American Journal of Epidemiology 91 (1970): pp. 260-72; and Jack Zackler, Samuel L. Andelman, and Frank Bauer, "The Young Adolescent as an Obstetric Risk," American Journal of Obstetrics and Gynecology 103 (February 1, 1969): pp. 305-312.

26. Barbara H. Kehrer and Charles M. Wolin, "Impact of Income Maintenance on Low Birth Weight: Evidence from the Gary Experiment," MPR Working Paper No. A-17, Mathematica Policy Research, Princeton, N.J., 1977.

27. Uninsured medical expenses above \$60 per year for an individual were subtracted from family earnings before calculating the NIT payment.

28. Kehrer and Wolin, op. cit.

DISCUSSION

[Charles M. Wolin presented the paper at the conference.]

JAMES WALSH: How many births did you have in the Gary experiment?

WOLIN: We identified a total of about 550 births, but rejected about 150 observations for a number of reasons--missing data on birth certificates, missing data in our own data files.

WALSH: Is that significant that you dropped these observations?

WOLIN: It may be. A larger sample might have improved the precision of our estimates.

J. M. ANDERSON: Will SIME/DIME help?

WOLIN: Yes, because the number of families in Gary was 1799; in SIME/DIME it's about 5000. Also, Gary was a three-year experiment, and we have both a three- and five-year component to SIME/DIME.

WENDY HOLDEN: Did you use alcohol consumption as a variable?

WOLIN: We would have if we could have, but we didn't have any data on alcohol consumption.

PATSY CARTER: What about nutritional patterns?

WOLIN: We do not have any nutrition data.

CARTER: Is there a chance of picking that up in SIME/DIME?

WOLIN: No.

PEGGY THOITS: The rural income maintenance experiment did ask about nutrition and found a significant effect on nutrition. With higher income, nutritional habits were better and people were eating better.

IMPACTS OF THE SEATTLE AND DENVER
INCOME MAINTENANCE EXPERIMENTS
UPON PSYCHOLOGICAL DISTRESS

by

Peggy Thoits
Sociologist
SRI International

A good deal of literature over the past 50 years has demonstrated an inverse relationship between income and psychological distress: the lower his or her income, the more likely the individual is to report high levels of psychological distress.¹ However, the studies documenting this relationship are cross-sectional in nature. Little is known about the effect that changes in income have upon psychological state. The income maintenance experiments offer a valuable opportunity to examine the effect of an improvement in financial circumstances upon psychological well-being.

The mental health literature offers several explanations of the relationship between income and psychological distress. Each has different implications for predicting the effect of an income maintenance program upon distress. Two of these theories are not considered in this paper, the drift hypothesis and the social mobility hypothesis. The drift hypothesis² suggests that individuals have "drifted downward" into a low-status or low-income position because their mental illness has impeded effective social interaction or job functioning. This theory applies only to cases of severe psychological impairment, or psychosis. Since the index of psychological distress employed in this report does not measure psychotic

symptoms, a test of the drift hypothesis is not possible.

The social mobility hypothesis³ suggests that a change in socioeconomic status may produce strain in the individual and thus result in increased psychological distress. But since experimental subjects do not receive enough income from the program to substantially alter their socioeconomic status, a test of this hypothesis is again not possible.

Two other hypotheses suggested by the mental health literature can be tested with the data from the income maintenance experiments. One is the "resource hypothesis" and the other is the "life events hypothesis."

According to the resource hypothesis, the low-income individual lacks sufficient resources to resolve the ordinary life crises which seem especially common in the lower class (for example, unexpected car repairs, job lay-offs, accidents and illnesses).⁴ Unresolved problems may overwhelm the individual's capacity to cope and produce symptoms of psychological distress. This reasoning suggests that an increase in or stabilization of income will facilitate coping with the exigencies of living and thereby reduce levels of psychological distress. Therefore, an income maintenance treatment will significantly decrease the psychological distress of financial subjects relative to control subjects.

The life events hypothesis implies an opposite effect. Life events are "objective events that disrupt or threaten to disrupt the individual's usual activities."⁵ These include death of spouse, start of a job, birth of a child, and so forth. The more life events a person experiences, the more his or her coping abilities are taxed, and the more psychological distress he/she will exhibit or report.⁶ The change in financial circumstances due

to the experiment may be a life event in itself. Furthermore, it has been shown that the experiment significantly increases the divorce rate of those receiving financial treatment, increases their rates of unemployment, and increases rates of migration. Because the experiment causes more life events to occur, these results imply that the experiment may significantly increase psychological distress, rather than decrease it.

Finally it is possible that alterations in the standard of living due to the experiment are so slight as to have no effects upon psychological distress.

What are the impacts of the Seattle and Denver Income Maintenance Experiments (SIME/DIME) upon psychological distress? Following a brief description of the methodology used in SIME/DIME psychological distress study, some of the results will be presented in an effort to answer that question.

RESEARCH METHODOLOGY

The psychological distress index used in SIME/DIME is a close variant of the MacMillan Health Opinion Survey index.⁷ It consists of a series of psychophysiological symptom items (listed in Table 1), and identifies individuals whose psychological state impairs their everyday functioning to some degree.

In general, the psychological distress index was administered at the first and fifth periodic interviews for men and at the second and sixth periodic interviews for women.⁸ This means that for men, the first measure was taken approximately four months after enrollment in the experiment and

TABLE 1

PSYCHOLOGICAL DISTRESS INDEX

	<u>Never</u>	<u>Seldom</u>	<u>Some- times</u>	<u>Often</u>
*1. How often do your hands tremble enough to bother you?	1	2	3	4
*2. How often do you smoke?	1	2	3	4
*3. How often do your hands or feet sweat so that they feel damp and clammy?	1	2	3	4
*4. How often are you bothered by your heart beating hard?	1	2	3	4
*5. How often do you have cold sweats?	1	2	3	4
*6. How often do you feel that you have several different ailments in different parts of your body?	1	2	3	4
*7. How often do you lose your appetite?	1	2	3	4
*8. How often has ill health affected the amount of work you do?	1	2	3	4
*9. How often do you have weak spells?	1	2	3	4
*10. How often do you have spells of dizziness?	1	2	3	4
*11. How often do you tend to lose weight when important things are bothering you?	1	2	3	4
*12. How often are you bothered by nervousness?	1	2	3	4
*13. How often have you been bothered by shortness of breath when you were not exercising?	1	2	3	4
*14. How often do you tend to feel tired in the mornings?	1	2	3	4
*15. How often do you have trouble getting to sleep and staying asleep?	1	2	3	4
	<u>Never</u>	<u>Not Very Much</u>	<u>Pretty Often</u>	<u>Nearly All the Time</u>
*16. How often are you bothered by having an upset stomach?	1	2	3	4
	<u>Never</u>	<u>A Few Times</u>	<u>Many Times</u>	
*17. How often have you been bothered by nightmares or dreams which frighten or upset you?	1	2	3	
	<u>Very Good Spirits</u>	<u>Good Spirits</u>	<u>Low Spirits</u>	<u>Very Low Spirits</u>
18. In general, would you say that most of the time you are in:	1	2	3	4

* Item is similar to MacMillan Health Opinion Survey item (MacMillan, 1957).

the second was taken approximately twenty months into the experiment. For women, the first measure was taken eight months after enrollment and the second measure, twenty-four months after enrollment. All analysis reported in this paper is based upon measures taken at the second observation.

Analysis Strategy

The intent of this research is to produce baseline findings that will become the focus of future work. Therefore, in the analysis, I attempt to estimate the effects of the experimental financial treatments upon distress scores, controlling for the effects of variables that determine assignment to treatments and for other variables (such as age, education, sex) that are known to affect distress. This strategy relies heavily on the experimental nature of the study. Controlling for the assignment variables, the control group and the various experimental treatment groups should not differ systematically on initial levels of distress. If they differ later in the experiment, the difference can be attributed to the experimental treatments.

Married men, married women, and single women were analyzed separately by site because we found that these groups cannot be pooled for statistical analysis.⁹

In each analysis, we controlled for the same set of nonexperimental variables or what I call "the background function." As mentioned above, the background function contains all those variables that were used in assigning a family to experimental treatments (site, race-ethnicity, normal family income, marital status) and several other variables related to distress.¹⁰

Treatment categories are also included to allow the effect of the experiment to vary by race, ethnicity, and the length of the experiment.

They are:

- a. F3*White: White three-year financial subject
- b. F5*White: White five-year financial subject
- c. F3*Black: Black three-year financial subject
- d. F5*Black: Black five-year financial subject
- e. F3*Chicano: Chicano three-year financial subject
- f. F5*Chicano: Chicano five-year financial subject

Each of these treatment categories and a measure of what the income maintenance "payment" would be if the family did not change their pre-enrollment behavior are examined for their effect on distress.¹¹

RESULTS

Table 2 reports the mean distress scores of the control and experimental groups, recorded at the second observation point. In every case, the mean scores of experimental subjects are somewhat higher than the comparable control subjects. The differences that are significant can be identified in Table 3. The estimates shown in Table 3 are the number of points an individual's distress score is raised or lowered due to the experimental treatment. For example, the distress score of a white husband on the five-year treatment in Denver is 3.22 points higher than the score of a comparable control subject. Another example: the distress score of a white husband on the five-year treatment in Seattle is 1.04 points lower than that of a comparable control subject.

TABLE 2
 MEAN DISTRESS SCORES AT OBSERVATION 2

	<u>MARRIED MEN</u>			
	<u>DENVER</u>		<u>SEATTLE</u>	
	<u>Controls</u>	<u>Experimentals</u>	<u>Controls</u>	<u>Experimentals</u>
Distress score	28.0	28.6	27.5	28.4
Number of cases	624	840	506	564

	<u>MARRIED WOMEN</u>			
	<u>DENVER</u>		<u>SEATTLE</u>	
	<u>Controls</u>	<u>Experimentals</u>	<u>Controls</u>	<u>Experimentals</u>
Distress score	30.0	30.6	30.2	31.2
Number of cases	547	734	394	443

	<u>SINGLE WOMEN</u>			
	<u>DENVER</u>		<u>SEATTLE</u>	
	<u>Controls</u>	<u>Experimentals</u>	<u>Controls</u>	<u>Experimentals</u>
Distress score	30.0	31.2	30.8	32.5
Number of cases	339	589	270	377

TABLE 3

EFFECTS OF THE THREE-AND FIVE-YEAR TREATMENTS
AND PAYMENT PSYCHOLOGICAL DISTRESS^a AT TIME 2,^b
HOLDING BACKGROUND VARIABLES CONSTANT

	Married Men ^c	
	Denver	Seattle
F3*White	.08	-.53
F5*White	3.22***	-1.04
F3*Black	-.03	2.17**
F5*Black	.14	-1.12
F3*Chicano	1.04	--
F5*Chicano	.48	--
Payment (in thousands of dollars)	-.06	.36
R ²	.06	.13
F-test for set of treatment variables	1.93*	2.39**
N	1228	793

	Married Women ^c	
	Denver	Seattle
F3*White	.63	1.19
F5*White	2.42**	1.16
F3*Black	-.33	.05
F5*Black	2.91**	-1.03
F3*Chicano	.95	--
F5*Chicano	2.06*	--
Payment (in thousands of dollars)	-.37	.12
R ²	.05	.08
F-test for set of treatment variables	2.02*	.87
N	1251	820

	Single Women ^c	
	Denver	Seattle
F3*White	.19	-.79
F5*White	-.51	1.51
F3*Black	2.28**	2.42*
F5*Black	2.69**	-.68
F3*Chicano	-1.10	--
F5*Chicano	.32	--
Payment (in thousands of dollars)	.08	.46
R ²	.04	.05
F-test for set of treatment variables	1.15	2.27**
N	889	615

* $.10 \geq p > .05$; ** $.05 \geq p > .01$; *** $.01 \geq p$

^aA positive regression coefficient indicates increased distress; a negative coefficient indicates decreased distress.

^c"Married" refers to married or cohabiting individuals. "Single" refers to divorced, separated or widowed persons.

We see that for most sex, marital status, and racial-ethnic groups the experimental treatments have no significant effects upon psychological distress. However, eight groups receiving financial treatment show a significant rise in distress level (at least two points over controls) in response to the experiment: white husbands on the five year treatment in Denver; black husbands on the three-year treatment in Seattle; black, white, and Chicana wives on the five-year treatment in Denver; black singles on both the three- and five-year treatments in Denver; and black singles on the three-year treatment in Seattle.

The Life Events Hypothesis and the SIME/DIME Results

The results from SIME/DIME are consistent with the life events hypothesis, which predicts an increase in distress due to a change in financial circumstances.¹² Participation in an NIT program significantly raises the distress levels of some experimental subjects relative to controls. However, we do not know whether the experiment affects distress directly or does so indirectly through its impact upon the occurrence of other major life events such as marital dissolution, unemployment, and migration. If I control for the effects of these and other life events and the significant experimental impacts disappear, I infer that the impact of the experiment is indirect through these intervening events. On the other hand, if they do not, I infer that participation in the experiment is a stressful life event in itself.

When the other life events (listed in Table 4) are controlled for, two of the eight originally significant estimates become nonsignificant. It can be said then that intervening life events account for the effect in those two groups--white wives and black singles on the five-year treatment in

TABLE 4

LIFE EVENTS* EMPLOYED IN THE ANALYSIS

Family Gain Events

Reconciliation/remarriage
 Pregnancy (or spouse pregnant)
 Birth of child
 Child or children arrive

Family Loss Events

Separation/divorce
 Child or children leave
 Other family member leaves
 Major residential move
 Spouse suffers health problems

Occupational Gain Events

Begin employment
 Begin school/training
 Increase in occupational status
 Spouse begins employment
 Increase in spouse's occupational status
 Family income increases 50%

Occupational Loss Events

Employment ends
 End school/training
 Decrease in occupational status
 Spouse's employment ends
 Decrease in spouse's occupational status
 Family income decreases 50%

* Death of spouse, death of child or other family member, retirement and institutionalization are other events for which data exist. However, these events are so rare that they are not included in the analysis for Table 5.

Denver. The impact of the experiment itself remains significant in six groups--four of which are black (see Table 5). In each case, the financial subjects' distress scores were raised two points or more above those of the controls. It is clear that for certain racial-ethnic and marital status groups, participation in an income maintenance experiment itself appears to be a somewhat distressing life change event.

CONCLUSION AND FUTURE RESEARCH

Given the considerations that the NIT support level guarantees are not high and the predicted payment to experimental families averages only \$1500 per year, it is surprising to find that the experiment significantly raises the distress levels of several groups, even after the effects of intervening life changes have been controlled. As the effect is evident approximately two years after enrollment, it does not appear to be a short-run reaction to a change in financial circumstances.

Why should an increase in income produce distress? Three possibilities occur to me. First, it may be that the experiment raises expectations but fails to provide enough resources for hoped-for improvements and only stimulates additional needs.

Second, the experiment may decrease wives' financial dependence upon their husbands, by providing an alternative source of economic support without the stigma attached to welfare. The wife's increased independence may stimulate marital conflict which in turn may increase distress.

Finally, there is some evidence that sharing whatever one has with relatives and friends is a norm in the lower class, and appears to be

TABLE 5

EFFECTS OF EXPERIMENTAL TREATMENTS
UPON PSYCHOLOGICAL DISTRESS^a AT TIME 2^b WHEN BACKGROUND
VARIABLES AND INTERVENING LIFE EVENTS ARE CONTROLLED

	Married Men ^c	
	Denver	Seattle
F3*White	.34	-.26
F5*White	2.52**	-.49
F3*Black	-.12	2.51**
F5*Black	-.003	.03
F3*Chicano	1.00	
F5*Chicano	.45	
Payment (in thousands of dollars)	-.16	.03

	Married Women ^c	
	Denver	Seattle
F3*White	.62	.50
F5*White	1.47	.98
F3*Black	-.29	.17
F5*Black	2.77**	-.98
F3*Chicano	.70	
F5*Chicano	2.08*	
Payment (in thousands of dollars)	-.25	.22

	Single Women ^c	
	Denver	Seattle
F3*White	.78	-.75
F5*White	.66	1.10
F3*Black	2.28**	2.57**
F5*Black	1.27	-.15
F3*Chicano	-1.02	
F5*Chicano	.15	
Payment (in thousands of dollars)	.20	.23

* .10 ≥ p > .05 ** .05 ≥ p > .01 *** .01 ≥ p

^a A positive regression coefficient indicates increased distress; a negative coefficient indicates decreased distress.

^b Time 2 occurs roughly 20 months after enrollment for males and approximately 24 months after enrollment for females.

^c "Married" refers to married or cohabiting individuals. "Single" indicates divorced, separated, or widowed persons.

especially true of black and Spanish-speaking groups.¹³ If this is true, it may be that experimental families experience increased demands upon their financial resources from relatives and friends, frustrating any hopes of getting ahead. Our finding that black groups more often have significantly increased distress is consistent with this. Although we have no direct way to measure the tendency to share resources with others, we can go back through the data and find the number of unrelated individuals and extended family members that join the experimental families and how long they stay. If more persons join experimental households and/or stay longer than in control households, we will have found some support for this hypothesis. This and the other two possibilities above will be part of the future research on psychological distress using SIME/DIME data.

At this point it is impossible to estimate the costs to society of the somewhat distressing effects of an income maintenance program. In the future, I hope to obtain funding to collect hospital and mental health agency records to assess whether stress caused by the experiment results in increased utilization of psychiatric facilities.

NOTES

1. See Robert J. Kleiner and Seymour Parker, "Social Structure and Psychological Factors in Mental Disorder: A Research Review," in H. Wechsler, L. Solomon and B. M. Kramer (eds.), Social Psychology and Mental Health, (New York: Holt, Rinehart and Winston, 1970); Norman M. Bradburn, The Structure of Psychological Well-Being (Chicago: Aldine, 1969); Norman M. Bradburn and David Caplovitz, Reports on Happiness (Chicago: Aldine, 1965); Derek L. Phillips, "The 'True Prevalence' of Mental Illness in a New England State," Community Mental Health Journal 2 (1966): pp. 35-40; G. Gurin, J. Veroff and S. Feld, Americans View Their Mental Health (New York: Basic Books, 1960); L. Strole, T. S. Langner, S. T. Michael, M. K. Opler and T. A. C. Rennie, Mental Health in the Metropolis: The Midtown Manhattan Study (New York: McGraw Hill, 1962); B. P. Dohrenwend and B. S. Dohrenwend, Social Status and Psychological Disorder: A Causal Inquiry (New York: John Wiley and Sons, 1969); Dorothea Leighton, J. S. Harding, D. B. Macklin, A. M. MacMillan and A. H. Leighton, The Character of Danger: Psychiatric Symptoms in Selected Communities (New York: Basic Books, 1963); "Selected Symptoms of Psychological Distress," Public Health Service Publication No. 1000, Series 11, no. 37, Washington, D.C.: U.S. Government Printing Office, 1970; and George Warheit, Charles Holzer, and John Schwab, "An Analysis of Social Class and Racial Differences in Depressive Symptomatology: A Community Study," Journal of Health and Social Behavior 14 (1973): pp. 291-299.
2. See H. W. Dunham, Community and Schizophrenia, (Detroit: Wayne State University Press, 1965).
3. See reviews by Robert J. Kleiner and Seymour Parker, supra (1970), and B. P. Dohrenwend and B. S. Dohrenwend, supra (1969).
4. See Dohrenwend, et al., supra (1969) and Melvin L. Kohn, "Social Class and Schizophrenia: A Critical Review" in D. Rosenthal and S. Kety (eds.), The Transmission of Schizophrenia (Oxford: Pergammon Press, 1968).
5. See Dohrenwend, et al., at p. 133.
6. E. S. Paykel, "Recent Life Events and Clinical Depression" in E. K. E. Gunderson and R. H. Rahe (eds.), Life Stress and Illness (Springfield, Ill.: Charles C. Thomas, 1974): pp. 134-163; Jerome Myers, J. J. Lindenthal and Max Pepper, "Life Events and Psychiatric Impairment," Journal of Nervous and Mental Disease 152 (1971): pp. 149-157; Myers, et al., "Life Events and Mental Status: A Longitudinal Study," Journal of Health and Social Behavior 13: pp. 398-406; Myers, et al., "Social Class, Life Events and Psychiatric Symptoms: A Longitudinal Study," in B. S. Dohrenwend and B. P. Dohrenwend (eds.) Stressful Life Events: Their Nature and Effects (New York: Wiley, 1974); and G. W. Brown and J. L. T. Birley, "Crises and Life Changes and the Onset of Schizophrenia," Journal of Health and Social Behavior 9 (1968): pp. 203-214.

7. A. M. MacMillan, "The Health Opinion Survey: Technique for Estimating Prevalence of Psychoneurotic and Related Types of Disorder in Communities," Psychological Reports 3: pp. 325-339.

8. Seattle single women actually answered the distress questions at the first periodic. They were then interviewed at the sixth periodic with other women.

9. See "Income and Psychological Distress: The Impact of the Seattle and Denver Income Maintenance Experiments" (Research Memorandum 50, Center for the Study of Welfare Policy, Stanford Research Institute, Menlo Park, California, 1978).

10. In all, ten background variables were included: (1) age in years; (2) black; (3) Chicano; (4) normal family income (expected income of the family in the year prior to the experiment, assuming normal family circumstances and adjusted for family size); (5) years of formal schooling; (6) whether the family has one or more children; (7) occupational status; (8) spouse's occupational status (for those married at enrollment); (9) working at enrollment; and (10) spouse working at enrollment (for those married at enrollment).

11. Analysis not reported here indicates that replacing these treatment variables with a larger set of support level and tax-rate dummies does not explain additional variance in distress.

12. Note that the magnitude of the financial change, measured by the payment variable, is less important than participation in the program itself.

13. See Lloyd H. Rogler and August G. Hollingshead, Trapped: Families and Schizophrenia (New York: Wiley, 1965); Oscar Lewis, Five Families (New York: Basic Books, 1959); Oscar Lewis, The Children of Sanchez (New York: Random House, 1961); Oscar Lewis, La Vida (New York: Random House, 1965); and Ulf Hannerz, Soulside: Inquires Into Ghetto Culture and Community (New York: Columbia University Press, 1969).

DISCUSSION

WENDY HOLDEN: What about the differences in the sites, the fact that there was significant distress in one population in Denver and not in the same population in Seattle?

THOITS: The only thing I can come up with is the employment situation in the two cities. Seattle is a highly depressed area, and Denver is a growing area. That may have something to do with the different distress responses.

CHUCK FROLAND: The controls should take care of that.

THOITS: That's true in part. I don't yet understand why there is a differential experimental effect by site.

FROLAND: If I understand you correctly, you entered in events as dummy variables which would seem to overlook the total number of stressful events.

THOITS: I've done the analysis both ways. And, in fact, it accounts for more variance if you enter them as dummy variables. The reason for not summing them up is this: As I mentioned before, the gain and loss events are not always gains and losses. Some gains can be distressing and some losses could be distress alleviating. To sum them up is to confound the effects.

HOLDEN: Holmes didn't make that distinction. Holmes created all distressful events as distressful events, positive or negative.

THOITS: There is a lot that I think is wrong with the Holmes-Rahe study. First, in general positive events decrease distress and negative events increase distress. Many events on the Holmes-Rahe scale are ambiguous. For example, one of his items is "changing work habits," and you don't know if it is a positive or a negative event. Second, if you get rid of the ambiguous events--those which aren't clearly culturally desirable or undesirable--there's a larger number of negative events on the scale. Thus, the relationship between total change and stress or total change and illness may be primarily due to the negative events on that scale. Third, there are only 43 events, including Christmas, on the Holmes-Rahe scale. There are many life events that we all agree on as major happenings that are not represented on those scales, such as infidelity of a spouse, victimization in a burglary or a major felony, and experiencing an abortion.

J. M. ANDERSON: In your overall study in stress, was suicide used as an indicator?

THOITS: There are very few deaths during the experiment, and we don't know whether they were suicides or not; we don't know about suicide attempts. We were developing a questionnaire for inclusion in one of the later interviews, and it was deemed too sensitive.

FROLAND: Just because there's more stress, doesn't mean people are worse off.

THOITS: It's clear that people do make the choice to divorce, which can be an extremely unpleasant experience for many people, and yet they make that choice consciously. It is a decision for a short run distressing event in hopes for a better life style.

FROLAND: If the differences in the NIT benefits didn't make a difference, and the life events didn't make a difference for most of the groups, what did make the difference?

THOITS: At this point I don't know what it is about the experimental treatments that caused increases in distress.

HOLDEN: Do you ask the people if they think they're better off or worse off?

THOITS: In the last interview, the people will be asked how they think income maintenance has affected their lives.

SIME/DIME HOUSING STUDY:
POLICY QUESTIONS AND PROGRESS SUMMARY

by

Cynthia Thomas
Project Director, Housing Study, MPR

The Seattle and Denver Income Maintenance Experiment (SIME/DIME) housing study seeks to determine the effects of income maintenance support payments on the quality of housing of individuals and families in the experiment. It was funded in 1975 by the Department of Health, Education and Welfare (DHEW) with a contribution from the Department of Housing and Urban Development (HUD). The results of this study, when compared with the findings from HUD's Experimental Housing Allowance Program (EHAP), will enable policymakers to assess the relative advantages of the various approaches for improving housing conditions for low income families.

Results from the SIME/DIME Housing Study and from related portions of EHAP are expected to be available in late fall 1978. This paper discusses the relationship between the SIME/DIME work and EHAP, plans for completing the SIME/DIME analysis, and preliminary findings from EHAP.¹

SIME/DIME HOUSING STUDY

In the SIME/DIME housing study, experimental and control families in Denver and in Seattle were administered a supplemental housing survey and dwelling assessment form to obtain information on the housing occupied at the years of participation in Seattle. Data from the SIME/DIME periodic interviews also will be used so that factors relating to housing outcomes

such as change in marital status or change in disposable income can be evaluated. Data from the early SIME/DIME instruments will be used to assess whether experimental and control households initially had the same levels of housing quality according to their rent levels and the presence of certain characteristics of the housing.

THE HUD EXPERIMENTAL HOUSING ALLOWANCE PROGRAM (EHAP)

Dissatisfaction with existing housing programs led Federal officials in the late 1960's to resurrect an idea that had originated in the 1930's --that of providing subsidies to enable low income renters or homeowners to select housing within neighborhoods of their own choice. In the early 1970's, HUD began evaluating the feasibility of these housing allowances, testing three major experimental components. A "demand experiment" is being used to determine how recipients respond to alternative payment formulas; a "supply experiment" measures the potential effects of a full scale subsidy program on rental and purchase prices and on construction and rehabilitation of housing; and an "administrative agency experiment" tests various approaches to program management. EHAP also offers an integrated analysis, combining results from each component to permit broader applications, such as estimates of likely costs of a national program.

Payments in the demand experiment are structured either to encourage households to obtain housing that meets minimum standards or to require them to do so. It is the component most directly comparable to SIME/DIME, as it both allows participants to make choices in spending a subsidy and then observe these choices. Instruments for the SIME/DIME housing study were designed to permit precise comparisons.

Three payment strategies are used. One, a "housing gap plan", requires households to spend a reasonable proportion of their income on housing and subsidizes the gap between the amount they pay and the cost of adequate housing. Families are expected to obtain housing that meets minimum levels of quality. This strategy resembles current housing programs, where households are required to contribute around 25% of their income toward the housing unit they obtain. A second strategy, a "percentage of rent plan", provides families with a subsidy equivalent to a predetermined proportion of their rent. This form of payment is designed to encourage households to spend more on housing than they might otherwise spend in order to obtain a larger subsidy; they are not formally required to obtain housing that meets specific requirements or to spend a minimum amount. A third provides a limited number of households with unrestricted cash payments similar to income maintenance payments. These payments are not structured, either explicitly or implicitly, to encourage households to obtain housing that meets minimum standards. All three Demand plans allow recipients the flexibility to select special attributes in their housing, including the locations of their dwellings.

POLICY QUESTIONS AND RESULTS FROM EHAP

SIME/DIME looks at seven questions about the housing occupied by individuals and families in the program. Four of the questions (1, 2, 3 and 5), are similar to those in the Demand Experiment of EHAP:

- (1) How have expenditures for housing increased or decreased during the experiments? To what extent are increases or decreases associated with participation in the experimental programs?
- (2) How have the experiments affected levels of housing consumption? To what extent has the proportion of families living in housing which is "standard" by commonly accepted definitions been affected

by the experiments?

- (3) To what extent were changes in housing quality induced by the experiments made by upgrading original dwelling units rather than by moving? What factors led to the decision to upgrade rather than to move?
- (4) How have choices between renting and owning been affected by the experiments?
- (5) To what extent have the experiments affected patterns of location? Are there observable effects on neighborhood quality? On distances from centers of work or from concentrations of various racial and income groups?
- (6) Are income maintenance payments treated the same way as income from other sources in housing consumption decisions?
- (7) To what extent has the duration of the experimental transfer program affected housing consumption?

Increased Expenditures for Housing

The first question concerns the rates of increase in expenditures for housing during the experiments. One would expect greater increases in housing expenditures under a housing allowance program than under an income maintenance program where the housing allowance payment formulas require or encourage minimum standards for consumption.

Both EHAP and SIME/DIME focus on the magnitude of differences between the experimental and control samples. Both experiments obtained similar data on rental costs two years after enrollment, including the information on such components of rent as furnishings and utilities. The data, however, were collected one year later at the SIME/DIME sites than at the Demand sites. In addition to nationwide inflation, which will produce different rental values, there are site specific factors associated with rental costs that will complicate comparisons between the two experiments.

In the Demand Experiment, one-third of households under the "housing gap" payment plan were already living in acceptable housing at enrollment and were spending 44% of their income on housing. These households could either spend the subsidy on additional housing or reduce the proportion of their income spent for housing by purchasing other goods and services. Generally they chose to spend the payment on commodities other than housing. The other two-thirds of "housing gap" households were required to improve their dwellings (by moving or upgrading) before they were eligible for payments. Half of this group (about one-third the total) had not done so by the end of the first year. The others increased housing expenditures, after accounting for inflation, by about 20%.

Households under the "percentage of rent" plan increased rental expenditures by approximately 2 to 16%, depending on the proportion of the rental payment that was subsidized. Although analysis of this result and results from the unconstrained payment plan are still in progress, it appears that households under the "percentage of rent" plan increased their housing expenditures by about three times the amount of the increases of households receiving an income maintenance payment. When all the data is in, we will determine whether this differential continues to exist after two years, and whether SIME/DIME results will support this finding.

Increases in Housing Quality

Housing expenditures do not necessarily indicate housing quality, for several reasons. First, a dwelling that has just been rented may cost more than a similar one occupied for a long time, because price adjustments are

more frequently made during turnovers. Second, some shoppers are more clever than others and get standard or better housing at lower prices. Thus, both experiments use three measures of housing quality: (1) a measure based on minimum quality standards; (2) an index providing continuous values of housing quality and incorporating a broad set of housing characteristics, permitting subtle distinctions between different dwellings; and (3) a measure based on families' own perceptions about their dwellings, including their satisfactions with certain features and their assessments of the condition of various elements, permitting assessment of characteristics more likely to be observed only after living in the dwelling for a period of time.

First year data from the Demand Experiment show that approximately 18% of participating households obtained housing that met minimum standards who otherwise would not have done so. An important question to be answered from the SIME/DIME data is how likely experimental participants are to obtain suitable and decent housing when it is not required. It will be important to evaluate the overall differences in quality improvements for control and experimental individuals and families in each experiment.

Method of Upgrading

When upgrading housing, households in the Demand Experiment have two options. They can move to better dwellings, or they can obtain repairs or improvements on their current dwellings. From a policy perspective, there is no strong reason to prefer one over the other. The preference to upgrade, however, could alter the characteristics of the housing stock in a nationwide program--perhaps reducing the rate of deterioration of older dwelling units.

Preliminary results suggest that, although Demand households prefer not to move when their dwellings satisfy program requirements, those seeking to qualify for payments are more likely to move than to upgrade current dwellings. Half of the households moving in the first year were able to meet standards for obtaining subsidies while only about 20% of non-movers upgraded their dwellings and became eligible for payments. We plan to obtain similar analysis of SIME/DIME households. Without the need to meet minimum housing requirements, SIME/DIME households may be more likely than Demand households to move for reasons other than to improve housing quality. Consequently, it will be necessary to distinguish upgrading from non-upgrading moves.

Renting Versus Buying

Home ownership is an important middle-class objective in the United States, not only because of intrinsic values associated with owning a home but also because of the potential tax advantages. The major thrust of EHAP has been to determine the effects of a housing allowance on the behavior of renters. SIME/DIME households can be studied as to how they used their experimental payments to change their tenure from renting to owning, although it will not be possible to compare this finding with results from the Demand Experiment.

Patterns of Moves

Family moves to satisfy housing requirements could substantially change the composition of neighborhoods by, for example, dispersing minorities and low income families throughout an urban area. If so, some greater social and economic integration could be achieved without intervention by public officials.

Analysis thus far does not indicate that the housing allowance program produces significant changes in the composition of neighborhoods. Both

experimental and control families were only somewhat likely to move to neighborhoods with smaller proportions of low-income residents than their original neighborhoods. Black experimental households may be somewhat more likely than black controls to move to neighborhoods with lower proportions of minorities, but evidence on this point is incomplete.

Both upgrading and non-upgrading moves can be taken into account for SIME/DIME in assessing whether households move into neighborhoods with smaller concentrations of low income and minority households. There is no apparent reason to expect mobility patterns to be different under a negative income tax (NIT) program than under a housing allowance subsidy, although different initial concentrations of ethnic and income groups in Seattle and Denver may result in different outcomes.

Impact of Payment Duration

The answer to the sixth question, concerning the impact of the duration of the SIME/DIME experiments on housing consumption, has implications for assessing the other five questions. Presumably if the program is viewed by recipients as short-term, they will be less likely to make substantial commitments to improve their housing. Renters who wish to purchase housing may be less likely to do so under a three-five year program than they would under a long-term program. The analysis should therefore determine whether results underestimate potential levels of housing consumption.

COST EFFECTIVENESS

An important question for policymakers at the conclusion of the SIME/DIME and EHAP studies is the relative cost effectiveness of the two approaches to

raising minimum levels of housing quality among low-income households. It is likely that housing allowances will produce greater overall changes in housing quality. There are, however, significant costs to administering a program that either requires the inspection of dwellings and/or certification that participants pay some minimum rent. Perhaps households responding to an NIT approach will be able to achieve satisfactory levels of housing quality without those constraints.

An NIT might be able to reduce the rent burden on low-income households to levels closer to 25% of income.

These and other important evaluations will take place when the results from the two sets of experiments are available and comparisons can be made from them.

NOTES

1. First year results from the Demand Experiment are described in Helen E. Bakeman, et al. "Housing Allowance Demand Experiment: Fourth Annual Report" (Cambridge, Massachusetts, Abt Associates, Inc., December 1977); David B. Carlson and John D. Heinberg, "How Housing Allowances Work: Integrated Findings from the Experimental Housing Allowance Program" (Washington, D.C., The Urban Institute, February, 1978); and James Wallace, "Housing Allowance Demand Experiment: Preliminary Findings" (Cambridge, Massachusetts, Abt Associates, inc., March 1978).
2. APHA-PHS Recommended Housing Maintenance and Occupancy Ordinance, Public Health Service Publication No. 1935, Washington, D.C. 1969.

DISCUSSION

Gladys McCorkhill: Is the amount of the HUD subsidy in the Experimental Housing Allowance Program related in any way to the inflation rate, as were SIME/DIME payments? If not, how do they adjust for inflation?

THOMAS: Under a housing gap payment plan, subsidies are not adjusted for inflation. Under a percentage of rent plan, they are adjusted because of the nature of the payment formula. Subsidies to households under a housing gap plan are the difference between the estimated cost of standard housing in an area (C) and some percentage of households' income (bY) where $\text{Payment} = C - bY$. The value of C was established only at the beginning of the experiment and was not adjusted to take account of changing housing costs. Under the percentage of rent plan, however, the Payment is equal to a fraction of the rent (aR) where $\text{Payment} = aR$. Although the fraction "a" does not change during the experiment, if rent (R) increases, so does the payment. This question has interesting implications for evaluating the effects of the two alternative payment formulas, especially rates of participation and quality of housing obtained under the two plans.

THE EFFECT OF INCOME MAINTENANCE
ON THE UTILIZATION OF SUBSIDIZED HOUSING

by

Marcy Avrin
Consultant
SRI International

Housing subsidies are a form of economic assistance provided to the producers and consumers of housing for the purpose of lowering the price or costs of housing. This is expected to increase the supply and use of decent quality housing by American households and to help achieve the national housing goal of "a decent home and a suitable living environment for every American family," as stated by Congress in the Housing Act of 1968. The main objective of this paper is to estimate the general effects that an income maintenance program could have on the utilization of subsidized housing. This is of critical importance as the national government revamps the welfare system and also attempts to meet the demand of many cities for more housing aid.

From the data gathered in the Seattle/Denver Income Maintenance Experiment (SIME/DIME), we have attempted to answer three main questions: (1) Does a negative income tax (NIT) cause families to move out of subsidized housing? (2) Does it decrease the probability of moving into subsidized housing? and (3) Do the effects vary with the level of support and family size?

INTERFACE OF SIME/DIME AND SUBSIDIZED HOUSING

In calculating the NIT payment in SIME/DIME, certain work and household expenses are deducted from gross income. This is an attempt to eliminate

the influence of transfer programs on the outcome of the experiments by fully "taxing" them. Although subsidized housing is a public transfer, it is an exception. Because it is an "in-kind" transfer, its value is difficult to determine. Changes in the amount of the subsidy during the experiment are also not considered in the determination of the NIT payment.¹ (In contrast, subsidized housing programs treat SIME/DIME income as any other income in determining both eligibility and the amount of the subsidy.²)

All housing programs base benefits on income and family size. Generally, families pay either 25% of their adjusted gross income or the fair market rent, whichever is less, for the housing.³ The majority of subsidized housing units are either Public Housing⁴ or "Section 236" Rental Housing.⁵ In 1973, the Denver Metropolitan Area had close to 7000 units for low-income families (non-elderly) in these two programs and Seattle had 8500 units. In addition, some units were available under Section 101 Rent Supplements to eligible families and individuals residing in multi-family housing whose owners have rent supplement contracts with the Department of Housing and Urban Development (HUD).

SIME/DIME did not obtain data about the type of subsidized housing in which the families are living. This lack of information is unfortunate because the quality of the housing may differ by the type of program. According to interviewers, the vast majority of families living in subsidized housing lived in Public Housing, as opposed to 236 Housing.

They reported that the subsidized housing in Denver is in better condition than that in Seattle. Denver's is described as one or two stories and well-maintained with considerable noise and no grounds. Denver families generally do not like living in the housing projects and the leased housing

is considered more desirable. Seattle's housing is also low-rise, but older and deteriorating. Some units are presently being remodeled and recreational grounds are being built. The housing projects have problems with insects, crime and fear. One controversial 236 development is now partly abandoned. The analysis takes these differences into account by looking at Seattle and Denver separately.

ANALYSIS

We are concerned with two effects in this study. The first is the overall effect of the NIT on the use of subsidized housing. We analyzed this by looking at data from month 18 of the experiment. We used standard statistical techniques to determine any treatment effect on the probability that a family resides in subsidized housing during month 18 after it enrolled in the experiment.

Second, we looked at the NIT effect on moves into and out of subsidized housing.⁷ This was done by looking separately at those families who lived in subsidized housing at enrollment separately from those who did not.⁸ We studied those who did not in an attempt to determine any NIT effect on the probability that they would move into a subsidized unit. Potentially, the NIT could affect moves both into and out of subsidized housing and it is important to know the probability of each.

We used the above two approaches to look at data from three sample populations taken from those enrolled in SIME/DIME. They were: (1) all originally enrolled dual-headed families who remained stable in terms of head of household through the first eighteen months of the experiment; (2) all originally enrolled single parent families headed by men that remained stable

in terms of head of household through the first eighteen months of the experiment; and (3) all originally enrolled single parent families headed by women that remained stable in terms of headship through the first eighteen months of the experiment.

We chose to look only at families that remained stable in terms of head of household. We expected stable families to behave quite differently from families that did not remain intact.⁹ For all groups we made comparisons with control families.

In both of the approaches, the NIT treatment plan in which a family was enrolled was specified in a way that identified the NIT-caused change in disposable income calculated on the basis of actual earnings in a single quarter before enrollment in the experiment. (Our analysis assumed no change from pre-experimental labor supply.)¹⁰

Besides the measure of change in disposable income, we also included a measure of the NIT-caused change in the marginal tax rate facing the family. This variable is intended to identify indirect effects of the NIT on subsidized housing through its effect on changing the relative price of leisure and other goods. We expected a positive relationship because other housing would increase relative to subsidized housing when work effort is reduced, as the cost of subsidized housing is generally 25% of a family's income.

The influence of the various NIT treatments was measured separately for families of different sizes. This was an important consideration for three reasons: (1) a larger family obtains a greater subsidy from the subsidized housing program than a smaller family with identical income; (2) NIT-caused income change is less significant to a larger family in terms of income

per person, and (3) larger families may find it more difficult to obtain adequate housing in the general housing market.

In addition to measuring the effect of actual income changes, we attempted to capture the average effect of the NIT by performing a separate analysis in which we replaced the income specification with a general indicator of whether the family was receiving an income maintenance treatment. Several other characteristics of the families that have not been mentioned so far are considered in the analysis, including: expected income of the family in the absence of the experiment; race; age of head(s) of household; and welfare status prior to enrollment.

FINDINGS

In general, the study shows that income maintenance substantially decreases the utilization of subsidized housing. The magnitude of the effect differs by site and how the household is headed and is sensitive to the level of support and family size.

We came to these findings by determining how much greater or how much less the probability is of each sample group (in each site) to reside in subsidized housing in the eighteenth month of the experiment. This was done by comparing them with selected control groups.

For example, in Seattle a woman without a family is 5.9% less likely to reside in subsidized housing with each \$1000 of NIT-caused increase in disposable income. (It should be noted that a single-person family does not exist in the experiment, but is merely used as a tool for analysis.) With each additional family member, the probability declines by 1.4%. Going through a few more steps with those findings implies that, in Seattle, for an average

family of four who are in subsidized housing at enrollment, the experimental average of \$1000 increase in disposable income and a 25% change in the tax rate causes a 4.3% reduction (from 70.6% to 66.3%) in the probability they will reside in subsidized housing in the eighteenth month of the experiment. If the family did not live in subsidized housing at enrollment, the probability (8.9%) that they would move into subsidized housing by the eighteenth month of the experiment decreased by .7%.

This type of analysis was also done separately on the probability of entering subsidized housing and of moving out of it as referred to when we discussed the method of analysis.

The results for the sample groups are described in more general terms in the next few paragraphs.

SINGLE PARENT FAMILIES HEADED BY WOMEN-SEATTLE

Single parent families headed by women in Seattle who are given an increase in income through the NIT are less likely to reside in subsidized housing 18 months into the experiment than those who are not. For those who did not reside in subsidized housing at enrollment, a NIT seems to increase the probability of moving into subsidized housing. The combined sample for both sites shows that a NIT-caused increase in disposable income is correlated with families who were in subsidized housing at enrollment moving out of it.

SINGLE PARENT FAMILIES HEADED BY MEN-SEATTLE

Single parent families headed by men in Seattle, whose income increases due to the experiment, are also less likely to reside in subsidized housing 18 months after the NIT treatment begins. It, however, does not cause those families who did not reside in subsidized housing at enrollment to move into it.

In the sample combined by site, there is also no significant effect on movement out of subsidized housing. A significant effect, however, could be occurring in Seattle, which is masked by the inclusion of Denver families in the sample.

SINGLE PARENT FAMILIES HEADED BY WOMEN-DENVER

Single parent families headed by women in Denver with NIT increases in disposable income are less likely to be in subsidized housing in month 18 than those without. The effect, however, becomes smaller with increases in the NIT tax rate. The NIT also seems to be related to the movement out of subsidized housing by those who lived there pre-experimentally.

In terms of entering subsidized housing, this sample group is more likely to do so given a tax increase, implying that the change in the relative price of leisure has an effect. Combined with Seattle, they are found to be more likely to leave subsidized housing with an NIT increase in income.

SINGLE PARENT FAMILIES HEADED BY MEN-DENVER

Results for single parent families headed by men in Denver were opposite those in all other sample groups. It indicates that they are more likely to reside in subsidized housing in month 18 given an NIT-caused increase in disposable income. This may mean that the indirect effects of the NIT on housing consumption through work effort changes and marital status changes prevail. In terms of entering subsidized housing, this sample, like the women heads of households in Denver, appears to be more likely to do so as the tax rate increases and the cost of leisure decreases.

DUAL-HEADED FAMILIES

The results for dual-headed families show that families receiving NIT financial treatments are much more likely to reside in subsidized housing in month 18 than those who are not. Families who were in subsidized housing at enrollment are more likely to move, but a financial treatment does not keep families who did not reside in subsidized housing at enrollment from entering.

SUMMARY AND CONCLUSION

The results of the study in general show a strong and significant interaction between the NIT as defined by the SIME/DIME experiment and the utilization of the type of subsidized housing found in Denver and Seattle. This interaction is seen in the generally lower probability of a family who receives an NIT residing in subsidized housing. Overall, the NIT also appears to be correlated with families moving out of subsidized housing and seems to inhibit families from moving in.

The diversity in response by site and by who heads the household shows that the NIT affects the utilization of subsidized housing differentially according to various characteristics. These differences are caused by several factors, including different labor supply effects, marital status effects, and different housing quality by site, and more complex modeling is necessary in order to understand more than the general orders of magnitude and the effects. Simultaneous models of labor supply (work effort), family stability and housing choice would be necessary to measure the effects according to various structural parameters. Also, more specific modeling of the housing subsidy programs would be useful and detailed knowledge of housing quality and housing markets at both sites would allow us to answer questions such as why single men who head households respond so differently by site.

NOTES

1. Attempts to identify changes in housing subsidies and to reimburse these changes through the grant have been ineffective. There has been only one grant change in response to an income-related change in rent in Denver and relatively few in Seattle.
2. In response to concern that enrollment in the experiment would cause families to lose their subsidized housing eligibility and create housing dislocations based upon fairly short-run income increases, SIME/DIME obtained promises from local housing authorities that experimental families would not be forced to move during the experiment. In fact, however, no one has been required to leave public housing in the Denver and Seattle areas since 1970, regardless of their participation in the experiment.
3. Adjusted gross income includes the total of earnings, social security, public assistance, state or private unemployment insurance, AFDC grants and SIME/DIME income. An allowance of \$300 is made for each dependent.
4. Public Housing, under U.S. Housing Act of 1937, consists of housing that is built or leased by the local housing authorities under federal cost limitations. It also includes Section 23 scattered site housing.
5. Section 236 of the U.S. Housing Amendment of 1968 provides for rental housing consisting of housing (five or more units) bought with interest reduction payments and rented out to tenants falling within the income limits.
6. A more detailed explanation of the methodology, with more detail on findings, can be found in the unabridged paper by the same title available from SIME.
7. In Seattle, approximately 23% of the families on the experiment were in subsidized housing at enrollment. In Denver, approximately 11% of the families resided in subsidized housing.
8. The study was designed in this way due to the form of the interview. A determination as to whether a family is in subsidized housing is based on the response to the question of whether the family lives in housing where the rent is subsidized. Because of the way in which the interview is constructed, if a family ever reported receiving a subsidy while living in a given house, that family is considered to be in subsidized housing as long as it remains in that house. Thus, a family may be in subsidized housing without actually being subsidized for a period of time.
9. These groups were stable in terms of how the household was headed at the time of enrollment in SIME/DIME. It does not preclude many changes, however, including other people leaving and joining the family.

10. Except for families on welfare before the experiment, the change in disposable income evaluated at initial equilibrium hours of work is equal to the payment (grant plus all positive tax reimbursements) the family would receive if it did not respond to the experiment or equivalently, the payment received by the family at enrollment. For families on welfare, their pre-experimental AFDC grant is subtracted from their payment in determining their change in disposable income. The reason the change in disposable income is not equal to the payment for families on welfare before the experiment is that families enrolled on SIME/DIME are required to give up their welfare status in order to receive payments.

DISCUSSION

DIANE BARTON: Did you do any geographical analysis? In Seattle, in particular, your choice of housing is not much. Segregation is such that you can always rent a house in South Seattle for a reasonable amount of money, but you could never rent a house in North Seattle for a reasonable amount.

AVRIN: One thing that we plan to do when data are available is to look at the quality of the housing units that our particular families were in, and including location.

LORRAINE R. WOJOHN: None of your analysis includes purchased housing at all?

AVRIN: We actually tried to analyze it, and just found that there wasn't enough data to permit it.

MYLES MAXFIELD: On your analysis of the flows into and out of the programs, if you were to have a national NIT, can you imply that after five years no one would be in subsidized housing?

AVRIN: No. Actually it seemed like the effects peaked at 18 months, and then the rate started declining after that. The effects are still significant if you look at it two years into the experiment, but the impact has declined-- so everyone won't leave.

MAXFIELD: The stock of people seems to go down. Would you say there was a certain fraction of public housing residents who are amenable to leaving, and a certain fraction which isn't?

AVRIN: Yes.

GLADYS MC CORKHILL: I have the impression nationally that the number of housing units being constructed is less than the need for the number of new families. Are we going to be in a housing shortage?

AVRIN: That is definitely true. Well, the problem is the baby-boom problem. Everybody born then is now about 30 years old or at prime home-buying age. So the demand is really increasing and will be increasing for awhile, and then there will be an over-supply or it will decline. There is also the trend of families splitting and requiring two homes instead of one, and single individuals buying homes. In Denver the housing market is just out of control. People can't find good places to live, so that is a problem, and also the cost of housing is going up faster than the consumer price index and people's incomes. This analysis was done a few years ago, in the first year and a half of the experiment.

THE EFFECTS OF ALTERNATIVE
NEGATIVE INCOME TAX PROGRAMS ON MIGRATION

by

Michael C. Keeley
Senior Economist
SRI International

Migration is an important equilibrating mechanism in a changing economy. It tends to reduce interregional differences in real wage rates, property values, and unemployment rates, and it is an important determinant of the geographical population distribution. The design of a nationwide NIT should, therefore, be concerned with migration. An important policy question in designing an NIT is the impact of alternative NIT programs on the rate of migration and the distribution of the population. A substantial redistribution of the population, especially the low-income population, would have important consequences for regional economic development. It also would have important fiscal effects on local government because tax revenues and expenditures probably would be affected. Also, we should understand the potential impact of welfare reform on migration to and from depressed regions with high rates of unemployment. Finally, the design should take into account not only the selection of appropriate support levels and benefit reduction (tax) rates, but in addition, determine whether to adjust benefits to account for regional variations in the cost of living.

Although existing studies of geographical migration that use nonexperimental data provide useful information regarding these policy issues, they do

not provide direct evidence regarding the likely impacts of a nationwide NIT on migration. In contrast, this paper uses data from the Seattle and Denver Income Maintenance Experiments (SIME/DIME) to analyze effects (on both migration rates and destination) of alternative NIT programs. In SIME/DIME, all sample members (both experimentals and controls) who move (within the United States) continue to be interviewed and, if eligible, to receive their NIT payments. Thus, SIME/DIME partially replicates a uniform NIT program in which the benefit structure is independent of region.

THE THEORETICAL FRAMEWORK

The decision of a family to migrate is the result of a complex decision-making process that depends on many factors, including the expected costs and benefits (pecuniary and otherwise) of the move. Economists traditionally have stressed the importance of wage differentials or other pecuniary gains as one of the prime motives for migration. In fact, there is considerable empirical evidence that supports this notion.¹ Viewed as a response to wage differentials, migration may be regarded as an investment in human capital because it may lead to a stream of benefits over time and because it entails current costs. The human capital investment model implies that the decision to migrate is similar to other investment decisions and depends on the rate of return to the investment.

An NIT program influences the decision to migrate because an NIT program affects the costs and benefits of an investment in migration. The basic functions and effects of an NIT are discussed in the paper by Spiegelman, above. Figure 1 in that paper displays the relationship between NIT support

and family income. As can be seen from that figure, an NIT program has several effects on benefits and costs and, hence, on the rate of return to an investment in migration. First, assuming hours of work or labor supply is fixed, the higher benefit reduction rate (tax rate) of the NIT reduces the rate of return if the monetary costs of migration are not deductible from income. (They are not in SIME/DIME.) This is because the higher tax rate reduces the income differential and hence the gain from migration, but does not correspondingly reduce the costs of migration unless they are deductible from income. If the monetary costs are deductible, then the tax rate of an NIT program reduces gains and costs proportionately and does not affect the rate of return to migration. Thus, the tax effect (without deductibility of costs) would tend to reduce the likelihood that people will move for monetary incentives and depressed regions would become even more permanent.² The NIT support level, assuming fixed labor supply and a perfect capital market, would not affect migration. However, it is likely that low-income persons would have difficulty borrowing the funds to finance a move; thus, higher income under an NIT should facilitate self-financing of migration. This effect would tend to offset the negative tax effect.

As shown elsewhere,³ both the support and tax components of an NIT reduce hours of work. This would reduce the likelihood of migration. For example, if an individual does not wish to work, even at the wage rates likely to be obtained by migrating, there is no financial reason for moving. Finally, both the support and tax components of an NIT reduce the variability of disposable income. This reduces the risk of migration and therefore should

induce migration if, on average, low-income people are risk-averse.

On net, the human capital model indicates two negative impacts: the tax effect and the reduced labor-supply effect; and two positive impacts: the reduced risk and increased effect availability of financing effects. Consequently, we cannot a priori predict the net effect of an NIT on migration.

An alternative, but not mutually exclusive way of viewing migration is to emphasize the nonpecuniary characteristics of different regions. In particular, one reason people move is to improve the quality of their environment. Generally, there is a negative tradeoff between the level of a region's real wage level and the region's environment. That is, locations with more desirable environments have lower real wage rates. This is because a region with a superior environment and equal or better wage rates would continue to attract new population until either wages fell or congestion reduced the level of the environment. Thus, migration is one important way individuals can substitute nonmonetary consumption (the environment) for income or market consumption.

An NIT increases the desirability of environmental consumption relative to market consumption because the high tax rate of the NIT reduces the return from market work but does not tax consumption of the environment. Thus, in much the same way as high positive tax rates encourage people to select jobs with lower wage rates and greater non-taxable "fringe" benefits, the tax rate in an NIT also encourages people to move to more desirable locations with, e.g., better climates, physical attractiveness, parks and other recreational facilities. The income effect of an NIT would also tend to lead to an increase in consumption of the environment. Consequently, this model would predict an increase in the rate of migration due to an NIT. It should be noted that in a

permanent NIT program, for each individual, this effect would be a one-time only effect and would not lead to a permanently higher rate of migration for a given individual.⁴ Thus, we would predict that an NIT would lead to a redistribution of the population to regions with better environments and lower wages.

A final issue concerns whether or not an NIT should use regional cost-of-living indices. From a strictly equity point of view, economists have little to say except that existing technology and information do not permit construction of meaningful indices.⁵ However, from an efficiency point of view, economic considerations indicate that resources would be wasted unless a uniform nationwide payment system were adopted.⁶ Currently, there are large inter-regional differences in welfare programs that probably induce migration. While individuals who move in response to welfare benefit levels increase their personal welfare, the welfare of society is reduced since the resources invested in migration serve no beneficial economic function. In fact, these moves would reduce total national real income if higher benefits are available in areas with a high cost-of-living (a reasonable assumption). Uniform nationwide NIT rules would eliminate this perverse incentive to migrate and could lead to a distribution of the population which maximizes each person's real income. A uniform nationwide NIT would provide incentives for people to move to or stay in lower cost-of-living areas. Such migration serves a beneficial economic function since each individual's increase in welfare contributes to a societal increase in welfare.

EMPIRICAL ANALYSIS

The effects of alternative NIT programs on migration are examined using SIME/DIME data. From an operational point of view, a person is considered to

have moved out of the area if he or she moves out of the greater Seattle or Denver metropolitan area and is gone long enough to miss a periodic interview.⁷ Records of the dates when persons move are kept in hard copy files by Mathematica Policy Research in both Seattle and Denver. With this procedure, a good approximation of an individual's date of move is obtained even if he or she should drop out of the experiment after migrating. Thus, attrition due to migration does not lead to a loss of information regarding the date of move although destination data is not known unless a post-move interview is completed.

Although some previous nonexperimental analysis suggests that the family is the appropriate decision-making unit, the analysis present in this paper concerns the response of individuals. This takes into account the considerable changes in families due to divorce, marriage, and remarriage during the time period of the study. Otherwise, response would have to be measured conditional upon the family remaining intact. Since there is considerable evidence that the experiment affects marital stability, using the family as the unit of analysis would lead to smaller samples and might seriously bias the estimates. Thus, all analysis is carried out on individuals classified by their original marital status regardless of subsequent marital status.

The data used for this study consist of all originally enrolled heads of households in SIME/DIME, except that data for Chicanos were not available when this study began. They are included only in the analysis of destination choice.

Table 1 contains tabulations of mobility data by race, site, original marital status and treatment status. In Denver, approximately 12% of the white married male and female controls move during the first nine quarters of the experiment. Their annual rate of approximately 5.4% is somewhat below

TABLE 1
 PERCENT MOVING BY RACE, MARITAL STATUS AND TREATMENT
 STATUS DURING THE FIRST 9 QUARTERS^a

	Seattle			Denver		
	Married Men	Married Women	Unmarried Women	Married Men	Married Women	Unmarried Women
<u>White</u>						
Financial Movers	17.2%	15.1%	5.8%	17.7%	18.1%	13.2%
Control Movers	59	52	13	54	57	25
Total	344	344	226	305	315	189
Financial Movers	15.7%	13.4%	8.1%	12.1%	13.9%	8.6%
Control Movers	51	44	13	28	33	10
Total	325	329	160	232	238	116
Difference (D)	1.46%	1.74%	-2.37%	5.64%	4.23%	4.61%
σ_D	2.91	2.72	2.84	3.29	3.35	4.08
$t (D/\sigma_D)$.50	.64	-.84	1.71*	1.26	1.13
Total Sample	669	673	386	537	553	305
<u>Black</u>						
Financial Movers	4.5%	5.6%	7.4%	8.3%	5.3%	7.3%
Control Movers	10	13	17	18	12	20
Total	223	231	230	216	225	273
Financial Movers	7.0%	4.7%	4.2%	6.6%	5.6%	12.1%
Control Movers	12	8	7	13	12	19
Total	171	171	168	198	214	157
Difference (D)	-2.53%	.95%	3.22%	1.77%	-.27%	-4.78%
σ_D	2.51	2.39	2.54	2.64	2.20	3.33
$t (D/\sigma_D)$	-1.01	.40	1.27	.67	-.12	-1.43
Total Sample	394	402	398	414	439	430

^aPersons who attrit before moving are eliminated from the sample.

*Significant at the 10% level

the national average for between-county moves. In Seattle, a somewhat larger proportion of controls (15.7% for married males and 13.4% for married females) move out of the area. Black controls move at about one-half the rate of the corresponding white groups reflecting national patterns of mobility by race. White single females move at lower rates than white married females, although black single females move at a greater rate than black married females. It should be noted that although these tabulations suggest that possibility of experimental effects, they do not control for systematic differences in income or family size between financials and controls caused by the assignment process.

Financial/Control Differences in the Rate of Mobility

Differences in the rate of mobility between controls and those receiving financial treatment are estimated using a maximum-likelihood statistical technique⁸ that uses data on duration until either a move occurs or until information is no longer available. Background and assignment variables are included in the equation to increase the precision of treatment estimates and to control for the stratified random assignment of treatment.

The estimates of financial impacts are presented in Table 2. On the average, white married males and females (both financials and controls) move out of the area at a rate of approximately 7% per year. The 5-year financial treatment increases the rate of mobility by 49% for white married males and 45% for white married females. These effects are statistically significant at the 10% level. For married black males and females, the average rate of mobility is only about 2.5% per year and experimental treatment has no statistically significant effect. Those on the 3-year financial treatments do not have statistically significant different rates of moving in any married subgroup.

TABLE 2

FINANCIAL/CONTROL DIFFERENCES IN THE RATE OF GEOGRAPHICAL MOBILITY^a

(Maximum likelihood estimated from the rate model)

	<u>R</u>	<u>5-year Financial Treatment Effect</u>		<u>Differential Effect for 3-year Treatment</u>		<u>χ^2 for site differences of treatment effects</u>
		<u>coefficient</u>	<u>ratio</u>	<u>coefficient</u>	<u>ratio</u>	
White Married Males	.072	.4023* (.2085)	1.495*	-.1180 (.2048)	.889	.22
White Married Females	.069	.3714* (.2150)	1.450*	-.1008 (.2105)	.904	.020
Black Married Males	.027	-.2013 (.448)	.817	.4529 (.4575)	1.573	1.78
Black Married Females	.023	-.08704 (.4616)	.916	.2352 (.4700)	1.26	.48
White Not Married Females	.037	-.6642 (.5134)	.51	1.156** (.4911)	3.177	1.81
Black Not Married Females	.034	-.1579 (.3995)	.85	.1479 (.4003)	1.15	3.77

^aR is the average rate per year of moving out of the area of both financials and controls. The ratio is equal to the exponential of the coefficient and gives the percentage effect of financial treatment on the rate of mobility. For example, a ratio of 1.495 indicates that 5 year financials move out of the area at a rate of 1.495 times greater than controls

* Significant at the 10% level.

** Significant at the 5% level.

The only statistically significant effect found for single females is that those on 3-year financial treatments have a greater rate of geographical mobility than controls.

For all subgroups there are no statistically significant effects of interactions between site and financial treatment on the rate of mobility.⁹ This suggests that the effect of the experiment does not depend on the unemployment rate which was much higher in Seattle.

An alternative way to measure the impact of the NIT on the rate of migration is to parameterize treatment. In this paper, NIT treatment is parameterized in terms of three variables; whether or not someone is above the break-even level, and if the person is below, the percentage change in disposable income and the percentage change in the net wage. The coefficient estimates are presented in Table 3. In addition, we use a dummy variable denoting enrollment in a 3-year financial program.

The results for white married males indicate statistically significant effects of income and tax changes on the rate of mobility. A 1% decrease in the net wage leads to a .94% increase in the rate of mobility. A 1% increase in disposable income leads to a .52% decrease in the rate of mobility. Similar results are found for white married females although the coefficients are smaller and the standard errors are larger. No significant effects for the other subgroups are detected. In this model, also, there is no indication of differential response depending on site.

As noted, the wage-differential model predicts that higher tax rates would make people less likely to move, while the environmental model predicts

TABLE 3
ESTIMATED COEFFICIENTS OF PARAMETERIZED TREATMENT MODEL
(Standard Error in Parentheses)

	\bar{R}	3-Year Treatment Dummy	Above Break Even Level Dummy	Percentage Change In Disposable Income	Percentage Change In Net Wage	χ^2 For Treatment Effects	F Tests For Site Differences In Response ^a
Married White	.072	-.0071 (.019)	.38 (.31)	-.52* (.31)	.94*** (.38)	8.36*	2.086
Married White	.069	.028 (.19)	-.066 (.36)	-.21 (.31)	.75* (.40)	4.79	2.053
Married Black	.027	.49 (.41)	-.210 (.62)	-.52 (.40)	-.048 (.079)	3.41	.610
Black Married	.023	.31 (.43)	.067 (.63)	-.87 (.58)	.11 (.85)	3.29	1.125
Single White	.077	1.07*** (.40)	-.70 (.75)	.041 (.13)	-1.09 (.75)	2.42	.086
Single Black	.070	.17 (.35)	-.49 (1.05)	-.33 (.66)	-.019 (.18)	1.43	1.94

^aTests of site-treatment interactions are based on ordinary least squares estimates of a model with a dichotomous dependent variable with 1 indicating a move during the first 9 quarters. The tests are performed by interacting each of the experimental variables with site dummies. The control variables are constrained to be the same in the tests.
*Significant at 10% level
**Significant at 5% level
***Significant at 1% level

that they would be more likely to move. Thus, our finding that higher tax rates lead to increased migration supports the notion that people are moving in response to the NIT for environmental reasons. The negative income effect, however, is not consistent with either model.

Effects on Destination

Out of the total sample of 384 married males who moved out of the area during the first nine quarters of the experiment, 71% or 271 completed post-move interviews that provide information on destination. Families headed by females also complete post-move interviews, but analysis of their behavior is not presented here.

Male heads of households moved from Seattle and Denver to all parts of the country. Table 4 gives the distribution of destination by treatment status and site. As might be expected, most people from Seattle move to the Pacific Region and most people from Denver move to the Mountain Region. In Seattle, financials and controls have a similar overall distribution of destination and our tests indicate no statistically significant difference. In Denver, however, financial treatment does seem to have an effect on the destination. In particular, financials are more inclined to move to the Pacific Region and controls are more likely to move to the West North Central Region.

Tables 5, 6 and 7 give estimated coefficients of the effect of financial treatment on climate, job-related variables and economic conditions at the destination for married males, with white married males shown separately. We find financial/control differences after correcting for assignment and background variables.

TABLE 4
 DESTINATION OF FIRST MOVE:
 DISTRIBUTION BY TREATMENT STATUS AND SITE--MARRIED MALES ALL RACES^a

Region	S E A T T L E		D E N V E R	
	Control	Financial	Control	Financial
1. New England.	0 (0)	3.8 (3)	0 (0)	0 (0)
2. Middle Atlantic	3.7 (2)	2.5 (2)	(2.2) (1)	0 (0)
3. South Atlantic	1.9 (1)	2.5 (2)	2.2 (1)	6.7 (6)
4. East North Central	1.9 (1)	6.3 (5)	2.2 (1)	3.4 (3)
5. East South Central	1.9 (1)	1.3 (1)	0 (0)	1.1 (1)
6. West North Central	5.6 (3)	6.3 (5)	13.0 (6)	9.0 (8)
7. West South Central	9.3 (5)	10 (8)	30.4 (14)	11.2 (10)
8. Mountain	16.7 (9)	10.0 (8)	39.1 (18)	43.8 (39)
9. Pacific	59.3 (32)	57.5 (46)	10.9 (5)	24.7 (22)
TOTAL	100 (54)	100 (80)	(46)	(89)

χ^2 tests of overall relationships between treatment and region

$\chi^2 = 4.9$

$\chi^2 = 13.65^*$

* significant at the 10% level

^aThe numbers in this table are the percentages of the total number of persons in the category who move to the designated region. Numbers in parentheses are actual numbers of persons in the cell.

TABLE 5

EFFECTS OF TREATMENT ON CLIMATE AT DESTINATION

<u>Dependent Variable</u>	<u>White Married Males</u>			<u>All Married Males</u>		
	<u>Mean</u>	<u>Coefficient of Financial Treatment</u>	<u>F-test for Seattle-Denver Differences in Response</u>	<u>Mean</u>	<u>Coefficient of Financial Treatment</u>	<u>F-test for Seattle-Denver Differences in Response</u>
1. January Mean Temperature	36.0 ^o	3.087* (1.70)	.169	36.6 ^o	1.60 (1.47)	.408
2. Precipitation (inches)	25.7"	2.42 (2.068)	3.84* (S=5.60 D=-2.62)	25.1"	2.38 (1.79)	2.45 (S=4.97 D=-.57)
3. How Important was Climate? 1 = important 0 = unimportant	.46	.083 (.078)	2.74* (S=-.021 D=.24)	.47	.12* (.067)	1.92
4. July Mean Temperature	72.8 ^o	-.38 (1.025)	.134	73.7 ^o	-.63 (.89)	1.4
5. Variance in January and July Temperatures	1509.3	-249.79* (131.90)	.261	1523.0	-186.76* (111.29)	.061
6. Variance in High and Low Temperatures	14619.4	-1514.9** (697.0)	.5	14579.7	-1076.0* (606.35)	.102

NOTE: The coefficient estimates in this table give financial (i.e., NIT treatment) - control differences in the dependent variable after eliminating differences caused by the assignment variables. For example, equation 1 indicates that white married male financials moved to regions with 3.087^o warmer January temperatures. In parentheses under some of the F-tests for Seattle-Denver differences in response, the effects for Seattle (S=) and Denver (D=) are presented separately.

TABLE 6

EFFECTS OF TREATMENT ON JOB-RELATED VARIABLES AT DESTINATION

<u>Dependent Variable</u>	<u>White Married Males</u>			<u>All Married Males</u>		
	<u>Mean</u>	<u>Coefficient of Financial Treatment</u>	<u>F-test for Seattle-Denver Differences in Response</u>	<u>Mean</u>	<u>Coefficient of Financial Treatment</u>	<u>F-test for Seattle-Denver Differences in Response</u>
1. Distance Moved (miles)	849.6	238.3* (121.63)	.059	831.1	149.1 (99.76)	.29
2. Costs of Moving (dollars)	364.1	168.9** (71.80)	.093	309.8	116.9** (54.78)	.021
3. Did Person Intend to Work?	.87	.0131 (.052)	.017	.86	-.02 (.045)	.12
4. Did Person Have Job Lined Up?	.31	.00005 (.074)	1.606 (S=.075 D=-.11)	.42	-.012 (.061)	1.69
5. Hours of Work Per Week Expected	37.8	3.45 (2.94)	.33	36.8	1.16 (2.45)	.33
6. Earnings Per Week	149.3	24.12 (15.22)	.06	146.	18.0 (12.47)	.00

* Significant at the 10% level
 **Significant at the 5% level.

TABLE 7

EFFECTS OF TREATMENT ON ECONOMIC CHARACTERISTICS OF THE DESTINATION

	<u>White Married Males</u>			<u>All Married Males</u>		
	<u>Mean</u>	<u>Coefficient of Financial Treatment</u>	<u>F-test for Seattle-Denver Differences in Response</u>	<u>Mean</u>	<u>Coefficient of Financial Treatment</u>	<u>F-test for Seattle-Denver Differences in Response</u>
1. Median Income (MEINTO)	\$9549.2	213.1 (190.3)	.17	\$9464.2	205.02 (173.34)	.012
2. Per Capita Income (PCMOIN)	\$3186.8	112.3* (67.75)	.23	\$3146.5	79.9 (61.36)	.058
3. Per Cent of Families Below Poverty Level (PERFBL)	9.54%	-.13 (.57)	.26	10.03	-.32 (.57)	.19
4. Per Capita City Government Expenditures (PERCAP)	\$17.9	12.65 (11.5)	1.16	117.6	14.88 (12.10)	.034

The results presented in Table 5 support the environmental model. Married males receiving financial treatment move to regions with warmer winters and smaller year-round variations in temperatures. In addition, financials are more likely than controls to indicate that climate was an important reason for their move.

Table 6 presents similar estimates of experimental effects for a set of job-related variables at the destination. White married males moved greater distances and all married males (including white separately) spent more on the move. This may indicate that the higher income caused by the NIT enables persons to move further to find a more desirable location. All the job variables are derived from questions relating to the person's job expectations prior to the move, not his actual job experience. (The investigation of actual job-related behavior is planned in a future study.) As can be seen, there are no experimental/control differences in expectations relating to work at the new location.

Table 7 presents the estimated effects of experimental treatment on the economic characteristics of the destination for married males. Although there is some suggestion that treatment induces persons to move to higher-income cities, the only statistically significant effect is on per capita income for whites.

SUMMARY AND CONCLUSIONS

Based on the findings presented in this paper, we can conclude that a uniform nationwide NIT benefit structure promotes economic efficiency, compared to the existing system which probably adjusts to some extent for regional differences in the cost of living.

The NIT programs tested in SIME/DIME increased the rate of migration for white, married males and females--by 49% and 45% respectively for those on the 5-year treatment. For this group, both income and tax rates influenced mobility and the tax rate effects were strongest. The positive response to the higher tax rate is consistent with the theory that people move to improve their environment.

None of the experimental groups examined were significantly different from controls in responses by site--some evidence that unemployment rates at the site do not affect migration due to an NIT.

For married males separately, there is some evidence that a uniform NIT would induce persons to move to better environments and to spend more on the move. White married males also moved further. Finally, there is no evidence that expectations about work, hours or pay in the new location was any different for controls or experimentals.

NOTES

1. See e.g., L. A. Sjaastad, "The Costs and Returns of Human Migration," Journal of Political Economy 70, no. 5, part 2 (October 1962, Supplement): 80; J. DeVanzo, "An Analytical Framework for Studying the Potential Effects of an Income Maintenance Program on U.S. Interregional Migration" (The Rand Corporation, Santa Monica, California, December 1972); A. Schwartz, "Interpreting the Effects of Distance in Migration," Journal of Political Economy 81, no. 5 (September/October 1973): 1153; A. Schwartz, "Migration, Age, and Education," Journal of Political Economy 84, no. 4, part 1 (August 1976): 701; M. J. Greenwood, "Research in Internal Migration in the U.S.: A Survey," Journal of Economic Literature 13, no. 2 (June 1975): 397; M. J. Greenwood, "A Simultaneous-Equations Model of White and Non-White Migration and Urban Change," Economic Inquiry 14, no. 1 (March 1976): 1.
2. This effect is one strong argument for allowing job related moving costs to be deductible from income in any nationwide NIT program.
3. See for example, Michael C. Keeley, Philip K. Robins, Robert G. Spiegelman, and Richard W. West, "The Estimation of Labor Supply Models Using Experimental Data," American Economic Review 68, no. 5 (December 1978): 873; and Keeley, et al., "The Labor-Supply Effects and Costs of Alternative Negative Income Tax Programs," Journal of Human Resources 13, no. 1 (Winter 1978).
4. The aggregate rate of migration, however, would increase since new individuals would be continually entering a permanent NIT program. In addition, changes in the environment or changes in individual characteristics over time might lead to a permanent effect on the rate of migration.
5. See Janice Peskin, "Geographical Payment Variation in a Federal Welfare System" (Technical Analysis Paper No. 14, Office of Income Security Policy, DHEW, January 1977).
6. This is because the cost of making a given payment is independent of the cost-of-living in a particular region.
7. In Seattle, moving is defined as moving out of the four-county area that includes King, Pierce, Snohomish, and Kitsap counties. In Denver, moving out of the area (MOA) is similarly defined. The Denver metropolitan area includes the cities of Arvada, Aurora, Broomfield, Brighton, Boulder, Denver, Englewood, Evergreen, Golden, Lafayette, Louisville, Lettleton, Longmont, Thornton, and Westminster.
8. Developed by Nancy Tuma in "Rewards, Resources, and the Rate of Mobility: A Non-Stationary Multivariate Stochastic Model," American Sociology Review 41 (1976): 338.
9. Statistical tests for the presence of site interactions are performed by interacting a site dummy variable with F and Y3.

EFFECTS OF INCOME MAINTENANCE ON
SCHOOL PERFORMANCE OF CHILDREN

by

Larry M. Manheim
Senior Economist, MPR

and

Mary Ellen Minchella
Statistician, MPR

The main purpose of income transfers is to alleviate current poverty status caused by a lack of current earning power. The question arises, however, as to whether such income transfers might also have positive effects on the human capital of individuals, which may in turn allow them to increase their long-term earning power. The specific goal of the current research on the school performance of Seattle and Denver school children is to determine whether income transfers to families increase the human capital of children in those families by improving their performance in school.

We limit ourselves here to three measures of school performance; standardized test scores, academic grades, and absence rates. Each of these performance measures has associated problems which make them somewhat ill-suited to the measurement task at hand. Standardized test scores provide a relatively objective indicator of performance, but will be much less sensitive to specific learning gains than categorical tests, designed to measure specific subject learning. Academic grades might fulfill such a function, but they are also likely to include teacher biases and be normalized by the immediate peer group performance. Finally,

attendance, which may be expected to reflect acceptance of the schooling process and thus be a measure of socialization (and perhaps a predictor of future achievement) will largely be a function of random illnesses. Subject to these caveats, it is still believed that each of the measures represents an important indicator of success in and adjustment to school, and may be a predictor of future success.

HYPOTHESIZED EFFECTS OF A NEGATIVE INCOME TAX ON CHILDREN'S SCHOOL PERFORMANCE

The effect of a negative income tax (NIT) on the school performance of children could depend on an income effect and a labor supply effect. An income effect could result if the family can purchase more goods and services which contribute to the school performance of their children--such as housing which reduces family crowding, or books and learning games. It may also reduce household stress, allow better nutrition, and decrease the need for older children to earn money to the neglect of school work. A labor supply effect could result if an NIT induces parents to convert work effort in the labor market into productive home services. It is hypothesized that increased parental time spent with children will have positive effects on the child's school performance; a result of increased parental teaching, discipline and support.

In this paper, there is no attempt to distinguish among these varied effects of an NIT. Rather, the question asked here is whether the total net effect of the transfer program is to improve school performance.

DATA COLLECTION FOR THE SCHOOL PERFORMANCE RESEARCH

The analysis of school performance presented here depended on the SIME/DIME data base, supplemented by data from school records, which contain information on student disabilities, academic grades, standardized test scores and attendance rates. A problem with the use of such a retrospective data collection effort is the inability to reconstruct missing data. For example, if tests were not given, or if any of the measures were not permanently recorded in a reasonably accessible form for many children in the experiment, the sample size may quickly fall. Even where tests and grades are available, they do not come in standardized forms, making comparisons somewhat tenuous. Despite such difficulties, sufficient numbers of observations are available, even when we require both pre - and during - experimental performance measures, to allow a relatively detailed analysis of experimental effects.¹

If there are experimental effects, performance measures may be expected to improve more for experimentals than controls during the period of the experiment. However, much of the variation in scores is due, not to such improvements, but rather to differences in human capital which exist at the beginning of the experiment. To eliminate as much of this variation as possible, pre-experimental measures of performance are controlled for whenever possible.

METHODS OF ANALYSIS

The method used here to estimate experimental effects is regression analysis.² Regression analysis allows us to control for variables used

in the SIME/DIME assignment model. In addition, the precision of estimates may be increased by controlling for factors which were fixed prior to the experiment (and thus not influenced by it) and which also affect school performance.³ Separate regressions are estimated for each year of data, since it is expected that time spent in the experiment is an important determinant of performance.⁴ Given the decision to run separate regressions by length of the experimental period, it is not possible due to limited sample sizes to also run a separate regression for each grade level, or for individuals with differing characteristics. However, since it is believed that high school represents a distinctly different period of development, with less possibility of altering cognitive levels of achievement, separate samples are run for students who are in high school (grade 9 through 12) at the time of the post-enrollment performance measure. For each of these regression samples an overall treatment effect is estimated. Then statistical tests are performed to determine whether there are differing treatment effects by grade level, value of pre-experimental performance measures, values of total pre-experimental income, and presence of a male head. The magnitudes of any differences are also estimated.

The main results reported in this paper are for school performance measures collected for the 1972-73 school year in Seattle and for the 1973-74 school year in Denver. As of the close of these school years the experiments had been running for close to two and one-half years. Also, both the three year and five year treatment families were still

in the experiment thus maximizing the number of experimentals.

It is also of interest to examine performance in 1974-75 (1975-76 in Denver) to see whether the five year sample either begins to demonstrate gains or has further increased any gains observed previously; and whether the three year experimentals have maintained any gains found while they were still in the experiment.

Results on Experimental Effects

Tables 1 and 2 summarize the effects of the experimental transfer on during-school performance measures. Two alternative specifications of experimental treatment are reported here. The first is simply a binary variable indicating whether the individual was enrolled as an experimental (=1) or control (=0) during the period. The second measure is the dollar increase in income which would occur under the experiment if there were no change in behavior. For controls, this is computed at zero dollars; for experimentals, as the difference between actual income in the year prior to enrollment and a calculation of experimental income (given no behavior effects).⁵ Sample sizes and the grade levels included in the regression samples are also reported. These regression results assume identical treatment effects for the different grade levels, pre-experimental income levels, pre-experimental performance levels, and for families headed by couples or a single parent. We also indicate whether allowing for differential treatment effects by grade level results in statistically significant treatment effects, for any specific grade levels. In regression results are not reported here, differential treatment effects by the above characteristics were allowed for, but almost never provided statistically significant differences, even at the 10% level.

TABLE 1
 NIT EXPERIMENTAL EFFECTS ON SCHOOL PERFORMANCE
 IN DENVER (1973-74)

	Math Tests	Verbal Tests	G.P.A.	G.P.A.	Attendance
Regression #1					
Coefficient on Treatment Dummy	-6.56	2.14	-.00	.30	-.24
(S.D. on Treatment Dummy)	(5.16)	(6.12)	(.63)	(1.11)	(.20)
R ²	.46	.27	.31	.28	.24
Regression #2					
Coefficient on \$1,000 Change due to Experiment	-.72	.73	-.37	.03	.13
(S.D. on \$1,000 Change)	(2.83)	(3.24)	(.39)	(.08)	(.72)
R ²	.46	.27	.31	.28	.24
Grade Levels as of Enrollment in 1972 ^a	3	3	2 - 6	7 - 10	2 - 4
Sample Size	94	85	711	376	335
Mean for 1974 Dependent Variable	30.87	29.95	77.80	71.96	11.38

^a Treatment was not significant for individual grades.

TABLE 2

NIT EXPERIMENTAL EFFECTS ON SCHOOL PERFORMANCE
IN SEATTLE (1972-73)

	Math Tests	Verbal Tests	G.P.A.	G.P.A.	Attendance
Regression #1					
Coefficient on Treatment Dummy	-.95	-1.58	1.41*	-.17	-.23
(S.D. on Treatment Dummy)	(2.24)	(1.85)	(.74)	(.94)	(1.62)
R ²	.64	.72	.14	.28	.22
Regression #2					
Coefficient on \$1,000 Change due to Experiment	1.15	-.77	.44	-.66	-.79
(S.D. on \$1,000 Change)	(1.60)	(1.30)	(.59)	(.61)	(1.17)
R ²	.64	.72	.14	.28	.22
Grade Levels as of 1971 ^a	2 - 6	2 - 6	7 - 10	2 - 6	7 - 10
Sample Size	234	261	453	451	210
Mean for 1973 Dependent Variable	31.28	33.16	75.66	12.78	14.78

^a Specific grade levels were not available for these grades and grade level dummy variables therefore could not be included as control variables.

* Significant at 10 percent significance level; two-tailed test.

The most striking finding is the almost universal lack of statistically significant differences between experimentals and controls in the regressions. This is true when using either specification of the treatment variable. An examination of treatment coefficients suggests that this lack of statistically significant treatment effects is not simply due to an inability to measure such effects precisely.

Grade Point Average

Grade point averages (G.P.A) exist for all grade levels in Denver, but only for high school students in Seattle. For high school students, the G.P.A. used is a weighted (by credit hours) average of non-vocational courses, excluding grades in non-academic classes, such as physical education. For students in grades K-8, the G.P.A. reflects the simple average of language arts, mathematics, science and social studies courses.

The coefficients estimated for experimentals in Denver are negligible for both the younger and older children. The effect is estimated as positive and more substantial for the Seattle children, who were high school students during 1972-73. This Seattle effect, while significant at the 10% level when a treatment binary variable is used, becomes less positive and statistically insignificant when the specification of the treatment variable used is change in income due to the experiment. This occurs because the treatment effect for these Seattle children is only substantial for those whose pre-experimental adjusted income from all sources was greater than \$4,800 (and therefore eligible for only a small experimental transfer). If we alternatively let this treatment effect vary by both

whether the family contains a couple at the head and whether the wife works, we find the largest (and only statistically significant) result is for families with couples, where the wife did not work during the year prior to the experiment. In families where the wife did work, the experimental effect is negative but insignificant. Thus, the estimated effect cannot be explained as due to wives who worked being induced to stay at home and increase their home productivity. The Seattle results do not, therefore, appear to be strong enough for us to alter our conclusion from the Denver sample; that there are no substantial differences between experimentals and controls.

For the Denver sample in Grades 2-5 as of 1972, we could also estimate possible treatment effects in 1976, after two more years had passed. In 1974 only the five-year subsamples were still on the experiment. We thus allowed both for treatment effects for all experimentals and for additional treatment effects for students who were still on the NIT experiment as of 1976. The treatment effect for those off the experiment as of 1976 is negative (-1.12) while that for the five year sample is positive (.62), but both coefficients are statistically insignificant.⁶

Finally, we may compare these results to past research using children in the Gary NIT experiment sample. Maynard and Murnane also found no positive treatment effects on grade point average, the simple treatment effect being generally negative and usually statistically insignificant.⁷

Absentee Rates

Absentee rates are more within the direct control of students and their families than grades or test scores. They thus may be more sensitive to policy intervention. This statement must, of course, be modified to the extent that illness determines the absence. We control somewhat for health using a pre-measure of both absence rates and disability but our main control against unobserved health biasing the treatment coefficient is the (conditionally) random assignment of individuals to either treatment or control status.⁸ Note also that the experiment may be hypothesized to decrease absences by increasing the expected health of the children through, for example, better nutrition and less crowding.

We observed absentee rates with pre-experimental measure for grade levels 5-8 for Seattle and for grade levels 5 and 6 for Denver. For Seattle, high school absentee data is virtually unavailable (because grade levels are not given); thus we do not control for them in the regressions for this sample. As with the composite G.P.A., none of the coefficients are statistically significant, and while all of the coefficients on the binary treatment variables are negative⁹ (treatments having less absences), the estimated effect is extremely small (less than a 2% decrease in absence rates in all three cases). Allowing the treatment effects to vary by grade level or by family characteristics also fails to provide any statistically significant effects.

For the Seattle sample, we are also able to estimate whether there are treatment effects on absences after another two years, during the 1974-75 school year.¹⁰ The results differ substantially by which spec-

ification of the treatment variable is used. Using the treatment variable measuring the expected dollar gain from the experiment provides no results of note.¹¹ However, use of the simple binary treatment variables for the three-year and five year experimentals results in a statistically insignificant reduction in absences for three year experimentals of 1.87 days and a statistically significant¹² reduction in absences of 3.70 days for five year experimentals. It is difficult to further decompose this effect since there are only 52 five year experimentals; to illustrate, only 10 of these have pre-experimental income less than \$4,800.

Given this large increase in experimental effects as we move from the 1972-73 to the 1974-75 post-measure, we also estimated the treatment effect on absence rates during the 1973-74 school year.¹³ One would expect the effects to be between those found in 1972-73 and 1974-75, and this is indeed the case. Use of the binary treatment variables results in a reduction in absences of 1.74 days¹⁴ for three year experimentals and a reduction of absences of 3.25 days¹⁵ for five year experimentals.¹⁶

In comparing these results with those for Gary, we note that only seventh grade treatments in the Gary NIT experiment had significantly fewer absences (third graders also had statistically significant effects for some subsets of individuals) with the overall results rather mixed.¹⁷ We conclude that any treatment effects on attendance are quite small in the short run, but may become fairly substantial as the experimental period increases.

Standardized Test Scores

The last set of performance measures examined in this paper are mathematics and verbal standardized test scores. In Seattle, these tests are available for most grade levels, with the exception of high school students for whom they were not collected. Seattle used the Metropolitan Achievement Tests in all grades. In Denver, the available test data is much more spotty, with tests being given relatively infrequently and for different grade levels by year. The types of tests given also vary substantially from year to year. Math and verbal results for third graders (as of enrollment) in Denver are provided in Table 1, but the sample size here was not large enough to provide precise estimates for this single grade. Test scores are coded in percentiles (the score most available). The scores may therefore be thought of as standardized, with treatment effects measuring increases in achievement relative to the national norm.

Neither the Seattle mathematics or verbal test scores in 1973 show statistically significant differences by experimental status. The actual coefficients are negative and positive for verbal and mathematics percentiles, respectively. Further, no statistically significant treatment differences emerge if we allow treatment effects to vary by grade level, pre-experimental test scores, income status, or by whether the child was in a family headed by a woman only.

Since high school test data do not exist in Seattle, we could only look at performance levels after another two years on the experiment

for second and fourth grades (as of enrollment). There is no change in the treatment coefficient in moving from the 1973 to the 1975 performance measures for these two grade levels.¹⁸

The Denver treatment effects for grade 3 (at enrollment) are also statistically insignificant, but in this case the small sample size¹⁹ limits our confidence in the estimated magnitudes of the coefficients.

The research on standardized test scores using the Gary NIT sample only looked at verbal achievement tests. Gary experimentals in fourth grade at the time of the post-enrollment tests had significantly higher test scores than controls.²⁰ By comparison, Seattle experimentals in fourth grade in 1972-73 actually had slightly lower verbal test scores than Seattle controls, though as noted the treatment effect was statistically insignificant.

CONCLUSIONS AND IMPLICATIONS FOR FUTURE RESEARCH

The estimation of the effects of NIT transfer on school performance began with the New Jersey and rural NIT experiments. In New Jersey, experimentals appeared to stay in school longer than controls. Other performance measures were not available. In the rural experiments in Iowa and North Carolina, the only positive results were for the North Carolina sample of younger children. It was argued that North Carolina families, poor and with limited welfare alternatives, contained a sample helped by experimental status to a degree found in no other NIT experiment. And the younger children in this sample, with a shorter history of pre-experimental deprivation, probably had the most potential to alter their

school performance in response to NIT transfers for the family.

The Gary, Denver and Seattle results have shown that experimental effects on grade point averages, absences, and standardized test scores exist, if at all, at a low level. Given these results, decisions about the value of switching from current welfare to proposed NIT programs should rest on grounds other than the expectation for indirect experimental effects on school performance. One caveat here is that in areas such as the rural South, where alternative forms of welfare are very limited, the introduction of a substantial welfare program may induce positive school performance for many or all of the measures considered here. Another point to be made is that even if these performance measures do not substantially change, attitudes about the type and extent of education to be obtained may change, reducing dropout rates and extending the quantity of schooling received by children from low-income families. These issues will be the subject of future research.

The question arises as to whether the lack of either substantial or consistent treatment effects in Seattle and Denver result because experimentals have not altered their inputs into schooling or because their inputs do not, in fact, substantially affect performance levels within the time period observed. Future research will consider both these possibilities, estimating experimental effects on certain aspects of home environment and using both during-experiment data on student families and school environment data to estimate the effects of various aspects of the home and school environments on the school performance

on the Denver and Seattle samples. This research should further define the possibilities and limits of policy in altering the school performance of low-income, urban children.

NOTES

1. For a complete discussion of the data including testing schedule in Denver and related missing data problems, see Minchella, "SIME/DIME School Performance Interview Files," Mathematica Policy Research, Inc., Memo: EP #601.
2. A more complete description of the methodology and results can be found in Larry M. Manheim and Mary Ellen Minchella, "The Effects of Income Maintenance on the School Performance of Children: Results from the Seattle and Denver Experiments," (MPR Working Paper, forthcoming). See Also, Rebecca Maynard, "An Analysis of the Home Environment Determinants of School Performance," (Ph.D Dissertation, University of Wisconsin, 1975).
3. See the paper by West, above for a more complete explanation of the reasons regression analysis was used. The variables used in the regressions included the three performance measures, NIT treatment, increased income due to treatment, race, family type, number of children under 16, working status of mother, education of mother, and several interaction variables and additional explanatory variables. For a complete treatment, see Manheim and Minchella, forthcoming.
4. Maynard and Murname, 1978, in a similar analysis, run separate regressions by grade.
5. Computed differences do not include the tax benefit which results from going on the experiment. Thus, benefits are understated to some extent. The average computed increase in income due to the experiment ranges from between \$700 to \$900 in the regressions reported here. Aside from the omission of the tax benefit, the computation is similar to that performed by Keeley, et al., 1978.

6. The sample size was 765 with 295 three-year experimentals, 134 five-year experimentals, and 243 controls. The differences in the coefficient for the three-year and five-year experimentals is only significant at the 10% significance level (t-statistic is 1.82).
7. They suggest some possible reasons why the coefficients may be biased downward, including experimentals moving to areas with more difficult schools, experimentals dropping out of school less frequently, and experimentals being induced to take more difficult subjects. They were only able to test the first two hypotheses and found these did not explain the negative results.
8. For the Denver sample, use of the \$1,000 change treatment variable results in a positive (but negligible) coefficient.
9. These results use ordinary least squares regression analysis. The number of cases with zero absences was always less than 5%, suggesting this lower bound was not enough of a constraint to switch to a more complex estimation procedure such as Tobit estimation. Absence rates greater than 50 were set to 50 to prevent outliers from completely determining the regression results. In fact, setting this upper bound had little effect on the results.
10. This sample consists of 354 children in grade levels 2 through 6 as of enrollment into the experiment as either a treatment or control. There are 157 three-year experimentals and 52 five-year experimentals in the sample, and the mean number of absences in 1974 is 14.49.
11. The coefficient for the three-year experimentals is -.99 while it is .00 for the five-year experimentals, both coefficients statistically insignificant.
12. T-statistic of 2.01.
13. This sample consists of 420 children in grades 2 through 6 as of enrollment. There are 186 three-year experimentals and 58 five-year experimentals in the sample, and the mean number of absences in 1973-74 is 12.83.
14. T-statistic is 1.62.
15. T-statistics is 2.07.
16. It is only possible to look at 1975-76 absence rates for Denver for those in grade two at enrollment. There were only 82 in the sample, of which 53 were three-year experimentals and 11 were five-year experimentals. The statistically insignificant coefficients on the treatment binary was -.71 for three-year experimentals and -2.14 for five-year experimentals, both measured rather imprecisely.

17. Maynard and Murnane, 1978.
18. One argument is that the treatment effects will be stronger for younger children. We look at the Seattle mathematics test scores in 1972-73 for children in grades 1, 2 and 3 at that time (there was no pre-enrollment measure of performance). The binary treatment effect was negative for these children (-3.84) though not statistically significantly (139 observations).
19. The sample size here was 144 observations for verbal tests and 124 observations for math tests.
20. They looked at five grades. Grade 3, 5 and 6 also showed (and statistically insignificant) positive treatment effects. Only grade 8 had a small negative (but statistically insignificant) treatment coefficient.

DISCUSSION

CAM DIGHTMAN: You've indicated that you ran into trouble in Seattle with academic performance measures in the lower grades because schools had different, noncompatible systems. It occurs to me that the teachers must have some way of coping with the fact that the child who comes into their classroom has been graded under some other kind of system, and they must be in a position to make some kind of judgment about the relative performance of the child. I wonder whether you considered using any "expert judges" to at least get you orders of magnitude or directional differences in terms of where a kid was.

MANHEIM: That would be very expensive, and we didn't have the money to do it.

JAMES SHORT: Have you coded any information on reasons for absence? You mentioned health reasons as being maybe predominant, but there may also be disciplinary reasons.

MINCHELLA: We don't have access to that information for our analysis, but on the Seattle transcripts that came in for students in high school, we do have transfer history with dates of suspensions and reentry. However, it was too difficult to get that information into the analysis format. The data came in a form that required extensive programming efforts by our technical people. We got as much as we possibly could from the transcript, but we couldn't get all of it. For example, we didn't get the grade level; that was one of our biggest disappointments. In the space where the grade level would be recorded, in some cases you would have years recorded, some cases, grade levels, in some cases nothing at all. It just didn't seem worthwhile to spend a lot of money at that point.

GARY REED: I understand that in the future work, you're going to try to disaggregate the reasons for performance and nonperformance. It seems that one of your causal links here is increased time of the mother in the house. There should be more direct ways of trying to test that than what you've done to date.

MANHEIM: Exactly. What we've done to date is just test whether there are experimental effects. What we want to do now is look at the amount of time the female head or the wife spent in the home and with the child; but additionally, using during-experimental data, we want to see if children perform better when wives spent more time with the child. There are questions, "Do you read to the child?" "How much?" "How much time do you spend on the child?" "Do you go on trips?"

SHORT: Actually, if you look at Phil Robins' paper, and those huge increases in employment for both wives and single female heads, that might be an intervening variable as far as school performance is concerned. The lack of a treatment effect doesn't tell you whether there might be something systematically related to that variable.

BURT BARNOW: Marital dissolution might also fit this description.

DIGHTMAN: Unfortunately, another moderating effect on all of those is the quality of the mothering that accompanies that additional time.

MANHEIM: But that's a good reason to use this experimental data: mothers who stay home anyway may be quite different from mothers who will be induced to stay home by an experiment.

BARNOW: Did you look at the effects of normal income on the family?

MANHEIM: We had a variable measuring whether the family earned a total income less than \$4800 in the pre-experimental year. You have to control for normal income because of the assignment problem. It's somewhat hard to disentangle the effects but people who had lower normal income generally had lower performance measures, though the effects often weren't significant.

BARNOW: You didn't put in occupation of the parents or anything like that?

MANHEIM: No.

BARNOW: You could, couldn't you? You have those variables.

MANHEIM: Right, we could code it and put it in.

BARNOW: I did some work on Headstart, and I found significant effects for occupation, especially for mother's occupation and education.

MANHEIM: I think that is a very important thing to do when we look at how the various elements of an NIT affect school performance.

BARNOW: On your tests for Seattle where you have four different grades that you're looking at, are you pooling all the observations and running one big regression or do you do a separate one for each grade level?

MANHEIM: We pool the observations, and run one regression, and then we put in an additional treatment interaction with each grade to look for differences.

SHORT: What are your next steps? What are you going to do now?

MANHEIM: The next steps are first, to look at the modules, at changes in attitudes due to the experiment; changes in time spent on homework; whether children in grades 9 through 12 have increased working or decreased working as a result of the experiment; and whether dropout rates have declined. We will then move on to the during-experiment data, and actually look at how these different variables might affect school performance.

SIME/DIME CHILD CARE AND PUBLIC POLICY

by

Samuel Weiner
Senior Economist
SRI International

Child care is at an important juncture in America today, with some groups seeking large-scale expansion of child care centers and a sharp increase in federal involvement. An opposing force argues that child care should be kept within the family. Between these forces are those seeking quality child care in whatever form users may desire, but free of any government interference. Until recently, arguments for and against these positions have been based on research and analysis on the effect of extra-family child care upon the child's emotional, physical and cognitive development. Research and analysis of related economic issues have been conspicuously missing, however.¹ Yet, if any significant change in public responsibility for child care is envisioned, we must attend to the economic implications.

Thus, this paper begins to bridge the gap. It provides an overview of the economic analysis of child care as an institution available to the working poor, and discusses generally the availability of child care, including factors encouraging or discouraging the private supply of child care. On more specific issues, the paper discusses child care services as utilized by Seattle and Denver income maintenance experiment (SIME/DIME) participants and the policy implications of changes in the pattern of child care choices attributable to a negative income tax (NIT) plan.

THEORETICAL FRAMEWORK

From an economic viewpoint, child care can be described as an industry made up of a large number of sellers of a differentiated product (by location, type of service, etc.). Supply and demand relationships can then be discussed, and price, quantity and quality adjustments in the child care market can be examined. This limits our viewpoint, of course, but it permits an understanding of price elasticities of supply and demand and how these affect the price of child care services.²

Although there is no theoretical model applicable to child care as an industry, we can identify important issues in the economics of child care. First, the analysis should consider a number of factors which affect the actual amount of child care provided as opposed to the amount for which payment is made. For example, the analysis must take into account that day care providers often care for their own children as well as non-related children. This free child care, although valuable, is often ignored since no money is exchanged.³ Also, since child care has both custodial and educational/developmental elements, analysis of any market adjustment process requires decomposing price, quantity and quality effects by both of these components. Another factor to consider in determining the economic value of care is the availability of unpaid volunteers, especially in centers, and the "free" use of a home or other facilities, where applicable. We have measured the value of the volunteer's services by determining the earnings the volunteer could have obtained through paid market activities, and if the volunteer's child is also enrolled free in the center, as often occurs, we have adjusted the market wage.

Second, because child care cost reduces the net wage of a working parent, it is a critical variable in that parent's decision to work. While

a subsidy would be expected to have some effect on this decision, we need to know more about the extent of the effect, as well as the impact of the subsidy on demand, and the nature of the industry's adjustment to an increased demand.

Third, this analysis should consider the provider's entry barriers, such as licensing requirements, zoning restrictions and capital needs. High barriers make an industry less competitive and therefore less efficient in resource allocation. Attention should be given to whether the various types of child care providers operate at full capacity.

Fourth, pricing policy must be understood, taking into consideration both quantity and quality of services provided. An adequate review of pricing practices would give us some insight into the extent of price competition in the industry, and we could better estimate how subsidies for child care would affect the price and the supply of services available to the poor.

Demand⁴

In order to estimate the demand for child care, an econometric model was specified in which the choice of type of child care depends on price, income, and other socioeconomic characteristics of the family. In the empirical analysis, using data from SIME/DIME interviews, three modes of child care were distinguished: (1) formal market care (centers and licensed family day care homes), (2) informal market care (all other paid forms of child care), and (3) non-market care (all non-paid forms). A competitive child care market was assumed and variation in prices was measured in a

cross section by differences in subsidy rates available to families.

In general the coefficients of the model used conformed to our prior expectations. The most important finding from the coefficients was that subsidization of market forms of child care leads to a reduction in the utilization of non-market care. This result was statistically significant in Denver and Seattle. Looking at the coefficients of the demand model is useful in determining the direction of the effect of any given variable on modal choice. However, the coefficients do not provide us with direct information about the size of the impact. In order to determine that effect, a series of probability estimates using those coefficients were evaluated.

Several important conclusions emerged from these probability estimates. The most important to us was that families appear to be very sensitive to price. Let us assume that the market price of child care is \$5.00 per day,⁵ that the subsidy rate under the Federal income tax is 16%,⁶ and that the subsidy rate under the SIME/DIME is 40%. As a result of these subsidies, the price of child care to the working parent is \$4.20 under the federal income tax scheme and \$3.00 under SIME/DIME. Shifting from the federal tax program to income maintenance increases the utilization of market care 14% in Denver and 18% in Seattle. These correspond to price elasticities of about -1.36 and -2.86 respectively. While the effects are about equally divided between formal and informal care, the relative impact is much greater for formal care. The estimated elasticities for formal care are -6.99 in Denver and -4.08 in Seattle. For informal care, the

estimated elasticities are $-.85$ in Denver and -2.40 in Seattle.⁷

We also found that the demand for market care increased with the mother's wage rate. For example, an increase in the hourly wage of about \$2.00 increases the utilization rate of market care by 28%. This corresponds to a wage elasticity of about $+0.91$. Again the elasticity is greater for formal care, with a wage elasticity of $+2.5$, than it is for informal care, where it is $+0.7$.

Third, we found that the demand for market child care varies systematically with family structure. In families with only younger children (under 5 years of age) the utilization rate of market care is estimated to be about 70%. In families with only older children (between the ages of 6 and 12), the utilization rate is estimated to be about 50%. If a teenager or other adult is present in the household, the estimated utilization rate of market care drops to 20%.

Finally, we found that the utilization rates differ by ethnic group. In particular, blacks exhibit a higher utilization rate of formal market care than do whites or Chicanos. This may be somewhat of an artificial effect, however, attributable to institutional factors. In many urban centers a large proportion of formal facilities, particularly day care centers, are located in areas heavily populated by black families. Since we have not controlled for supply conditions in our demand analysis, we may be picking up location effects in our ethnic variables.

Supply⁸

Several observations from the Seattle and Denver data with regard to providers of child care should be emphasized. First, the majority of formal

child care providers appear to be part of the regular labor force, whereas a significant portion of the informal providers seem to be (at least temporarily) out of the regular labor force. This means, however, that there is a fairly elastic supply of labor in the informal sector. Since the care given is very similar to work generally done by parents staying home to care for their own children, the supply of informal sector service can be expanded easily, assuming no special academic or experiential achievements are required.

Another important observation is that the proportion of black and Chicano providers in the informal sector in Denver was much greater than in the formal sector. This is also true in Seattle, except for providers in the child's home, where the proportion of the providers who are from minority groups is approximately the same as in the centers. Moreover, the racial/ethnic composition of day care users was approximately the same as that of providers. However, within the centers we found that racial minorities composed a large percentage of the public non-profit staff and children, while they were only a small proportion of the users and staff in the private profit-oriented centers. We concluded that there was no apparent restriction in entry into the field of non-profit day care by minority group members.

Finally, in the sectors for which the data were available, we found that providers and users are more likely to be related in Denver than in Seattle. For the licensed family day care homes, about one-fourth of the children using day care services were related to providers of those services. However, in the informal sectors there was a much larger percentage of providers who were related to the children for whom they provided care in

Denver relative to the percentage in Seattle. For example, almost two-thirds of the unlicensed family day care home operators in Denver were related to the children using their services, whereas in Seattle almost one-third of the unlicensed family day care home operators were related to the children under their care. Furthermore, we found that over four-fifths of the unlicensed Chicano operators in their own homes in Denver provided services for related children. It appears that a more liberal policy toward subsidies in Denver, including related unlicensed providers, resulted in a far greater use of relatives for unlicensed day care.

To measure "quality" of care, we estimated the proportion of time devoted to educational-developmental and custodial type care, based on interviews with providers. We found that providers in the child's home, and unlicensed family day care homes in Seattle devoted approximately one-tenth, and in Denver, one-fifth, of their time to educational/developmental care; and in the licensed family day care homes both in Seattle and Denver about one-fifth of the time devoted to child care was reported as spent in educational/developmental activities. The proportion of the week spent on educational/developmental care was somewhat greater for centers--overall, one-third in Seattle and slightly less in Denver. In Denver there was very little difference between different proprietary types, but in Seattle the public centers reported spending almost half of their time in educational/developmental care, while other centers reported slightly under a third of their time allocated to that type of care.

Another dimension of the service concerns care when a child was ill.

We asked whether a child with a minor illness, other than a simple cold, would be accepted during regular day care hours. Providers in the child's home almost all reported that they would take care of such children. About three-fifths of Seattle unlicensed family day care homes and over two-fifths in Denver said they would provide care for an ill child. Among their licensed counterparts, only about one-fourth to one-third of the providers in either city provided such care. A very small percent of the centers in either city would accept children with a minor illness for care. On the other hand, for the most prevalent of the minor childhood illnesses (a simple cold) almost all providers in both cities said that they would accept such children for regular care. For working mothers, being able to leave their children with a day care provider when the child has a minor illness is an important consideration in maintaining a steady work record.

Concerning restraints on the potential supply of child care services, we found that the licensing procedure does not seem to be a major barrier to entry into the day care industry. The majority of providers waited less than two months to obtain their licenses and few family day care homes spent more than \$100 in complying with licensing requirements. However, there is some indication that the cost of compliance, especially as it concerns the new Title XX child/staff standards, may present a significant financial burden for the private for-profit centers, if enforced. Zoning restrictions may also sometimes present a barrier to entry for centers. Approximately one-third of the centers in both cities and a smaller proportion of family day care homes had to obtain zoning variances. These licensing and zoning

requirements contributed noticeably to the cost of entry into the day care market. However, these are costs which are under the control of the local authorities.

There is also some evidence that the child care markets in Seattle and Denver operate with excess capacity, despite the existence of waiting lists at many facilities. In general, we found that there was about a 15-20% under-utilization of measured capacity. This may indicate some frictions in the child care market due perhaps to parents' desires for a supplier who fits their special needs or to differences in the type of care and hours. We did find that care was more difficult to obtain for very young children, and for children with any but the most routine illness. We found that a substantial proportion of the children cared for by licensed and unlicensed family day care homes in both cities, as well as public centers in Seattle, are toddlers under age two. Only a small percentage of the children cared for in other sectors are toddlers. Yet another possible reason for friction in the day care market is the lack of widespread information about available suppliers. Although both cities have free referral services, only 10 to 25% of enrollments result from the use of such service.

For centers a relationship was also found between subsidies and revenues. For example, in Seattle, less than 20% of those centers enrolling fewer than half partially or fully subsidized children had gross annual revenue per child that was more than \$900. Of the centers with more than half the children on a subsidy, two-thirds had gross revenue per child of \$900 or more. This relationship holds for both cities and is somewhat

more pronounced in Denver. This finding may indicate that differential pricing according to subsidy status may exist, or that public centers, which have a much larger proportion of their users subsidized, pay higher wages. The data also suggests a possibility that revenue from subsidized children is more reliable than payment from parents, and is more likely to be made even if the child is absent for a short period. Thus the subsidized children yielded higher average revenues over the period of a year.

Analysis of profits and expenses revealed a wide range of earnings per child. For unlicensed providers, this was \$20-30 monthly, and for licensed family day care home operators, \$42 per month. For centers there is a substantial increase in the average gross monthly earnings per child,⁹ with Denver public non-profit centers receiving an average of over \$200 per month per child. The average fees paid to all providers ranged from 45 to 63 cents per hour in Seattle and 33 to 60 cents in Denver. In both cities over 90% of all children were cared for at a rate of less than 75 cents per hour.

We found that the annual cost per currently enrolled child in public centers was from 2-1/2 times to 4 times higher than for other centers in Seattle, and from 2 to 3-1/4 times greater than for other centers in Denver. The cost discrepancy appears to be due to three main factors: (1) a lower average ratio of children to staff in public centers, (2) a somewhat higher average number of hours worked per week by public center employees, and (3) a much higher average hourly wage received by public center employees.

We also estimated cost equations for all child care sectors, separately identifying the custodial and educational/developmental components of child

care costs. To permit comparisons, a cost based on a standard forty-hour week was calculated. All remaining variation in cost was attributed to variation in quality. Regression analysis revealed costs are positively associated with education of the provider and negatively associated with the average capacity utilization rate¹⁰ for providers in their own or the childrens' homes. In both Seattle and Denver, providers in the child's home have the highest cost while licensed family day care homes cost the least; the cost for unlicensed family day care homes is also small. Among day care centers we found that cost is positively associated with the average level of education of staff members. For each additional year of education, the cost per child increases by almost two dollars per week. We also found that for-profit day care centers have a significantly higher cost than private non-profit centers, holding constant all other characteristics of the centers. We did not, however, find a significant relationship between cost and the percentage of time spent in developmental activities.

POLICY IMPLICATIONS

From this analysis we are able to provide some answers to important policy questions. For example, as noted, we found some barriers to entry in the child care market, but they were not substantial ones. Regulations could be simplified and procedures streamlined if the decision was made to increase the availability of day care. For example, in Denver there are a number of different agencies involved in the licensing process, including health, sanitation, zoning, building, and fire. These somewhat overlapping jurisdictions delay the licensing procedure and most certainly impose an additional, if only psychic, cost to the potential entrant into the day care market.

We found regulation of the child care centers in both cities seemed moderately successful. Yet there appeared to be a substantial number of family day care homes in both cities that were unlicensed and therefore unregulated. There were fewer unlicensed homes in Seattle, where licensing authorities were more diligent.

The evidence from our survey is especially equivocal on the efficacy of direct subsidies. Our data on subsidies is for 1973, while our cost data is for 1974. However, if it can be assumed that subsidy levels remained relatively fixed for the two years, then the survey indicates that direct subsidies are not an efficient means of reducing costs to users of day care. This does not necessarily argue against the use of such subsidies, as it may be that they were spent to upgrade the quality of the service provided. However, the same result could be obtained by indirect subsidies to users, combined with greater regulation. Such a policy would give more control to parents, and so would seem to be preferable.

Finally, using data from our 1974 survey we were able to estimate the cost to the child care industry of meeting current and proposed federal standards for the provision of licensed child care services. Estimates were made for centers and family day care homes. We found that between 43 and 59% of the private for-profit centers in our sample were not in compliance with the federal standards. For those centers, cost of compliance would be largest, going up by 5 to 27%, depending on whether it was in Seattle or Denver or the specific set of assumptions used with regard to the relevant standards.

For private non-profit day care centers the cost of compliance ranges from 2 to 13%; between 15 and 55% of those centers are not in compliance. For public day care centers, the cost of compliance is about 8%; between 10 and 38% of those centers are not in compliance. For family day care homes, we estimated the cost of compliance in terms of the number of additional family day care homes needed to serve the same number of children while complying with the newly legislated child/staff standards. In Seattle an increase in supply of about 8% is required; in Denver an increase in supply of 6% is required. A significant proportion of licensed family day care homes in Seattle and Denver are not in compliance with existing and proposed standards.

To demonstrate the potential national implications of compliance with the federal child care standards, we extrapolated the Seattle and Denver cost estimates for centers to a national total. We found that under one set of assumptions, compliance cost nationally would range from \$31 million to \$51 million. Under another set of assumptions, the cost of compliance would range from \$79 million to a high of \$110 million.

One possible reform in the area is subsidization of child care services. This is likely to increase labor force participation rates for the primary parent, as it increases the net wage. It is also likely to change the pattern of usage of child care providers. Unpaid forms presumably would not be subsidized, and would perhaps give way to paid forms. Depending on the policy choices made, centers or other forms may experience new growth, although at present less formal modes seem to be preferred by parents. A subsidy program, if adopted, would also have to deal with the

rules on care of ill children, minimum wage laws and other standards raising costs, and similar issues raised in this paper.

NOTES

1. For an excellent review of some notable exceptions, see C. Russell Hill, "The Child Market: A Review of the Evidence and Implications for Federal Policy", (a report presented to DHEW, OS/ASPE/SSHD, January 1977).
2. The total cost of child care can be divided into three components: the price per unit of child care services (or the quality adjusted price), the quality of child care (or child care services provided per hour of care), and the quantity of child care (or the number of hours of child care provided). The economics of child care involves understanding of how public policy affects each of these three components.
3. In economic terms the value of unpaid or nonmarket care is called the shadow price of nonmarket care.
4. For a more rigorous exposition concerned with the demand for child care services, see Philip K. Robins and Robert G. Spiegelman, "Substitution Among Child Care Modes and the Effects of a Child Care Subsidy Program", in Philip K. Robins and Samuel Weiner eds., Child Care and Public Policy: Studies of the Economic Issues, D. C. Heath & Co., Lexington, Mass., 1978.
5. This is the average cost of child care experienced by families in our sample during the period studied.
6. This is an overestimate to the extent that families with child care expenses do not itemize their deductions or are ineligible for a Federal income tax subsidy because the expenses are incurred for care by a relative (which are not deductible).
7. It should be pointed out that these price elasticities may be overestimated to the extent that our price variable is measuring effects of being in the experiment rather than to changes in the price of child care. A similar upward bias exists if we have overestimated the size of the child care subsidy under the federal income tax. For example, under the extreme assumption of a zero federal income tax subsidy rate, the price elasticities for market care are $-.97$ in Denver and -2.04 in Seattle. On the other hand, if child care subsidies are available (other than in the federal income tax) the price elasticities are underestimated. None of these biases are believed large enough to qualitatively affect our estimates.

8. For a more rigorous and extensive discussion of the supply of child care services, see Samuel Weiner, "The Child Care Market in Seattle and Denver" and Arden Hall, "Estimating Cost Equations for Day Care" in Philip K. Robins and Samuel Weiner eds., Child Care and Public Policy: Studies of the Economic Issues, (D. C. Heath & Co., Lexington, Mass. 1978).
9. For centers we really refer to gross revenue since subsidies and donations are included; whereas for the other sectors there is very little involved other than earnings.
10. Since cost is measured on a full-time equivalent basis, the negative relationship between the average utilization rate and cost suggests that there are economics of scale in the provision of these types of child care services.

DISCUSSION

BETH HARRIS: How did you come up with your figures on under-utilization of day care providers?

WEINER: For example, for centers we estimated the capacity and the full-time equivalent enrollment and divided one by the other; the proportion was the rate of utilization.

HARRIS: Did you include the hired staff?

WEINER: Yes, that's correct. For centers, for example, there are requirements for the staff/child ratio. If it is 6 to 1 and you have only 100 children with 30 staff members, then you are under-utilizing. However, this may simply be an indication of much higher quality service.

HARRIS: Yes, I question the use of that figure under those circumstances. Do you use it also in terms of day care homes?

WEINER: Yes. It was much more difficult in day care homes because one child would come for two hours, one for four hours.

SUSANNE MARTIN: There are studies other than this one that report that day care centers are under-utilized: this should be of concern to policy makers. The majority of public funds for child care have been allocated to institutional care rather than the informal types of care.

WEINER: I have an interesting analogy. When you look at industry--the steel industry or the auto industry--some normally operate at about 85% of capacity. This is about what we find in the day care industry, about 85% utilization. The reason industry adopts 85% utilization is because it is more efficient. This may also be true for day care; we haven't got the data yet to confirm it. Under-utilization is not a derogatory term.

HARRIS: I strongly disagree with your conclusions about the best way to subsidize child care, and some of your assumptions about the preferred method of child care. I have read other national surveys and it doesn't follow that just because people are using a certain kind of care that that is the preferred method. And if you don't directly subsidize centers, some will have to close; this is especially true for centers among low-income populations.

WEINER: If it's true that people would probably prefer center care, if you had indirect subsidy--some sort of a voucher system--parents would direct the subsidy to their preferred choice, which then would be center care.

JIM ANDERSON: We don't have a national child care policy, although we have been trying for the last several years to develop one. I wonder if we seem to be going through so much agony formulating a national child care policy because we are not all talking about the same thing. Your paper indicated

that different economic groups seem to use different levels of child care-- such as in-home care or licensed home care for families with two employed parents, whereas the subsidized day care centers, where I think there is more developmental work going on, were utilized more by children who were economically or educationally deprived. These are two different things and perhaps we ought to have both.

WEINER: I agree, and I'm not sure what to say about it. I would add, parenthetically, that Senator Cranston is sponsoring a Comprehensive Child Development Act, which may do this to some extent. From my conversation with his staff, I believe that he strongly desires some new direction.

ANDERSON: In today's world, at least in Washington, "comprehensive" is the key word. If you want to say something good about something you want to do, you use the word "comprehensive". Maybe that's a problem here. Maybe we are trying to be too comprehensive.

MARY JANE CRONIN: On that point, it does seem that whatever federal policy that has emerged in the child care area has to be loaded onto the day care scene. If you really are considering the separate components of child welfare services that you want, it shouldn't be connected to day care. For example, if you are making choices about where to put your child welfare dollars, why take children of employed mothers--and from probably the most functional families--and put all of your limited child welfare resources there? It seems to me the federal policy should address the most seriously deficient in terms of child welfare and find some vehicle for reaching out to those children.

IMPACT OF FUTURE INCOME UNCERTAINTY
ON SAVING DECISIONS

by

Matthew Black
Economist, MPR

Unexpected changes in income may create psychological and economic hardship for families which varies by the magnitude, frequency and sources of the unexpected income variations. Unanticipated falls in income flows may leave households over-extended with regard to debt obligations and disappointed in not achieving consumption, saving, and leisure goals. Unexpected rises in income may have the opposite effect. Over time we might expect households to alter their behavior in response to previously unstable income patterns. In addition, we might expect families to modify their future behavior according to their perceptions of future income instability. The purpose of this paper is to explore the phenomenon of future income uncertainty by analyzing its repercussions on family decisions to accumulate financial assets and non-mortgage debts.

Consumer liquidity and credit liability have direct implications for monetary and fiscal policy. Higher than normal levels of income uncertainty might reasonably be expected in periods when the national economy is suffering from recessionary pressures and/or experiencing higher than expected rates of inflation. Both situations may erode consumer confidence in their expected flow of future income or the real purchasing power of subsequent income. If greater uncertainty encourages greater accumulations in liquid assets and/or decreases in non-mortgage debts by families, it can

have a profound effect on the national economy. Moreover, what appears rational and prudent for the individual may not be good for the economy when all individuals pursue such activity.¹ In a recessionary climate, for example, a drop in aggregate consumer demand for commodities in reaction to greater income uncertainty (i.e., an increase in liquid assets and/or decrease in debt accumulations) will tend to accelerate the decline in household income and increase unemployment.

Future income uncertainty also has implications for a family's standard of living. On the presumption that households are risk averse,² the distribution of welfare, as implied by the distribution of income across families, may be exacerbated or ameliorated according to the incidence of future income uncertainty over income groups. An effective public transfer system should not only seek to provide adequate levels of support, but also to be responsive to changes in a family's economic status.³ To the extent that a welfare system adjusts to these changes in a timely fashion, then both the economic hardship facing clients and the costs of providing assistance are reduced.

This paper develops an empirical approximation of future income uncertainty in order to analyze the key implications of theoretical models of saving decisions under uncertainty. The empirical analysis is based on interview data collected for the control households in the Seattle and Denver Income Maintenance Experiments (SIME/DIME) over the first 19 months of the experiment (i.e., the first three periodic interviews).⁴

MEASUREMENT OF FUTURE INCOME UNCERTAINTY

A central element in decision-making is the expected outcome of the decision. In addition, the uncertainty surrounding expectations will play a role in the decision. For example, a risk-averse individual may not select the alternative offering the highest expected value if its reliability is low. Thus, some economists utilize both the mean and the variance associated with variables in predicting economic behavior, assuming that the smaller the observed variation about the mean, the more confidence a family has in its future expectations.

Future income uncertainty arises because the sign and magnitude of errors associated with forecasting income are unknown to households. To operationalize future income uncertainty, it is assumed that a household's income expectations are related to its recent experience in predicting past income. The emphasis of the analysis is thus on intertemporal rather than interpersonal income uncertainty. The latter reflects unexpected changes relative to cohort income while the former refers to unanticipated changes that are household specific. In particular, intra-year fluctuation in monthly income is examined as opposed to inter-year variation. For the purpose of this paper, income uncertainty due to uncertainty about life expectancy is not considered.⁵

Previous approximations of future income uncertainty have used either the variance or standard error of estimate from a single time-series regression on past income observations.⁶ These approaches assume that a family's uncertainty about the future is based on recent instability with respect to

a known income stream over the recent past. Because future income uncertainty arises from the unknown accuracy of future income predictions, it is as reasonable to assume, as we do, that expectations about income are related to past errors in forecasting income. The more accurate were a household's previous predictions, the more confident is it expected to be in projecting subsequent income.

Within this framework, a number of forecasting periods can be simulated for each family. The prediction errors associated with each forecasting problem are then combined and used as a proxy for a family unit's perception of its future income uncertainty. In particular, monthly forecasts are made on the basis of a moving average of the past seven months of income, adjusted for a linear time trend estimated for that lagged period.⁷

An important weakness is inherent in this approximation. By using an "adaptive expectations" hypothesis of how units predict their next period's income (i.e., formed by a moving, weighted-average of past incomes), the measure of future income uncertainty may ignore relevant information about the time path of income. The adaptive expectations hypothesis presumes that because units cannot see into the future with much accuracy, they look backward in time and use historical information to project future values of income. This results in a gradual correction of past forecast errors.

It sometimes may be rational for units to correct for certain factors in advance. For example, a plan to quit a job will undoubtedly be considered to predict future income. However, a voluntary spell of not working (e.g., pregnancy, return to school) can generate some degree of future income

uncertainty, since the implications of the income loss may not have been anticipated, and return to work may be dependent on factors beyond an individual's control. Hence, it is exceedingly difficult, if not misleading, to net out changes in income that may appear to be planned. Thus, in this paper we create a measure of future income uncertainty based on observed income alone. We recognize it to be an imperfect measure because it does not incorporate a household's subjective appraisal of how observed income fluctuations affect their future uncertainty.⁸

THEORETICAL IMPLICATIONS OF FUTURE INCOME UNCERTAINTY FOR SAVING DECISIONS

The most intuitively appealing hypothesis of the effects of future income uncertainty on saving decisions is that the greater it is, the more it encourages savings and discourages consumption.⁹ The reason would be twofold: First, increased assets can be invested in low-risk interest bearing assets (e.g., savings accounts) which would partially offset the less certain income expected in the next period. Second, the increased assets can serve as a financial cushion with which to meet desired consumption needs in the future.

Leland¹⁰ has developed a two-period saving/uncertainty model where income in the first period is known with certainty while that in the second has an unknown variance about its expected value. Hence, consumption/saving decisions are made a function of a fixed current income and a subjective probability distribution with respect to the second period's flow. Leland seeks to derive behavioral conditions to insure that an increase in future income uncertainty (analytically represented as the unknown variance of

next period's income) will result in an increase in savings in the first period. He finds that a sufficient behavioral response is that an individual becomes "less risk averse in a variable as that variable becomes increasingly predominant in a constant utility bundle."¹¹ In Leland's two-period case, the greater the ratio of income expected in period two compared to period one, the less averse is an individual to a given level of expected risk. This implies that the impact of future income uncertainty on levels of current saving should decline as this ratio increases. Our empirical analysis attempts to verify the appropriateness of this implication.

Sandmo¹² derives a similar behavioral prerequisite to obtain a positive relation between future uncertainty and current saving. He labels this condition the hypothesis of "decreasing temporal risk aversion." Sandmo also makes a distinction between sources of income, with degree of uncertainty varying across sources. He finds that an increase in uncertainty about income derived from work elicits an increase in saving (also subject to some income uncertainty) to offset the larger but unknown variance in the labor-produced income, by increasing the expected value of asset income (the precautionary effect). Our analysis will focus on the precautionary effect by excluding asset income from total household income. As low-income families in SIME/DIME have few capital investments, they do not provide a good opportunity to estimate the net impact of the asset income uncertainty on saving decisions.

IMPACT OF FUTURE INCOME UNCERTAINTY ON
FINANCIAL ASSET AND DEBT ACCUMULATIONS

Estimating Strategy

Our analysis focuses on the change in families' financial assets (cash, bank deposits, securities) and the change in non-mortgage debts (excluding real estate mortgages). This section seeks to explain variation in these stock changes that occurred between the first and third periodic interviews of SIME/DIME families. Separate but parallel analyses are conducted for financial asset and debt accumulations. As the impact of future income uncertainty may affect changes in liquid assets and debt changes differently, we use separate regression equations. In both analyses the two change variables are regressed on identical sets of explanatory variables.¹³

Families are assumed to have desired stocks of financial assets and non-mortgage debts, given their demographic and economic characteristics. Deviations from these norms are expected to encourage saving decisions that bring actual stocks into closer accord with desired stocks. Differences between actual and desired levels as of the first periodic interview are imputed on the basis of preliminary regressions reported in the longer version of this paper.¹⁴

The first step investigates the additive impact of future income uncertainty on asset and debt changes, along with the effects of other independent variables (such as permanent and transitory income,¹⁵ initial differences between actual and desired asset and debt stocks, and demographic variables), and then we estimate the extent to which the level of imputed permanent

income influences the impact of future income uncertainty.

The disturbance terms in the two equations are likely to be heteroskedastic. Thus, an ordinary-least-squares regression (OLS) will yield inefficient estimates if the problem is left uncorrected. This problem is dealt with in a preliminary stage on the assumption that the variance of the disturbance term is positively related to the square of household income. OLS can then be used for the amended asset/debt change equations to obtain efficient, generalized-least-squares estimators.¹⁶

We also had to account for the distribution of interview dates across time. Households were interviewed over an interval of approximately six months rather than simultaneously. This problem is minor for most time-variant variables such as income, family size, and spells of unemployment because retrospective data was collected. However, interview dates present an acute problem for our financial asset and debt accumulation variables which are equal to the change in stocks between the first and third interview dates. There can be from six to 18 months between these interviews. Our solution is to divide the observed changes by the number of intervening days and to multiply by 365 days--deriving a measure of annual asset/debt accumulation. This assumes that accumulations occur at a constant rate over time. Errors in our assumption are embedded in the disturbance terms, but they will not affect the unbiasedness and consistency of our estimators.

Discussion of Empirical Results

The regression equations of asset and debt accumulations are applied separately to samples of families headed by two parents or a single female.

Both groups consist of observations pooled from the SIME/DIME data files. Descriptive statistics for the two samples are presented in Table 1. The regression results for changes in assets are in Table 2 while those for the non-mortgage debt accumulation are in Table 3. Results from both equations are given.

In general, there does not appear to be any statistically significant difference in asset accumulation by race or ethnicity. Nor does the average size of the household during the analysis period appear to affect asset decisions. Using equation 1, we find that the age of the male head has a negative influence on asset accumulation for two-parent families. For each year of age, two dollars less is saved each year, suggesting that units save more in their earlier years, but tend to save less with the changing financial needs over the life cycle as family responsibilities decline. There also appears to be some site difference as two-parent families appear to accumulate fewer savings in Denver, compared to Seattle.

The pervasively strong stock-adjustment effect is immediately apparent. The imputed difference between actual and desired assets is intended to measure the extent to which a household's initial level of assets varies from what it is expected to desire, given its demographic characteristics, human capital, and normal income. The measure partly controls for asset preferences and the level of transaction balances desired by particular cohort groups. The imputed difference between actual and desired non-mortgage debts has an analogous interpretation.

For both samples, liquid assets are adjusted by approximately 15 cents for every dollar that the initial stocks deviate from the cohort average.

TABLE 1

DESCRIPTIVE STATISTICS

(Sample Proportions or Means with Standard Deviations in Parentheses)

Variable	Households Headed by Two Parents	Households Headed by Single Female
Black	.34	.46
Chicano	.23	.14
White	.43	.40
Denver Site	.61	.45
Age, years*	35.0 (11.9)	31.0 (10.8)
Education, years*	10.6 (3.6)	10.4 (3.5)
Mean Family Size	4.0 (1.5)	2.8 (1.3)
Homeowner (1st periodic)	.30	.15
Windfall Income**	\$ 14 (153)	\$ 13 (232)
Observed Income	\$ 5720 (3065)	\$ 3186 (1764)
Permanent Income	\$ 8010 (1478)	\$ 4475 (1166)
Transitory Income	\$-2177 (2471)	\$-1244 (1456)
Change in Assets	\$ 34 (323)	\$ 1 (146)
Change in Debts	\$ -14 (984)	\$ -109 (656)
Future Income Uncertainty	\$ 106 (71)	\$ 52 (42)
Number of Observations	603	389

*Age and education are with respect to the male head in two parent households.

**All monetary variables are measured as annual equivalents.

TABLE 2

REGRESSION COEFFICIENTS ON CHANGE IN FINANCIAL ASSETS
(Absolute Value of t-ratios in Parentheses)

Variable	Two-Parent Households n=603				Single Parent Households n=389			
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Constant	167.91	-77.139	27.988	2.465				
Black ^d	-16.704	(.559)	-10.170	(.341)	20.299	(1.27)	20.855	(1.30)
Chicano ^d	15.065	(.396)	22.125	(.581)	11.238	(.47)	10.305	(.43)
Age	-2.218*	(1.92)	-1.801	(1.549)	-.486	(.64)	-.517	(.68)
Mean Family Size	-3.531	(.344)	-4.057	(.397)	-2.214	(.34)	-2.310	(.35)
Actual Desired Assets	-.155***	(7.25)	-.154***	(7.24)	-.142***	(11.24)	-.142***	(11.19)
Actual Desired Debts	-.018	(1.62)	-.018	(1.62)	-.038***	(4.14)	-.038***	(4.09)
Windfall Income	.153*	(1.86)	.149*	(1.82)	.038	(1.34)	.037	(1.27)
Site ^d (Denver = 1)	-96.364***	(3.32)	-106.145***	(3.63)	-10.759	(.69)	-9.292	(.59)
Normal Income	.007	(.71)	.035**	(2.24)	-.007	(.82)	-.002	(.146)
Transitory Income	.023***	(4.21)	.023***	(4.19)	.014***	(2.74)	.014***	(2.69)
Future Income Uncertainty	-.126	(.66)	2.218**	(2.20)	.321*	(1.92)	.754	(1.12)
Future Income Uncertainty x Normal Income	---		-.0003**	(2.37)	---		-.0001	(.66)
F-Ratio for the Regression	7.739		7.615		13.299		12.209	
R ²	.07		.08		.22		.23	

^dDummy variable

*Significant at 10% level

**Significant at 5% level

***Significant at 1% level

TABLE 3

REGRESSION COEFFICIENTS ON CHANGE IN NON-MORTGAGE DEBTS
(Absolute Value of t-ratios in Parentheses)

Variable	Two-Parent Households				Single-Parent Households			
	(1)	n=603	(2)		(1)	n=389	(2)	
Constant	-42.464		-70.184		422.52		-29.380	
Black ^d	159.774**	(2.09)	160.513**	(2.09)	-93.920*	(1.70)	-84.076	(1.55)
Chicano ^d	-18.342	(.19)	-17.544	(.18)	-89.758	(1.10)	-106.274	(1.32)
Age	-.054	(.02)	-.007	(.003)	2.005	(.77)	1.467	(.57)
Mean Family Size	28.707	(1.09)	28.647	(1.09)	-38.000*	(1.69)	-39.706*	(1.79)
Actual Desired Assets	-.028	(.52)	-.020	(.513)	-.013	(.30)	-.006	(.12)
Actual Desired Debts	-.427***	(14.87)	-.427***	(14.86)	-.639***	(20.26)	-.632***	(20.290)
Windfall Income	.146	(.69)	.146	(.69)	.100	(1.02)	.067	(.69)
Site ^d (Denver = 1)	681.158***	(9.16)	680.052***	(1.09)	275.333***	(5.15)	301.307***	(5.66)
Normal Income	-.071***	(2.65)	-.068*	(1.69)	-.104***	(3.38)	-.005	(.13)
Transitory Income	.057***	(4.03)	.057***	(4.03)	.091***	(5.03)	.087***	(4.90)
Future Income Uncertainty	.026	(.05)	.291	(.11)	-1.118*	(1.94)	6.548***	(2.88)
Future Income Uncertainty x Normal Income	---		-0.00	(.10)	---		-.0016***	(3.48)
F-Ratio for the Regression	33.448		30.610		46.881		45.248	
R ²	.33		.33		.41		.42	

^dDummy variable

*Significant at 10% level

**Significant at 5% level

***Significant at 1% level

Households also exhibit a tendency to reduce financial assets by 2-4 cents for every dollar that debts exceed the initial desired level. This adjustment probably represents an effort to use liquid assets to pay off some existing debts when debts become undesirably large. Alternatively, fewer debts seem to lead to greater asset accumulations because either there is no pressing need to pay off existing debts and/or liquid assets are being accumulated in anticipation of future purchases. The sign and significance of the stock adjustment variables are expected because the level of initial stocks is contained in the dependent variable as a benchmark against which the change is calculated. The consistency of our estimators, however, is not endangered because the initial stocks are predetermined in our estimating framework.

Future income uncertainty is estimated to have a significant impact on asset accumulation. For single female-headed households, the greater the imputed value of future income uncertainty, the greater is the change in liquid assets, but the saving-uncertainty relationship does not appear to vary by earning capacity.

The opposite effects are found for households headed by two parents. Future uncertainty does not have a significant impact on asset changes, on average. However, when future income uncertainty is interacted with permanent income, we find substantial empirical corroboration of our a priori expectations. First, greater uncertainty encourages a larger accumulation of liquid assets than would ordinarily occur (which is consistent with the precautionary saving-uncertainty effect). Second, the increase in savings as a result of greater uncertainty diminishes as permanent income rises,

all other things being equal. This finding offers some empirical verification that families are less averse to risk as expected future income rises relative to current income, insofar as reflected by asset accumulation. The lack of a significant effect for single-female-headed households may be due to their lower range of permanent income.

The coefficient on windfalls is positive and significant for two-parent households. The large size of the coefficient supports an implication of the permanent income hypothesis, namely that consumers tend to save a larger fraction out of unexpected income receipts than they do out of regular income flow. This relationship, once again, does not appear to hold for the one-parent family.¹⁷

Change in non-mortgage debts are treated in Table 3. For families headed by women, blacks appear to accumulate fewer debts than whites. On the other hand, black two-parent families seem to incur more debt. For the single parent sample, family size tends to have a depressing effect on debt accumulation, but the relationship does not hold for two-parent families. Again, site differences are significant, with Denver residents accumulating considerably more debts than similar households in Seattle. It is possible that the stronger local economy in Denver created a more stable and confident climate that encouraged or permitted greater debt holdings and fewer liquid assets than in Seattle.

As shown in Table 3, future income uncertainty does not have a significant impact on debt accumulation for the two-parent sample. Its effect on saving decisions appears to be channeled through changes in financial assets. Single-parent households, however, appear to adjust their debt

holdings in response to increased uncertainty. This is consistent with the precautionary effect predicted by theoretical models of saving under uncertainty in that reduced debts in the current period result in lower repayments in subsequent periods characterized by greater income uncertainty.

When future income uncertainty is interacted with permanent income, the results appear to be contradictory. Both estimated coefficients have the opposite of the expected signs. An explanation may be that the additive effect is capturing an ex post reaction to increased levels in measured future income uncertainty in the sense that low-income households may be forced to incur debts to meet consumption needs when income falls unexpectedly. However, as the interaction term suggests, the need to borrow, for a given level of future income uncertainty, diminishes as earning capacity increases.

The stock adjustment variable again plays a major role. One- and two-parent households reduce debt holdings by approximately 64 and 43%, respectively, of the amount by which initial debts exceed the imputed desired amounts. Given prior information on a household's current non-mortgage liabilities, we can predict rather closely the subsequent changes in debts by comparing the initial level with our estimates of what amount of debt is typical for that household's cohort.

In sum, the empirical results offer some useful insights into household decisions regarding financial asset and non-mortgage debt accumulations. The most powerful explanatory variables in our analysis are the imputed differences between actual and desired asset/debt levels. Households display a very strong tendency to adjust both in order to bring them into closer accord

with what we estimate to be their desired levels. Changes in assets and debts are also sensitive to transitory income but there is no consistently significant relationship with normal or permanent income.

IMPLICATIONS

Our estimates confirm the major implications of theoretical models of household saving decisions under conditions of uncertainty. In general, households tend to accumulate more assets and fewer debts when they perceive greater future income uncertainty. The results suggest that households may act in a precautionary manner to increase their net worth in anticipation of less certain flows of future income.

These effects are important to the national economy. In a recessionary climate that leads to greater uncertainty, households tend to consume less, which subsequently causes a drop in aggregate demand and acts to accentuate the downward drift of the economy. Other analysis of ours indicates that the existing welfare system tends to counteract this feedback chain by providing households with an income floor which lessens the magnitude of income fluctuations over time.¹⁸ Hence, transfer payments not only provide a means of income support, but they tend to bolster the overall level of economic activity indirectly by reducing future income uncertainty.

NOTES

1. See Thomas C. Schelling, "On the Ecology of Micromotives," Public Interest 25 (Fall 1971): 61.

2. In this study, we do not distinguish between risk and uncertainty as defined by Frank Knight, where risk refers to situations of known possibilities and uncertainty is with respect to unknown probabilities. Frank H. Knight, Risk, Uncertainty, and Profit, (New York: Hart, Schaffner and Marks, 1921).

3. There are, of course, other features of the welfare system that are relevant in an appraisal, such as work incentives, equity, and costs.
4. We do not distinguish between families and households in this report. The two do not necessarily coincide for there are many cases of multiple families living within a single household. The data, however, do not permit this distinction.
5. Example includes D. Levhari and L. J. Mirman, "Savings and Consumption with an Uncertain Horizon," Journal of Political Economy 85 (1977): 265-281, and L. J. Mirman, "Uncertainty and Optimal Consumption Decisions," Econometrica 39, no. 1 (January 1971): 179-85.
6. Recent empirical investigations on this issue include Jacob M. Benus, "Income Instability: An Empirical Analysis," (Ph.D. dissertation, The University of Michigan, 1974); J. R. Johnson and J. M. Benus, "The Impact of Income Instability on the Consumption of Durable Goods," (Draft Research Memorandum, Stanford Research Institute, 1977); F. S. Mishkin, "Illiquidity, Consumer Durable Expenditure, and Monetary Policy," American Economic Review 66 (September 1976): 642-654; and Raymond J. Uhalde, "An Empirical Analysis of Monthly Income Instability and Its Impacts on Saving and Labor Supply Behavior," (Final Report, Washington, D.C., Mathematica Policy Research, Inc., June 1976).
7. The choice of a seven-month forecasting period is arbitrary. A shorter or longer period could have been used but the measured differences in the respective future income uncertainty would probably have been minor. Using intra-year income fluctuations does enable us to examine uncertainty with respect to short-run income flows and its impact on economic decisions that tend to be short-term, such as consumption, changes in liquid assets, and labor supply. Purchases of durables like cars and houses may be based on longer time horizons which suggests that uncertainty with respect to inter-year income streams may be more relevant.
8. The formula used to measure future income uncertainty are found in M. Black and J. O'Hare, "Determinants of Future Income Uncertainty and its Implications for Savings Decisions," (Mimeo, Mathematica Policy Research, June 1977).
9. See Kenneth J. Arrow, Aspects of the Theory of Risk Bearing (Helsinki: Yrjo Johansson Foundation, 1965); M. J. Block and J. M. Heineke, "The Allocation of Effort Under Uncertainty: The Case of Risk-Averse Behavior," Journal of Political Economy 81 (March/April 1973): 376-385; Jacques H. Dreze and Franco Modigliani, "Consumption Decisions Under Uncertainty," Journal of Economic Theory 5 (1972): 308-335; Hayne E. Leland, "Saving and Uncertainty: The Precautionary Demand for Saving," Quarterly Journal of Economics 82 (1968): 463-473; John W. Pratt, "Risk Aversion in the Small and in the Large," Econometrica 32, no. 1-2 (1964): 122-236; Michael Rothschild and Joseph Stiglitz, "Increasing Risk I: A Definition," Journal of Economic Theory 2

(1970): 225-243; "Increasing Risk II: Its Economic Consequences," Journal of Economic Theory 3 (1971): 66-84; and Agnar Sandmo, "Capital Risk, Consumption and Portfolio Risk," Econometrica 38, no. 4 (October 1969): 586-599.

10. Leland (1968).

11. Leland (1968), p. 469.

12. Sandmo (1969).

13. An argument can be made that the efficiency of our estimation can be enhanced by incorporating system information contained in the covariance matrix of the disturbance terms for the two change equations. In other words, if disturbances in the debt change equation are somehow related to disturbances in the asset change equation, then additional information is available with which to obtain better estimators. Although our two equations may be "seemingly unrelated" it is not sufficient to have information to convey. We must be able to actually take advantage of the relationship between the disturbances. In our situation, however, the set of independent variables are identical. Hence, the GLS coefficients collapse to the OLS coefficients (after correcting for heteroskedasticity) because there is no way to transmit the system information contained in the covariance of the error terms. Because we have no a priori justification for specifying different sets of predictors for the asset/debt change equations, we estimate the two relationships separately.

14. See Black and O'Hare (1977).

15. Transitory income is defined as the difference between actual and imputed permanent income. The latter is estimated from preliminary regressions reported in Black and O'Hare.

16. The reported R^2 's for the equations are not those obtained from the respective OLS regressions on the transformed data (correcting for heteroskedasticity). The dependent variables in the equations have been divided by observed household income and no longer equal the original stock change variables. To circumvent this problem, we used the estimated changes obtained from our regressions and then calculated the squared correlation between the original dependent variables and the fitted values to obtain the respective R^2 's. The F ratios for each equation, however, are with respect to the OLS regressions on the amended analysis variables.

17. Our dependent variables, however, do not permit us to conclude that the marginal propensity to save out of normal income is zero. Our analysis focuses on particular facets of net worth changes that prevent us from ascertaining the effect of normal income on saving behavior in a manner that is consistent with the permanent income hypothesis.

18. See Black and O'Hare (1977).

DISCUSSION

(John O'Hare presented Matthew Black's paper in his absence and answered questions.)

JOHN MC COY: What was the upper level of income that you were looking at?

O'HARE: Well, the experiment was with people at \$9,000 to \$11,000 annual income. However, we did find family incomes that went up to \$14,000. That's right about the national median for 1973--the year we looked at. It's important to note that the earlier research found that as income rose, uncertainty went down, but we found an opposite effect. Possibly earlier work was measuring instability rather than uncertainty, and I believe we have a better tool to measure uncertainty. We found that the original technique that we used was very, very sensitive to, for example, a one month erratic shift in income. Imagine an income stream that is mostly horizontal, but with one month of no income: we would predict a very large error for that month. Under our measure of uncertainty, the error would get progressively smaller. So, we modified that by creating a "tolerance band" for income. That is, if a person is off of normal income by a few dollars, we ignore it. But if he gets outside the "tolerance band" we consider that a critical error, and include it in our measure of future income uncertainty. With this procedure, we still got the same results. We found that as income rose, uncertainty also rose.

RICARDO SPRINGS: Maybe the higher income goes, the more room there is for fluctuations. It can fall further, which could effect uncertainty.

O'HARE: Exactly. This is one of the problems. We tried to correct for that in our estimation, but again, the problem is you have someone who earns \$200 a month, and they drop down 50%; that's a big drop, but it's only \$100.

JOHN W. MOUNT: Weren't you adjusting your tolerance bands to account for that?

O'HARE: Yes.

MICHAEL LINN: How, over what period of time did you draw your sample group?

O'HARE: The pre-enrollment and the first twelve months of the experiment were used.

LINN: The reason I'm asking is because the data from Seattle may well be funny.

O'HARE: Exactly. The paper mentions some site differences, particularly in the debt and asset accumulation. We find that the people in Denver were liable to incur more debts. We explained that it was the stronger economy in Denver which might have made them more liable to incur more debt.

LINN: Now, the other question I had, are you planning to look at income uncertainty with a group that, for instance, has been involved in a transfer program where there is a one-year accounting period versus the welfare one-month accounting period?

O'HARE: As far as looking at people who are on different plans, it is intimately tied with the labor supply. People drop out of the labor force and their income goes down. You really have to control for that, which we did. However, in some preliminary work with the Gary income maintenance experiment, we found that families on the experiment suffered greater income uncertainty. This was disconcerting, and it was contrary to what you would expect. However, we investigated it further, feeling that the results had much to do with the labor supply--not just primary labor supply but secondary and tertiary labor supply. If the husband cuts down a few hours on his work so he can spend more time with the family, or the wife quits her job to take care of the children, the income stream looks more uncertain. One of the things that we want to investigate this next year, and are approaching D.H.E.W. about it, is research on where the labor supply effects are concentrated. Is it concentrated in the primary labor participation or is it concentrated in a reduction in labor supply?

SPRINGS: Do you use the model to predict future income?

O'HARE: Not at all. We use the time series of data to construct a measure of income uncertainty. This measure is related to the ability of the family to predict its income over a yearly period. Families who consistently predict well have virtually zero values of income uncertainty. Families who have an erratic income stream have higher levels of uncertainty. So, we use that measure, that one-dimensional measure of uncertainty in savings and debt equations to find out what effect this had on savings and debt decisions.

MOUNT: Can you use your model to predict when a family's savings are down to a certain level, that they will go back to the labor market?

O'HARE: That was indeed the title of the paper that we did on Gary: "Income Uncertainty and Its Effects on Labor Supply." The results were not statistically strong, but we did find that people with uncertain income did have a tendency to increase their labor supply. In Gary, we also found that larger families experienced relatively smaller levels of uncertainty. The larger the family was, it seemed to be the more the parents knew they had to work, the more stable the income stream was over time. But, again, you have to ask whether we were really measuring uncertainty in that analysis.

MC COY: You indicated that you measured uncertainty looking at income and then looking at the fluctuations of income within a range? So where does uncertainty come in? How would you then translate this fluctuation into uncertainty?

O'HARE: We tried to predict every month what the family's income would be. Then you take the difference between the prediction and what it actually was and you just keep that as a residual component; then you calculate these residuals over the year. Some months we can predict exactly what a family's income will be; other months we'll be way off.

MC COY: So uncertainty is the degree of error between the actual and the estimate?

O'HARE: Essentially. The error band that we used was the absolute value of these residuals. Some, of course, are going to be high, some are going to overshoot, undershoot, and we took twice the standard deviation, if you want to get into technical terms, around the income stream and adjusted the absolute values. Then we just used those that fell outside that stream--the real outliers. And, again, I'll emphasize that this is a technique that we felt captured the flavor of uncertainty a little more than previous research.

MC COY: Having done that, those cases that fell outside those standard deviations, those then you defined as being cases of income uncertainty?

O'HARE: No, we summarized those, and again, we took the mean of those outliers and took the standard error of those; that was an arbitrary decision. It was a quadratic function of these errors. We could have just summed them up (and did in an earlier paper, or at least took the mean error). We could have done any number of things, and believe me, we tried many.

LINN: I don't really see what your prediction tells us, compared to what you might get, were you to ask for a prediction from the individual. Because if I planned to take the summer off, I have no uncertainty about my income.

O'HARE: I agree. The ideal way to do it would be to have the interviewers go to that house every month and ask. But we didn't have the resources, so we used our imperfect measure as the best available.

MC COY: You indicated that the higher the income you find greater uncertainty and where there was greater uncertainty you found a proclivity to save associated with a proclivity to reduce debts. That indicates then the lower the income, the greater is the certainty. You said that was reversed.

LINN: It's hard to diverge from nothing.

O'HARE: That's true, relatively speaking.

MC COY: It has to be. And then as you get down to, say, welfare cases where, as you indicated there is great certainty--a constant source of income. That would lead me to conclude, based on what you said, that the welfare recipient doesn't save and may enter more debt too.

O'HARE: You might be able to say that. But we were estimating changes in debts -- not actual propensities to take on debt.

MC COY: You didn't look at growth in debt?

O'HARE: No, just changes in debt. For example, a 19-year-old, single woman with one young child may have a certain amount of debt that she felt she could handle. We looked at deviations from that amount, not at deviations from the amount of debt a two-parent family would incur.

MC COY: Isn't that like saying we get a correlation of all the people who are not married and had babies and found that there was almost a direct relationship, one to one relationship, between unmarried parents and illegitimate children?

O'HARE: There are actually two relations in the paper that are estimated regarding savings. One looks at just the straight average, overall average result the income uncertainty has on savings behavior, changing assets behavior. And we did find a positive correlation between uncertainty and income. However, we went one step further and interacted uncertainty with savings. Then we interacted uncertainty with income and we found that if the income gets to a certain point, the savings no longer increases. In other words, as the income goes up, families facing greater uncertainty had to save more. However, at very high income levels you find a levelling-off effect.

MC COY: Let's go back to your sample. Your sample dealt with, what were the income ranges that you said were in the sample?

O'HARE: Up to \$14,000.

STELLA B. JONES: Did your extreme cases then cause the relationship which you found?

O'HARE: We don't know. I tend to doubt it because we did try throwing out outliers with just -- there were a couple of cases in there that had extreme amounts of savings and debts computed to them, and the results weren't really changed all that significantly, but I really can't say. But we found a nonlinear relationship between income and savings--debt accumulations.

MC COY: Now, can I infer that people with \$7,000 income--at the bottom end of your income range--are the least likely to save?

O'HARE: They perhaps did not save as much as persons at higher income, but they did save. In fact, we did find that the highest predictor of savings behavior was this adjustment model of bringing savings back in line to what you would expect a person of that cohort, that demographic group, age, education, race, sex, whatever, family size to do. So they did save. However, as income rose, they tended to save more.

MOUNT: Have you tried to see whether the instability measure is correlated somehow with other events in the family's history such as separations, alcoholism or anything else behavioral like that?

O'HARE: Not at all.

LINN: So how do we use your output in the area of welfare policy development?

O'HARE: Economists indicate that there are certain characteristics associated with an uncertain income stream. People have expectations that they like to fulfill. One of the ways to fulfill expectations is to save. A family that has a constant, fluctuating income is going to continually find that their expectations and their hopes are depleted because of an erratic income stream that forces them to continually draw down on savings. It just makes no sense for someone to save if they know that they may have to get rid of their savings in another three or four months. A welfare system that can smooth out, that is responsible enough to smooth out short periods of low or no income and kind of smooth out these rapidly fluctuating income streams would, in my opinion, contribute to the welfare of the family, both psychic welfare and overall community welfare, social welfare.

LINN: Okay. So what you're proposing would be a welfare system in which you would have a very short accounting period.

O'HARE: Certainly.

LINN: Another way you could do it might be to use the income imputed from assets and a longer accounting period so that you don't force people to sell off assets.

O'HARE: Right.

LINN: Well, with an NIT, you don't need it, you don't worry about assets, you just worry about income so that you don't assault assets.

O'HARE: There is an asset test in the NIT. I believe it's \$1500.

LINN: There's not an asset test. One of the things you're saying is you would begin to look at getting away from the asset test then and just impute income to available assets?

O'HARE: Right.

MC COY: Going back to the paper, what you're saying, it seems, is if you establish some kind of floor--not an immodest floor--of income so that it at least wipes out the bottom end of uncertainty and it would reduce the amount that a family is otherwise made to feel obligated to save. If you have that kind of system, not only would you have psychically healthier arrangements for the family itself, but there would be more money since you're not saving the money, you're actually spending and consuming--along with whatever tendencies there would be for this income to impact on inflation.

You've really encouraged spending and you've exacerbated inflation.

O'HARE: I think we're feeling some of that now--exacerbating inflation--but I was in general speaking about the other side of the coin, the fact that when the G.N.P. starts going down, people start losing jobs, they become more uncertain about their income, they again start saving more to hedge against the possibility of losing their income, which, again, throws more people out of work because there's not enough money being spent. So, as far as the macroeconomy goes, there are tremendous implications for uncertainty. We're focusing on the savings/debt aspect of it, but again, if you look at a lot of interesting work now that's being done in inflation is being done on this uncertainty with regard to future price levels.

LINN: The kind of things you're saying lead to some conflicts with other recommendations stemming from SIME/DIME, such as an annual accounting period.

SPRINGS: I don't think anyone has proposed that. That's just what they use for the NIT program. We definitely talked about monthly retrospective reporting, but not a long carryover--three or six months.

ROBERT HOREL: Six months, at the most.

MC COY: I understand from Henry Arron that Congress has considered a year-long accounting period.

LINN: I'm all for a year myself.

SPRINGS: That's tough on families.

LINN: I think you can do a couple of things to help. You bifurcate the program, you give some assistance in money managing and supplement with catastrophe-aid programs.

O'HARE: But one of the things Congress is interested in now is thinking of having a shorter accounting period at the end of the year through the tax system.

MC COY: That was in the sub-committee's bill. Is that least efficient, in terms of coping with uncertainty?

O'HARE: I would say it tends to help it. But no matter how long your accounting period or carryover period is, a family with absolutely no income has to be helped.

MC COY: Even if they haven't been prudent enough to save?

O'HARE: As we know, not everyone is prudent enough to save.

LINN: Then have a categorical program for nonprudent people--the NP program, but it's a very costly category.

MC COY: That's what most states were concerned with in their analysis of the six month or longer accounting periods--the nonprudents that were going to be ineligible for this mainstream program. The states would have to take care of their immediate problem, and they saw an immense cost burden.

VI

ADMINISTRATIVE
IMPLICATIONS

EFFECTS OF INCOME ACCOUNTING AND REPORTING
PERIODS IN TRANSFER PROGRAMS

by

Ricardo C. Springs
Researcher, Policy Studies Division, MPR

In early 1970, a computer simulation model which had the capability of reproducing the income reporting and accounting procedures of government transfer programs was developed. This computerized design, called the Accounting Period Simulation (APS) model, permitted the modeling of both prospective and retrospective income accounting systems, as well as several other hypothetical systems believed to be in the realm of possibility for inclusion in the administration's welfare proposal. It was developed to model the Family Assistance Program (FAP) then under consideration, but it has since been extended to other transfer programs, such as Aid to Families with Dependent Children (AFDC) and Food Stamps. These modeling operations have used data from the Seattle and Denver income maintenance experiments (SIME/DIME) as it becomes available.

This paper reviews several applications of this model, the implications that the results may have for public policy, and some further areas of research that may shed additional light on the income accounting/reporting issue.

ACCOUNTING SYSTEMS

Any accounting system for a transfer program contains an accounting period, a reporting period, a payment period, and a set of rules for determining income for eligibility purposes, which can rely on prospective or retrospective income reporting. The manner in which these elements are specified determines whether or not a reasonable balance can be struck between three crucial policy issues: (1) horizontal equity--the treatment of program recipients with like or differing needs; (2) reasonable control over costs and caseloads; and (3) responsiveness to the changing needs of the target population.

One of the most important elements involved in an accounting system is the length of the accounting period over which income is measured for determining benefit entitlements. This could be, for example, the recent past month, the upcoming quarter, or a moving average of two or more past months. This accounting period may be lengthened considerably by the use of a carryover of past income that exceeded entitlement, which would be used to offset current and future entitlements. The components of such a carryover system are: (1) the length of time over which past income in excess of current entitlement is "remembered"; (2) whether past excesses are applied against entitlements on a "last-in, first-out" basis (LIFO) or on a "first-in, first-out" (FIFO) basis; and (3) the amount of income in excess of entitlement which is remembered.

Since income may be reported on a more frequent basis (monthly, for example) than is actually used for accounting period purposes (benefits may

only be adjusted quarterly, for example), the reporting period may be shorter or equal to the accounting period. Once accounting and reporting periods are established, however, the system should be further designed to provide the recipient with relatively frequent payments (monthly, as is the case in most transfer programs) in order to alleviate the financial burdens on low-income families. Finally, the system must either use forecasts of future income (prospective reporting) or records of past income (retrospective reporting). With a prospective system of reporting, current payments represent only estimates of entitlements based on income forecasts. Thus, over- or under-payment may occur, depending on the accuracy of the forecast. Retrospective systems do not produce over- or under-payments, since current payments are based on past income which can be reported with certainty.

APPLICATIONS OF THE APS MODEL

Simulations reviewed in this paper all make use of various combinations of the accounting system elements discussed above. For example, the early FAP simulations utilized several alternative versions of prospective and retrospective systems. For simulations of AFDC and the Food Stamp program, we exerted considerable effort to first, produce the accounting system described by the authorizing legislation; secondly, to reproduce the observed accounting system as it currently operates in the real world; and thirdly, to examine several hypothetical accounting systems holding constant the benefit levels and other eligibility criteria to measure the differential effects on the above two programs. For the AFDC program, this requires the use of current-month income to determine "current need", which then remains

in effect until the recipient voluntarily reports a change in circumstances or until the redetermination period runs out. In either case, the "current need" benefit initially in effect becomes a "prospective needs" benefit by virtue of the accountable system; hence, the use of a quasi-prospective accountable system for AFDC. The Food Stamp simulations employ principles of a pure prospective accounting system as Food Stamp rules require a forecast of income for an upcoming month to establish a current month's entitlement.

FAP and AFDC

Alternative accounting systems using FAP's income guarantee and tax-rate parameters were simulated by Jodie Allen in 1973. The data used was from SIME/DIME, as well as the New Jersey, Pennsylvania, and the North Carolina and Iowa income maintenance experiments. There were 5,200 observations, subsequently weighted to reflect urban national totals for 1974. She simulated 19 different accounting systems, spanning the range of all possible alternatives under serious consideration by the administration at that time. Included among them were eight variations of the quarterly prospective system, from the pure quarterly prospective system to the quarterly prospective plan with intra-quarterly reporting, likely administrative lags, LIFO carryover, and automatic recapture of overpayment.

Analysis of the results of these simulations generally revealed that it would be quite expensive to achieve high responsiveness in the prospective systems. For example, costs and caseloads for the pure quarterly prospective system increased by 16% and 41%, respectively, above the annual

entitlement system--which was used as a benchmark to determine equity and the least-cost alternative. Furthermore, when it was realistically assumed that the system would rely on voluntary income reporting between redeterminations, costs and caseloads over annual entitlement increased even further. The best overall balance between costs, caseloads, and responsiveness was obtained by monthly retrospective accounting with a 12-month carryover provision. Cost and caseload differentials were only .5% and 1.6% above the annual entitlement benchmark system.

The first attempt to simulate the cost and caseload impact of existing practices on state-run AFDC programs was undertaken by Allen in 1975. She sought to explore the potential gains which might be realized if states introduced more rigorous income reporting and accounting procedures into the AFDC program. The data base used in these simulations was the same sample employed in the FAP simulations--5,200 low-income families with children involved in the income maintenance experiments. It was not possible to weight the data by state.

The AFDC simulations required several minor modifications in the logic and operations of the original model's quarterly prospective income reporting and accounting system in order to more accurately simulate real world conditions.

Thus, initial eligibility and subsequent redetermination had to be made on the basis of "current" income received in the month instead of forecasts of future income. Second, adjustments were necessary to account for a discrepancy between the legal requirements and actual administrative practice. For example, the median redetermination period for AFDC recipients is roughly nine months, while legislation requires that

it occur at least every six months. Consequently, four alternative simulations with redetermination periods of three, six, nine and twelve months were run in order to bracket the range of variation in current state practices. Finally, since there exists no study which explains recipient behavior with respect to their willingness to report changes in income amount or source, or family circumstances, three sets of simulations were employed to show the effects of different assumptions about income reporting.

Results of these simulations revealed several interesting findings. First, the monthly retrospective plan with a six- or twelve-month carryover option was shown to be most compatible with the goals of equity and to cost the least, in comparison with annual entitlement. Secondly, varying the voluntary reporting assumptions and the length of the redetermination period had little impact on annual caseload, but a substantial impact on benefit costs. Finally, it was shown that the most realistic simulation of the current AFDC program was attained by the prospective plan which utilized a six to twelve-month redetermination period and the high cost forecast probability assumptions. While the model was not intended to predict annual cost and caseload amounts, simulated values were reasonably in line with published government statistics.

Food Stamp Program

Analysis of the Food Stamp Program was conducted with unweighted DIME data. A total of 2,390 low-income families comprised this sample for 1970-1971, which included Chicanos, whites and blacks. The simulations take

account not only of changes in family eligibility induced by alternative accounting/reporting procedures, but likely changes in program participation as well.

Implicit in the design of this program is a three-month prospective reporting and accounting period as described in the Food Stamp Certification Handbook. Quoting from this handbook, we find that "income and expense figures used to determine adjusted monthly income are those anticipated during the certification period" and that "the normal certification period for . . . households shall be three months." However, a recent survey of Food Stamp recipients provides strong justification for a more lax assumption about the length of the certification period. In this survey it is revealed that the median redetermination period is five months for non-public assistance families and nine months for public assistance families. Accordingly, we simulated the base plan with both lax and stringent periods for recertification.

To determine the potential impacts of income accounting/reporting procedures on costs and caseloads, several alternative plans were simulated and compared to the base plans. Although the current Food Stamp program with its prospective accounting system is responsive to the needs of eligible families in that few such families are underpaid, results indicate that this system is both costly and inefficient. Further, several alternative income accounting and reporting systems performed as well as, or much better in some areas than the current program in terms of costs, caseloads and responsiveness.

Foremost among these alternative accounting systems was the monthly retrospective plan (without carryover), which provides benefits in the current month based on income in the previous month. Overall costs and caseloads were significantly below those of the current legislated Food Stamp Program by this system, and the average level of overpayments was substantially lower than in the current program. Specifically, for the non-public assistance and public assistance families, respectively, program costs were reduced by 6.9% and 7.1%, and annual program caseloads (persons ever receiving benefits during a year) were reduced by 1% and 2% under the monthly retrospective plan, as compared with current program operations on the assumption of stringent current program administration. If a more lax current program operation is assumed, and the survey of recipients suggests this is realistic, potential cost savings rise to 13.5% and 8.3%, and caseload reductions increase one percentage point for each group.

The analysis also indicates that potential savings due to more rigorous income accounting and reporting procedures are sensitive to the assumed level of participation among program eligibles--the higher the level of participation, the greater the potential savings. This effect is produced by the fact that, given observed patterns of Food Stamp Program participation, families with potentially more income to report are more likely to participate when the aggregate level of program participation is higher. Thus, even greater savings in costs and caseloads could be expected when participation levels are higher than 1971, the year simulated.

Better Jobs and Income Program (BJIP)

BJIP simulations used the approximately 5,200 control and experimental treatment family observations from DIME for the years 1970-1974 representing ten thrice-yearly interviews. Since the model generally requires two years of data on families present for each month of a two-year interval, the sample was subsequently restricted to approximately 2,000 observations for analysis purposes with the simulation years being 1973-1974. Weights for each family stratified by income/family status/family size category were then derived by dividing each raw DIME experimental cell into the corresponding weighted Survey of Income and Education (SIE) matrix. Gross family income amounts measured on the SIE were deflated to 1974 values while retaining the original 1975 demographic distributions.

BJIP would base eligibility on a monthly retrospective accounting principle and rely upon monthly reporting so that benefits may be adjusted promptly. It also adopts a carryover principle which reduces the current entitlement of recipients (based on last month's income) by the "excess available income" from each of the five preceding months.

The use of a carryover principle serves two purposes. First, it adjusts the benefits recipients receive so that like families receive roughly similar benefits over the course of a year. This produces more horizontal equity among similar families participating for the same length of time, but with varying fluctuations in income over that time period. Secondly, it reduces the overall cost of the transfer program in a manner that is not totally inconsistent with other program goals, such as responsiveness to changing needs of recipients or administrative requirements of simplicity of operation.

In addition to simulating the proposed BJIP with a five-month carryover provision, five other monthly retrospective plans with variable lengths of the carryover period were analyzed. To test responsiveness, the pure monthly entitlement plan which provides benefits based on income in the current month was also simulated. Results show that the monthly retrospective accounting plan with no carryover is nearly 10% more expensive than the BJIP's carryover provision. Other systems which utilize carryover of shorter lengths are also more expensive than the base plan, but only marginally so in several instances. For example, the four- and five-month carryover options increase costs, but by less than 1%, while the percentage differences in caseloads are slightly larger. Caseload differences between the base plan and monthly retrospective without carryover are a striking 46%. The wide disparity between cost and caseload percentage differences reflects the attainment of eligibility by many more higher income families, i.e., those families where annual income exceeds the poverty level.

The relative responsiveness of each accounting system was also simulated. Responsiveness in this, as well as all of the preceding results presented, is defined to be the rapidity with which each alternative system pays benefits with the occurrence of monthly fluctuations in family size or income level. The measure of responsiveness which best captures this concept is obtained by measuring the difference between the benefit paid in the current month based on income received in the current month and that provided by each alternative plan simulated. As expected, monthly retrospective accounting without carryover provided the lowest level of under-

payment relative to entitlement. Average underpayments for the base plan were considerably higher. However, overpayments for all retrospective systems were much smaller and less divergent among each other. Systems with a carryover option of any length reduce the average overpayment level by virtue of their "smoothing" of income over several months.

Policy Implications

The policy implications of the above findings are unequivocal. First, the use of an accounting system which employs prospective principles such as those now in effect for the AFDC and Food Stamp programs unnecessarily boosts the aggregate level of program costs, compared to alternative accounting systems. In fact, virtually every alternative plan simulated for all of the above programs was less costly than the prospective accountable system, notwithstanding the other policy considerations of equity and responsiveness.

It may be argued by some that the increased level of costs associated with the prospective accountable system buys a higher level of responsiveness in the sense that fewer families are underpaid relative to their entitlement. Yet, the evidence provided by these simulations strongly supports the observation that an equal level of responsiveness is provided by retrospective accounting of income and family circumstances in addition to the attainment of a lower level of benefit overpayment relative to entitlement. It is also noted that the annual caseload level under retrospective accounting was equal to or less than that produced by the prospective systems.

Secondly, considerations based on horizontal equity require the introduction of a carryover option into the basic accounting system. Two purposes are served as a result: first, it adjusts benefits that like recipients receive in the current month for differences in their income streams. This produces more horizontal equity from an annual perspective among similar families participating for the same length of time, but with varying fluctuations in income over that time period. At the same time, it reduces the overall cost of the transfer program in a manner that is not totally inconsistent with other program goals, such as responsiveness to changing needs of recipients or the administrative requirements of simplicity of operation.

It should be fairly apparent that the three issues of cost and caseload control, equity, and responsiveness are inextricably intertwined. Given an overriding goal of maintaining or decreasing the overall program costs of administering a welfare system, it is clear that some sacrificing of the highly desirable goal of responsiveness must be offered in exchange. This may be accomplished in two stages by first, adopting a pure monthly retrospective accounting system which significantly reduces program costs and marginally reduces annual caseload without sacrificing responsiveness considerations. Secondly, further cost savings and caseload reduction may be obtained by appending a carryover option of appreciable length onto the basic system. The only social cost associated with this second stage relates to the decrease in system responsiveness as a result of larger underpayments. However, this important drawback is sharply mitigated by two positive

outcomes of this operation--overpayments are smaller and more equity is obtained. Hence, a reasonable balance is struck between the three conflicting goals of cost and caseload control, equity, and responsiveness.

FUTURE RESEARCH

In the future we will, on one hand, attempt to estimate more precisely the impacts of alternative accounting systems through the use of larger and more representative data bases. In particular, analysis of the BJIP accounting system is currently in progress with the purpose of expanding the weighted sample of observations from DIME to include SIME. Also, newer accounting systems which involve, for example, pure annual entitlement (equity and least-cost comparison), moving average, and monthly prospective, are being designed for incorporation into the model.

On a more fundamental level, research efforts will focus on the empirical support for some of the underlying behavioral assumptions contained in the APS model. For example, there is some debate about the ability of recipients to forecast income and to voluntarily report known changes in income or family circumstances under a prospective system. Also open to question is the assumption that recipients respond to changes in benefits, rather than changes in income, when it becomes necessary to produce an estimate of the probability of current or future income (or a change in income) will be reported. This assumption suggests that recipients implicitly know and calculate effective benefit reduction rates. One could assume that they look only at income levels. Sensitivity analysis may also be required, therefore, to test the impact of either assumption.

Lastly, research on the effect of using 24 months of data to simulate the impact of a nine-month prospective system is needed. This is important because the full long-run effect of the nine-month redetermination period is not captured by current model operations.

DISCUSSION

BERNIE STUMBRAS: When you say that you're going to estimate three months in advance, are you going to recoup any overpayment?

SPRINGS: For the AFDC, the benefit is determined based on a current month at the start-up, and this remains in effect for three or six months, or twelve months, and only changes when we allow in the model for a family to voluntarily report a change in income or circumstances. If the family does not report any changes in their circumstances that differ from what they were when they initially came on the system, then the benefit remains the same for the entire year.

STUMBRAS: There would be no incentive to report unless your earnings or other income went down.

SPRINGS: That's true. We model the probability of a family reporting a change in income or circumstances. We assign a probability to different income changes, relative to total family income, and if the change in income is large enough, we will assign a higher probability; if it's lower, we'll assign a much smaller probability. Then we compare this number to a randomly generated number, and if in fact the probability is more than the random number, we will allow that change to occur and then recompute the benefit. However, these probabilities have not been empirically substantiated.

STUMBRAS: Again, one of the things that we (in Wisconsin) are now changing is the recoupment policy for overpayments. Before, we could find that a family got \$275, on the average, more per month than they should have. All we could do is ask them to voluntarily return that \$275, and the client would smile and say no. From that point, that person would be asked in every month, with their pay stubs.

SPRINGS: That's one of the main problems generated by the prospective accounting system. You're going to generate much larger levels of overpayments simply because families fail to report, or even if they do report, the agency doesn't act on it for a couple of months.

STUMBRAS: But if you model a prospective three-month accounting period and you also recoup all overpayments, you shouldn't treat the clientele any more generously than retrospective accounting.

SPRINGS: We have simulated recoupment options with the prospective systems, and you are right in that it will reduce overpayments. The problem is administrative. In practice in the real world I'd expect that recoupment is much more difficult, and can create unnecessary burdens on families who have little or no income.

STUMBRAS: It's operating in a lot of places.

KAREN DEASY: Do you recoup for both client and agency error?

STUMBRAS: Right now we don't recoup at all except voluntarily, but a new system will go into effect July 1, when we'll recoup only from earned income, for either client or agency error, and transfer income only for willful neglect on the part of the client.

GARY CHRISTOPHERSON: What about changes in family size?

STUMBRAS: You still have to develop some evidence of willful neglect.

CHRISTOPHERSON: You mentioned a lack of responsiveness with a six-month accounting period, or a five-month carryover. What percentage of the families received a payment that was less responsive, or that was less responsive by 10%, or some substantial amount?

SPRINGS: For the most part the lowest income families are affected the most by the differences in responsiveness between pure monthly retrospective accounting and the six-month accounting system.

STUMBRAS: I guess I need a definition of responsiveness.

SPRINGS: To measure it, we determine what the family is entitled to under a hypothetical base plan, and consider this the most responsive system that we could have.

STUMBRAS: So there's greater fluctuation of income among lower-income level families?

SPRINGS: Right, and because of that they are not always given a benefit in the current month.

STUMBRAS: Right, while in fact the six-month retroactive accounting was adapted in PBJI to affect the high income families.

SPRINGS: Right, it does affect the high income families too.

STUMBRAS: But nobody talked about the side effects. You can have a responsive system and reduce inequity between similar families with a carryover option and a recoupment of anything that is overpaid. For example, if a family earns more than \$20,000 in a year, but in the first three months of the year had almost no income, you will pay benefits for those months. Then you can either carry that forward and recoup the benefits when income is earned, or you can wait until the end of the year at income tax time.

SPRINGS: And merge the two. We at Mathematica will be starting some work in that area, particularly with the Food Stamp Program. Some people are interested in seeing how recoupment of food stamp bonuses merge with the tax system.

THE COLORADO MONTHLY REPORTING PROJECT

by

Alan M. Hershey
Associate Director, MPR

The Colorado Monthly Reporting Project is a project that takes the administrative techniques the income maintenance experiments used for income reporting and applies them to the Aid to Families With Dependent Children (AFDC) program. In all of the experiments, those receiving payment were required to complete and submit monthly reports on their income and family circumstances. This was used as a basis for monthly computation and issuance of their benefit payment. This reporting approach was viewed by many observers of the experiments as an administrative tool, which if applied in the AFDC program, could allow more responsive and accurate adjustment of grant levels as recipients' circumstances changed.

Using the monthly data submitted by the people in the experiments, simulations were done to test this hypothesis. Estimates were made of the likely effect on aggregate benefits and their distribution if monthly reporting and a variety of retrospective accounting techniques were substituted for the usual biannual or quarterly reporting and projection of circumstances used by most welfare agencies.¹ These simulations indicated that AFDC benefits would be more efficiently distributed if monthly reporting and retrospective accounting were used. They also predicted that the

total amount of benefit payments would be reduced with monthly retrospective reporting. Monthly reporting in combination with a one-month retrospective accounting period would result in a decrease in aggregate benefits of between 5 and 19% in Colorado. Use of longer accounting periods would result in even greater reduction.

Because the simulations indicated apparent potential for improvement in AFDC administration, the Department of Health, Education and Welfare (DHEW) and the Colorado Department of Social Services decided to test effects of monthly reporting and retrospective accounting in the field through the Colorado Monthly Reporting Project. The project had two major purposes. The first aim was to develop and implement an administratively feasible system for monthly reporting using one-month retrospective accounting, and to measure the system's impact on benefit payments; second, the effect of this system on administrative functions and costs was to be evaluated.

There were two distinct parts of the project. In Denver County, the monthly reporting system was implemented in a controlled experimental setting. A random sample of approximately 10% was selected from the AFDC caseload and placed on the monthly reporting system beginning in March 1976. A comparable sample of control cases, who continued to receive benefits under conventional reporting procedures, was selected for observation throughout the experiment.²

In Boulder County, the monthly reporting system was implemented for the entire AFDC caseload in November 1976, and changes in administrative

functions and costs were assessed. (This administrative pretest was conducted in a second county because of the difficulty of measuring the administrative impact in Denver, where the agency's conventional procedures were still in use for most of the caseload.)

DESIGN AND OPERATIONS

Three major features distinguish the system designed for the Colorado monthly reporting project from conventional administrative methods for AFDC: mandatory monthly reporting, retrospective accounting, and a high level of automation.

Each month, recipients are mailed a comprehensive monthly status report (MSR) form that collects data on most factors of AFDC eligibility. This form is designed to be easily understood and completed, and to elicit unambiguous responses. AFDC recipients must complete and submit this form to remain eligible and receive their next check, whether or not they think their circumstances have changed. In contrast, under conventional procedures in most areas, AFDC recipients are required to report formally on their income and family circumstances at infrequent intervals (six months in the regular AFDC program, and three months in the AFDC-Unemployed Father program). If changes in circumstances occur between these scheduled reports, recipients are expected to contact the agency and report the change.

In the Colorado system benefits are computed each month using retrospective accounting. Each MSR filed by the recipient lists income received, family composition and other circumstances in the monthly reporting period just ending at the time the recipient completes the report. That month's

data are used as a basis for determining continuing eligibility and the grant to be issued in the following month. Under conventional AFDC accounting techniques, grant amounts and eligibility are based on prospective estimates of need. Benefits issued in June, for instance, are based on projections of income and family composition made at some earlier time.

Finally, the Colorado monthly reporting system is distinguished by the high level of automation used in the collection and processing of information and administrative functions. The system automatically pre-prints and mails the report forms on scheduled dates; monitors the receipt of completed reports; and issues notices on several filing deadlines that warn of discontinuance if the MSR is not filed. MSR's are entered directly to the computer system without any review or examination for eligibility by staff. Extensive computer edits detect missing answers, internal inconsistencies, erroneous entries, and major changes in circumstances from the previous month. The results of this editing process are reported to eligibility technicians, who focus their attention on those reports that require correction or clarification. Based on this MSR data, the system computes the authorized amount of assistance by calculating earnings and then taking into account changes in family composition, school attendance status, required withholding adjustments, and special pregnancy allowances. If the processing of an MSR results in reduction or termination of benefits, the system immediately produces notice of such action to the recipient. The system also issues a variety of reports to eligibility staff, supervisors, and agency managers. Although many welfare agencies use computer technology to maintain case files and reduce clerical activities, the Colorado system represents a considerable

advance in the level of automation.

The schedule used in the Colorado system is designed both to provide rapid issuance of benefits in response to recipient reports, and to structure agency workload in a manageable fashion. Recipients can be issued assistance checks as soon as sixteen days after the end of the reporting period covered by the MSR. Three payment dates are provided. Those recipients who file rapidly receive early payment, and those who file later in the month receive payment on the second or third issuance date, at the latest one month after the end of the reporting period. Two overlapping cycles of reporting and payment are used to even out the workload of the eligibility staff. Half of the recipients report on circumstances by calendar month. The other half reports on circumstances in periods extending from mid-month to mid-month.

RESULTS

Analysis of the effects of the monthly reporting system has focused on patterns of benefit payments and on administrative processes. In both cases, research has been designed to answer important concerns from the standpoint of both program administrators and AFDC recipients. Administrators need to know the effect on total benefit costs and whether payments can be more accurate and responsive to changes without serious increase in administrative costs. From the perspective of recipients, it is important to look at the burden of monthly filing compared to its possible advantages. It is also important to know how retrospective determination of grants will work in meeting current needs of the recipients and how the variable payment

dates will affect budgeting.

Comparison of payments to experimental and control cases in the Denver experiment has shown that the monthly reporting system reduces aggregate benefit payments. There is also a marked increase in the benefit adjustments made in response to changes in circumstances. Over the first eight months of the Denver experiment, benefit payments were approximately 8% lower for experimental cases than for control cases. This reduction is due primarily to a reduction in the average length of stay on assistance. The monthly reporting system more rapidly identifies changes in circumstances that cause ineligibility and discontinues such cases more promptly than the conventional system. Analysis of benefit payments has recently been extended through the first complete year of the Denver experiment, and continues to show similar results. Further research will look at a full two years of data.

The monthly reporting system also appears to implement adjustments in benefits more promptly following changes in circumstances than does the conventional system. In the Denver experiment, it was found that in any given month, approximately 20% of ongoing experimental cases reported changes that resulted in grant size changes, whereas such changes occurred for only about 8% of control cases. This sharp rise in grant changes holds true for both increases and decreases.

These improvements, the reduction in total benefits and the improved

responsiveness to changes, are obtained for a relatively small increase in administrative cost. This is found in the Boulder pretest where the entire AFDC caseload was on the experimental reporting system. The primary cost increase is for additional data entry staff and the more extensive computer processing required. However, since the system automates a number of previously manual tasks, there seems to be no increase in need for eligibility staff. Increased costs for forms and postage are offset by savings in clerical and bookkeeping time, where the system reduces workload significantly. The net effect of these changes is an increase in total administrative cost of 0 to 8%. It appears that the increased administrative cost of a monthly reporting system yields a relatively high return in reduced assistance costs.

These changes in the administrative costs need to be interpreted with some care. It has been measured only in one relatively small county, and was measured against its normal set of administrative functions and associated costs. However, since Colorado's administrative costs for AFDC are below the national average, it seems fair to assume that many other areas may be able to change to a monthly reporting system with even less impact on administrative cost. Measures of administrative cost impact in this study do not take into account the potential advantages of automated reporting of other assistance programs along with AFDC. In the Boulder County experiment, for instance, AFDC eligibility technicians process Food Stamp actions

separately for participating cases. A system that automated both might result in an overall administrative saving in the two programs.

With the monthly reporting system there has also been a sharp decline in the need for corrective measures such as recovery actions, retroactive payments, and check cancellations. Since the beginning of the Denver experiment, the frequency of overpayments has been 74% lower in the monthly reporting group than in the control sample. Since most recovery actions for overpayments do not result in actual collection, the reduction of overpayments themselves is also an important source of savings. Recipients are also, therefore, less likely to be faced with legal or other action forcing repayment.

The incidence of retroactive payments made because of underpayments, and cancellations of erroneously issued checks, has also been considerably reduced by the system. In Denver County, retroactive payments have been issued to only 1.6% of all cases, as compared to 4.1% of the control group, a reduction of 60%. Similar results have been obtained in Boulder. Check cancellations have fallen by 78% in Denver and 80% in Boulder with the use of monthly reporting.

It is clear that a monthly reporting system can have important administrative advantages. For the AFDC recipient, monthly reporting becomes an added, but manageable, responsibility with potential advantages. Most file MSR's promptly and adequately. Between 85 and 90% file them by the first filing deadline. Although about 50% of all MSR's contain some kind of edit problem or major change, corrections and clarifications are made quickly and

about 85% of all recipients receive their grants on the first payment date.

However, there is serious concern felt by both administrators and recipient advocate groups that retrospective accounting will have adverse effects on recipients when a sudden drop in earnings or other income may leave recipients with an insufficient grant until a month after the income decline. Clearly, retrospective grant computation can pose this danger, but safeguards can be designed. Sudden large decreases in income do not happen often.³ Given the rarity of such occurrences, it appears that inadequate grants due to retrospective accounting should be dealt with by emergency assistance, rather than by abandoning retrospective accounting and the advantages it offers in improved accuracy and suitability for automated processing.

That, however, does not answer all the questions about the impact on recipients. The information gained thus far is indirect. A survey is now being done of both active and discontinued recipients, of attitudes about the burden of monthly filing, the clarity and reliability of the monthly reporting mechanism as compared to conventional reporting procedures, the problems posed by variations in payment date from month to month and the effects of retrospective accounting. In addition, more work is being done to determine how often severe problems occur due to inadequate income as a result of retrospective reporting.

FUTURE RESEARCH

Although analysis of the monthly reporting system's impact on total benefits and administration is still being refined and extended, results

to date suggest that the techniques developed in the Colorado project can make a substantial improvement in the administration of public assistance programs. Further work needs to be conducted in two directions. First, it is important to determine whether the Colorado results can be replicated in other areas, particularly in the larger cities with what are seen as the most severe administrative problems. Recent proposals by DHEW for a series of experiments and administrative demonstrations using AFDC monthly reporting will address this need. Second, projects should be done that integrate a full range of assistance programs. Such a program is being designed now for testing in Colorado that will integrate reporting, eligibility determination, benefit issuance and data base management functions for financial, medical, and Food Stamp assistance programs. Evaluation of this system's operations and impact, combined with the Colorado recipient survey and analysis of further demonstration projects, should provide a sound basis for administrative change at the state level and for federal welfare reform.

NOTES

1. Jodie Allen, "Simulation of the Impact of Income Reporting and Accounting Procedures on AFDC Costs and Caseloads", (Urban Institute, Washington, D.C., February 1973).
2. Both samples were replenished each month with randomly selected approved applicants to prevent sample depletion due to case discontinuances.
3. Of the approximately 9% of grants increased in any given month due to changes in family circumstances, 6% are due to income changes, and only about 1/3 or 2% of those are adjustments greater than \$50. Less than 1% of the active caseload in a given month has changes in income that should result in a grant increase of more than \$100.

DISCUSSION

(Robert Williams presented the paper for Alan M. Hershey at the conference.)

ROBERT SPIEGELMAN: It seems to me one by-product of this might be the use of a body of data for research and evaluation. Are these data useful for that, more useful than they would be under a normal system?

WILLIAMS: Most definitely. I think there is increased potential for doing research and recipient-characteristics analyses because of the data base that has been generated. One thing that is coming out of the project is information on patterns of recipient assistance. Most of the data that have been recorded nationally on length-of-stay on assistance, for example, is based on a cross-section estimate. It does not really give you an expected length-of-stay for the average recipient; whereas in Colorado we are looking at it on a longitudinal data basis. Our data show the expected length-of-stay, at about 18 months for AFDC regular cases and six of seven months for AFDC-U cases.

SALLY ANDERSON: I assume that HEW excused you from such things as quality control and fair-hearing procedures during the course of your experiment?

WILLIAMS: No.

ANDERSON: In our income reporting system in Kansas, one of the biggest jobs we had was to educate the clients. During the time that we were educating the clients we had a long series of fair hearings. I think as the clients have become used to the reporting system, we've had fewer, but we continue to have fair hearings over income reporting forms.

Income reporting forms have to be very sensitive. For instance, it doesn't do you any good to have a form that only shows the amount of what you would call "gross income"--that is, your social security and your federal income tax--because for quality control you must show any changes from hourly wages, going from part time to full time, etc. In the beginning, the administrative procedures in our urban agencies was overwhelming and resulted in changes in the total agency operation. Income reporting takes a tremendous amount of time on the administrative level and many more things go on in an agency. I sound negative about it; I am not. I wouldn't get rid of it for anything; but I would want you to understand, it's not as clean and easy as it sounds.

Income reporting for Food Stamps needs to be considered carefully. It is hard enough with migrants, for example, under the present system to get reporting on a migrant situation. With the new rules regarding income reporting and the different ways migrant income is treated, there will be problems in both the regular and emergency issuance of stamps.

Also, it's not hard for me to figure out why you have fewer people applying for AFDC in your experimental program. It has nothing to do with income reporting. It has to do with retrospective accounting. If I've earned \$700 this month, I know there is no sense to go and report next month.

WILLIAMS: How about the following month?

ANDERSON: Lots of things can happen in a month's time. I could have found a job, I could have moved in with somebody I don't think that you can pinpoint that it's due to monthly reporting. I don't think that any agency, unless they plan to use tremendous amounts of state money, can go to initial retrospective accounting; because you have somebody standing in front of you with nothing to eat, no roof over their head, and no money. Now, it's useless for me to average six months back and say, "You should have saved \$700", when he hasn't got \$700. I fully approve of retrospective accounting once a case is open. But there is no way, unless you are willing to start all sorts of expensive emergency assistance programs such as the welfare reform acts are proposing, because for the most part the people that come to you have nothing.

MASAKO DOLAN: What about the ten-day notice requirement?

WILLIAMS: There are two components to a notice requirement under AFDC, one is adequacy and the other is timeliness. Under existing HEW regulations timely notice is not required if you are reducing or terminating assistance based on information provided by the recipient in written form with a signed statement that they understand that this information is going to be used to affect their grant, and that it may reduce or terminate the grant. Every time a recipient signs an MSR, they sign this statement. Therefore, for reductions and terminations based on information provided in the MSR, we do not provide timely notice. We do provide timely notice for all other types of adverse actions. For example, if people are being discontinued for failure to file they get ten days notice.

I must say that aside from the questions of legality, I don't think we have run into any real questions in terms of fairness to recipients. The timely notice requirements were built into the existing system, which is basically prospective. If we had to build those into a retrospective system, it would make it extend the processing cycle and therefore be less responsive to the recipients.

PARTICIPANTS' RESPONSES TO AND PERCEPTIONS OF
THE OPERATIONAL AND ADMINISTRATIVE ASPECTS OF
THE INCOME MAINTENANCE EXPERIMENTS

by

John Hall
Consultant, MPR

and

Billy J. Tidwell
Senior Sociologist, MPR

The results of the income maintenance experiments can provide guides not only for designing a more efficient use of welfare support levels and tax rates but also a more humane and better system for administration of the welfare system. The latter perspective calls attention to the administrative aspects of the experiments and the ways in which the participant families perceived and responded to them. Policy interest in this area stems from the need to control the administrative costs. And in a related need to devise an equitable system which would minimize, insofar as possible, the potential psychological burdens that might be imposed on participants.

GENERAL FRAMEWORK

First, we are concerned with participant compliance with income reporting requirements, including: (1) the extent to which the parti-

cipants filed their income reports on time, and (2) the extent to which these reports could be processed immediately for payment by program staff. In this connection, a key administrative feature of the experiments has been the use of monthly income reporting and retrospective accounting--that is, computing program payment amounts on the basis of data submitted the preceeding month. In order to operate efficiently and responsively, therefore, the timely receipt of reports has been essential. By the same token, reports that contain errors, inconsistencies, or omissions create processing delays by requiring the expenditure of staff time to effect the necessary adjustments. Either circumstance--late filing--may be costly to the participant as well as the program itself. Examining the participants' performance in respect to the income reporting requirements of the experiments provides a basis on which to anticipate the level of compliance that might characterize a national income maintenance program.

Second, we are interested in contacts between the participants and program staff, as this is probably intimately related to compliance behavior. Many contacts may be generated by the need for program staff to communicate personally with families in order to clarify, correct, or complete information contained in their income reports. However, other contacts might be initiated by the participants themselves--either in relation to income reporting procedures or a variety of other matters that may or may not be program-related. In any case, the design for a national income maintenance program will have to provide for a certain

amount of personal worker-participant interaction. Again, observations from the experiments should help planners to estimate the types and volume of participant contacts that are likely, and to better estimate staffing needs.

Third, the participants' perceptions of a social program may influence greatly their behavior toward it. Feedback from participants, therefore, is a potentially valuable source of insight for purposes of program planning and management. Indeed, feedback data on program operations and administration might illuminate dysfunctional or undesirable aspects that other, more objective evaluative data may fail to disclose. These observations are particularly applicable to new or innovative programs. Fortunately, certain feedback data are available that allow us to examine the administration and operation of the income maintenance experiments from the viewpoint of the participants themselves. Although it is pursued separately, this component of the discussion will give added meaning to the more focused analyses of participant compliance and contacts.

METHOD

Our examination of the above topics relies on data collected in the Seattle and Denver income maintenance experiments (SIME/DIME), and the Gary experiment. With regard to participant compliance and contacts,¹ the SIME/DIME data were drawn from the administrative files of individual families. Only experimental families who were not on Aid

to Families with Dependent Children (AFDC) most of the time and usually received more than the minimum transfer payment are included in the study. The sample thus consists of 447 families in Seattle (for 1972) and 385 in Denver (for 1973). The different time frames permit us to examine families who have been enrolled in the experiments for roughly an equivalent period of time.

The data base for Gary differs, in that it is comprised of aggregate data extracted from monthly administrative reports and covers all but the first 12 months of the experiment. Thus, the Gary analysis does not focus on the compliance behavior and contacts of individual families, but the sample as a whole. This difference obviously makes comparisons with the Seattle/Denver results problematic and should be kept in mind.

Concerning participant perceptions,² the analysis is based on certain feedback data elicited from the Gary participants six months after they had been disenrolled from the experiment. The study sample consists of 637 respondents and is limited to experimental households. The absence of similar feedback data from the Seattle/Denver experiments precludes any kind of cross-experiment comparisons in the area of participant perceptions.

In terms of analytic procedures, the investigation of participant compliance and contacts involves the use of descriptive and multiple

regression techniques. The participant perceptions data are presented in the form of percentages and percentage differences.

PARTICIPANT COMPLIANCE

Timeliness of Filing

Data indicate that the overwhelming majority of the income maintenance participants filed the required Income Report Forms (IRF's) on time at all sites. Of the 8,880 monthly reports filed by the sample families in SIME/DIME in the months for which data are available, 93% of them were filed on time. The analysis revealed no important differences by ethnicity or number of heads of families, indicating that the high rate of on-time filing was universal. Denver had slightly more on-time filing among one-parent families and white families, while in Seattle there was very little difference. Among the possible determinants that were examined, the amount of payment was the only variable that accounted for any part of the small variation in on-time filing. Each \$100 increment in benefits increased the Seattle IRF's filed on time by just over 3%. At average payment levels of \$400, almost all IRF's were filed on time in both sites.

A comparable average of 93% of the experimental families in Gary filed their reports on time (more than 94% during the period just before disenrollment). There is no doubt that the basic schedule used in the Gary Experiment, providing participants with five days to a

week to complete and submit the forms, was manageable for most families. It might be expected that the timeliness of reporting would improve over time, and the Gary data do show a modest decline in the percentage of late reports. However, the differences across months are slight and not significant, except for the months just before and during the disenrollment period. Generally, then, we find virtual constancy in filing performance over time. On the other hand, variability of late filing and failure to file (by month) is considerable. Even excluding the disenrollment period, 1.4 to 6.3% per month file late; 1.2% to 6.4% fail to file each month. (Later, we examine the extent to which these results can be explained by administrative changes.) To assess the effects of size of payment on filing, we did not have individual family data for Gary. Data on mean monthly payment by experimental and AFDC status was available, however, and they show that the larger the mean monthly payment, the better the filing performance of the group!

	<u>Percent Filing On Time</u>	<u>Mean Monthly Payment</u>
Experimentals on AFDC prior to experiment	96.7%	\$291
Non-AFDC Experimentals	89.0	161
Controls on AFDC	84.0	10
Non-AFDC Controls	76.7	10

These differences in filing behavior are moderately large and tend to support the Seattle/Denver finding that the amount of payment is an

important determinant of timely filing. However, the difference in filing performance between AFDC and non-AFDC control families suggests that additional factors may be at work which cannot be captured in a cross-sectional analysis of aggregate data.

Regression analysis of SIME/DIME showed that variation in timeliness of filing was not explainable by administrative changes. Only one administrative change--in the IRF in September, 1978--had any significant impact. It increased late filings among experimental AFDC families during the month the change occurred.

To reiterate the principal result, on-time filing was clearly the norm in all three experimental sites. Thus, there would appear to be little question but that monthly income reporting is, at least in this sense, administratively feasible for a national income maintenance system.³

Processability of Reports

In filing their income reports, participants were required to provide detailed information about the amount and source(s) of their income and their family composition in the preceeding month, along with appropriate documentation (i.e., paystubs for earned income and receipts for expenses). The precise content and format of the reporting form varied between sites and, in Gary, changed over time, but the basic requirements were essentially the same. We will examine, first, the

extent to which these requirements were met in Seattle and Denver-- that is, the extent to which the IRF's were immediately processable.

To be classified as immediately processable, experimental IRF's in Seattle and Denver had to be signed, all YES/NO boxes had to be checked, all required zero entries had to be made, information had to be legible, and types of income received had to be entered in the correct spaces. By these criteria, 47% of the samples IRF's in Seattle and 61% of those in Denver were immediately processable. (As will be seen below, when a more liberal definition was used, the rate of immediately processable IRF's increased greatly.) The difference between the rates in Seattle and Denver cannot be explained adequately by the available data, but can probably be attributed largely to differences in procedures in the two sites.

There were some differences in rates of immediately processable reports for different family types. In both SIME and DIME two-parent families and white families filed immediately processable IRF's much more often than other types of families with otherwise similar characteristics. Since the analysis controlled for differences in income, years of education, number of earners, self-employment status, and number of IRF's filed, the ethnic differences are perplexing. However, quality of education was not controlled for.

Returning to the main finding, it is unlikely that a national program would require IRF's as detailed as those required by the ex-

periments. Accordingly, the Seattle/Denver IRF's were examined not only in terms of the experimental criteria discussed above but also by ignoring those items on the IRF's which, in our judgment, would probably not be used in a national program.⁴ If the remaining data were reported in such a way that the IRF' would be keypunched immediately, the IRF was considered processable. Under these new conditions, the estimated processability rate rose dramatically: in both sites, over 90% of the IRF's were defined as immediately processable.

Since aggregate data were used in the Gary analysis, it is not possible to compare results with the Seattle/Denver study in any but the most general way. Nonetheless, as in the other experiments, a fairly large number of IRF problems were identified. We found, on the average, almost 37 problems per 100 IRF's submitted. However, the proportion of families filing defective IRF's is probably lower, due to the likely occurrence of multiple problems within single IRF's.

The number of problems in the IRF's of non-AFDC families (who are more comparable in family composition to those of the other experiments, including Seattle/Denver) is more than a third higher than for the AFDC families. One might, in part, attribute the superior performance of the AFDC families to their prior experience in dealing with a welfare system or to their greater financial incentives. On the other hand, the bulk of the problems involve erroneous reporting of income

and non-AFDC families were much more likely to have earned income as well as multiple earners. (The bulk of these families had at least two adults.) Each of these factors would probably work to complicate the filing process and this is the more plausible explanation of the difference.

Another perspective on problem rates--and one which is directly relevant to the administrative effort required to process IRF's--is provided by grouping the IRF's according to need for staff to contact filers for clarification or additional information before processing payments. Based on the rules and procedures used in the Gary Experiment, we group IRF's into three problem types:

- 1) Problems not requiring contact: staff can make correct entries from data provided, especially with the help of an historical record of the family's prior reports.
- 2) Problems probably requiring contact for experimental purposes rather than administrative purposes: IRF's provide the data required for payment calculation but deficient procedurally (e.g., missing signature or documentation).
- 3) Problems requiring contact: data provided on the IRF were inadequate to calculate the payment.

The data show that the non-AFDC experimental families in Gary generated significantly more IRF problems for all three types. We also attempted to determine whether certain variables relating to the administration of the Gary Experiment influenced the rate of the IRF reporting problems. Generally, the number of problems increased over time.

The ratio of staff to clients had a small positive effect on the overall number of problems, but this variable also produced a small decrease in specific types of problems experienced by AFDC experimental families. By far the largest systematic impact was caused by the switch to a caseload management system, under which one person was responsible for both processing IRF's and maintaining contact with a specific group of families (rather than different groups of workers). The change reduced the incidence of IRF problems, particularly among AFDC experimental families. Finally, changes in the report form itself increased problem rates during the months they were implemented.

To summarize, the results appear to be somewhat less favorable with respect to the question of processability than they were concerning timeliness of filing. However, the more liberal this definition, the lower the incidence of IRF problems and, thus, the amount of staff time needed to correct them.

PARTICIPANT CONTACTS

As already noted, an important aspect of planning for a national income maintenance program is to anticipate the level of personal staff-participant interaction that may be required. Assuming that income data is self-reporting some amount of staff-initiated contact activity will be necessary in order to clarify, correct, or complete reports. Other contacts will be initiated by the participants themselves--either

in connection with their income reports or other matters. To determine the probable staffing requirements of a national income maintenance program, planners must first estimate the level and kinds of contact activity that may be expected.

The Seattle/Denver Results

In Seattle there were 34 contacts (initiated by either staff or participant) per 100 IRF's filed; in Denver there were 73 such contacts per 100 IRF's. Although we were not able to test the possibility, differences in office procedures and administrative practices are thought to account, in large measure, for the different contact rates. Some effect may also be due to the stronger labor markets in Denver, where changes in income may have been more frequent. (Contacts related to income reporting and payments account for a large proportion of the total.)

Relatively few contacts were disruptive (held up payments processing). Disruptive contacts occurred for 7% of the sampled Seattle IRF's and 14% of the Denver reports. When contacts which dealt with matters specifically related to the experiments are not counted, the data indicate that, annually, 5 to 10% of the reports would have been disruptive. We believe that this range is indicative of the rate of disruptive contacts that would occur in a national welfare program using self-reported income and family status for establishing benefit

levels.

In both Denver and Seattle, family-initiated contacts were more common than contacts initiated by field office staff. Problems related to check-cashing, incorrect payment amounts, and other payments problems were major reasons for contacts in both sites. In Denver a relatively large proportion of contacts involved questions about income and expense.

Some racial differences appear, with fewer contacts for white families than black or Chicano families. In Seattle, the white/black differential is thought to be due largely to the fact that the field office was more conveniently located for the black families and they were more likely to deliver their IRF's in person. In Denver, problems and contacts related to payments were more prevalent among black families than among white families. Family type was a significant determinant in both sites, with two-parent families showing a higher rate of contacts than one-parent families.

As one might expect, contacts to explain the rules of the program, and to resolve income reporting issues declined as families became more experienced with the programs. The decline was more rapid in Denver than in Seattle. Problems with payments declined over time in Seattle, but not in Denver. It could be that the Denver program, having started later, was able to take advantage of Seattle's experience with its payments system and thus had a lower number of "start-up" problems.

The Gary Results

In Gary we found a mean monthly incidence of contacts of all types of 56.1 per 100 families--falling between the estimated incidence of contacts for Seattle and Denver. As in Seattle and Denver, most contacts were client-initiated, although the proportion of contacts initiated by staff rose dramatically during the disenrollment period. The majority of contacts (65.1%) were related directly to IRF processing and payments.

We were somewhat surprised to find that timeliness in filing IRF's and the incidence of problems were not good predictors of the total contact rate, particularly in comparison with various administrative variables. However, we are convinced that this circumstance is attributable, in large part, to certain technical limitations of the study. The relationship between filing difficulties and participant/staff contacts is more complex than we originally believed, and there are many important factors which the available data do not allow us to investigate.

We can say that certain administrative variables were highly influential. For example, increasing the staff size for a given number of participants produced a large increase in contacts, while switching to the caseload management system produced a large decrease. Administrative procedures in the income maintenance experiments were designed largely to facilitate family contacts. It seems unlikely that a larger

and more impersonal national program would involve as much personal interaction between clients and staff as did the experiments. Also, automated IRF processing in a national program would probably result in fewer worker-initiated contacts than was the case in the experiments. Thus, one might expect the rate of participant contacts in a national program to be lower. We would add, however, that the case for minimizing contacts is not clear-cut. Certainly, it does not pay to eliminate those contacts which are productive in terms of reducing payment error. Unfortunately, present data do not allow us to estimate what level of staff-client interaction might be desirable in terms of productivity.

PARTICIPANT PERCEPTIONS

Perceptions of the Rules of Operation

In the most basic sense, the character of any social program is embodied in its rules of operation. Among other things, the rules govern one's eligibility for continued participation in the program, the amount of benefits received, and, occasionally, how one disposes of those benefits. Although we have data only on Gary the basic rules of all the income maintenance experiments were essentially the same. Gary participants were asked for their perceptions of program rules-- in terms of (1) intelligibility, (2) coverage, (3) enforcement, (4) strictness, and (5) continuity.

The findings reveal that only 16% thought the rules were too hard to understand, and a comparable percentage believed they changed too much over time. On the other hand, about 44% thought the rules intruded too far into their lives--and 48% thought they were stricter than need be. Apparently, then, the participants experienced little difficulty in understanding and staying abreast of the rules⁵--while substantial proportions found the nature of the rules--in respect to coverage and/or strictness--objectionable.

Unfortunately, we can at the moment only speculate as to whether the objecting families were actually less willing to comply with the rules. However, it is plausible to hypothesize that they did.

Finally, on the question of enforcement, about 20% of the participants did not believe the rules were equitably enforced. Although this was a distinctly minority opinion, one wonders how it might have developed. Program operators were acutely aware of the need to avoid differential application of rules and penalties--a need made all the more urgent by the experimental nature of the program. Possibly many of the persons who perceived inequitable enforcement simply did not understand how the rules worked, but perhaps some differential treatment in this area did occur. Needless to say, this is an issue to which the operators of any social program must be highly sensitive.

Obviously, it is difficult to evaluate these results in the ab-

sence of some baseline. Thus, the participants who were an AFDC prior to enrolling in the income maintenance project were asked to assess the rules associated with AFDC in terms of three of the five attributes-- i.e., intelligibility, coverage, and equity of enforcement. In each instance, the participants were substantially more critical of the AFDC program than they were of the Gary rules and procedures. On the matter of intelligibility, for example, 48% expressed a negative view toward the AFDC rules, while as we have seen, only 16% judged the rules of the Gary NIT to be deficient in this respect. The differences concerning the coverage and enforcement aspects of the rules were even more pronounced. One must, of course, be appropriately cautious in interpreting these comparative findings. Nevertheless, they are at least consistent with the judgment that the regimen of an income maintenance program of the sort that was implemented in Gary and elsewhere is, from the beneficiaries' viewpoint, a preferred alternative to AFDC.

Perceptions of Administrative Procedures and Practices

As was intimated previously, in cases where there was some question or problem regarding a family's compliance with requirements--particularly those associated with income reporting--the standard practice in Gary and the other experiments was to contact the affected family as often as necessary to resolve the situation. To determine the extent

that the Gary participants felt the program operators were over-reaching in executing this practice, they were asked whether they thought the program staff "bothered" them more than or as much as necessary about their income reports and "other things." Only 14% claimed they were bothered more than necessary about their income reports, and an even smaller porportion (11%) felt they were bothered more than necessary about other things. Although we have yet to test the possibility, one suspects that the more exasperated participants were less conscientious in filing their income reports.

A related issue in the area of standard administrative practice is the extent to which such practices were perceived as being legitimate or justified. Conceivably, one might recognize the necessity of certain administrative practices but still question them on ethical or moral grounds. In this connection, only 14% of the Gary peticipants believed the program operators did some things that they did not have a right to do. One might wonder, of course, whether these findings reflect some characteristic deference to officialdom or a genuine appreciation of the necessity and legitimacy of the program's administrative procedures.

Propensity to Contact Program Staff

As observed earlier, most of the contacts were initiated by the participants themselves--either in connection with their income re-

ports or some other matter. It is instructive to note, first, that a full 66% of the participants claimed they never experienced any difficulty with their income reports and 61% indicated they never had a question or problem concerning their transfer payment. In other words, to the extent that such difficulties imply a need to contact the program staff, these data suggest that a solid majority of the participants never felt a need to do so. Among the participants who did experience reporting- or payments-related difficulties, about 67% stated they "always" or "almost always" brought their difficulties to the attention of the program staff. Thus, it appears that approximately a third of Gary families were not inclined to present their concerns or problems to the staff even though they had a problem. From this we can at least tentatively recommend that program operators be somewhat more active in encouraging solicitations and spot-checking to uncover possible unreported problems. It is not uncommon to find that some participants in public programs are reluctant to solicit help from program operators when they need it.

Perception of Worker Performance

One factor that surely influences participant-initiated contacts-- although the direction of effect may be uncertain--is the extent to which program staff actually solve the problems that participants present to them. In this regard, some 65% of the Gary participants said

they were "always" or "almost always" satisfied with the way in which the staff dealt with their concerns; the rest were rather critical of the type of service they received. Unfortunately, we cannot present any analysis to the propensity of the two groups to solicit assistance, but it seems unreasonable that the more satisfied participants would be more active in their solicitations.

A related dimension of worker performance has to do with the exercise of discretion. Workers in a social program typically function within a range of discretion when servicing the needs of clients. In a given instance, this discretionary behavior may or may not be to the participant's benefit. Feedback data on this issue show that about half of the Gary participants thought the program staff performed some services for them that were, in their judgment, above and beyond the call of duty. Conversely, only 21% believed that staff neglected to do things that should have been done.

In short, the findings on perceptions of staff performance are more often favorable, but suggest that, from the participants' standpoint, relations with program staff could have been better.

AFDC/Non-AFDC Comparisons

Many families in the income maintenance experiments were on AFDC prior to enrollment in the NIT. We have already commented upon certain differences in compliance behavior of the AFDC and non-AFDC families

and, in respect to program rules, compared certain of the AFDC participants' attitudes toward the NIT and toward AFDC. We now want to extend the earlier discussion by examining briefly the effects of pre-experimental AFDC status on some of the other kinds of participant perception. Given the wide-spread criticisms of AFDC, one might expect that families having fresh experience with that system might tend to perceive the income maintenance program more favorably than other families.⁶

In general, however, the data do not reflect any consistent pattern of differences. On the question of staff-initiated contacts, for example, AFDC and non-AFDC families were almost equally inclined to believe that they were not "bothered" more than necessary. On the other hand, the AFDC families were significantly more likely to experience payments- or income-related difficulties. The latter finding would appear to contradict an earlier result which showed that the income reports of non-AFDC families were less often immediately processable. At the moment, we cannot fully explain this inconsistency. However, it could be that the AFDC families were simply more willing to acknowledge problems. The data also indicate that the AFDC families were somewhat more likely to present their problems to the program staff. It is possible that these families were more accustomed to dealing with officialdom, and consequently, less reluctant than non-AFDC families to reach out in time of need.

Summarizing the participant perceptions results, the data indicate that the Gary participants had rather definite views about the operation of the experiment and their experiences with it. A worthwhile objective of future research would be to investigate the influence of selected perceptions variables on the objective measures of participant compliance and contact rates.

SUMMARY AND CONCLUDING COMMENTS

The present study examined participation in the income maintenance experiments, focusing on the participants' compliance with income reporting requirements, contacts with program staff, and perceptions of program operations and administration. In summary, we found that in all of the experiments--Seattle, Denver, and Gary--only a small proportion of income reports were filed late, the incidence of on-time filing being well over 90%. Second, a large percentage of the income reports--due to errors, inconsistencies, or omissions--were not immediately processable. However, the application of a more liberal definition of processability resulted in a drastic reduction in the frequency of unprocessable reports. Some variations across experiments was observed. Third, staff-participant contacts were most frequent in Denver (73 per 100 IRF's) and least frequent in Seattle (34 per 100 IRF's). In all three experiments, the majority of contacts were participant-initiated

and involved payments or income reporting difficulties. Last, the perceptions on data showed that the Gary participants, while critical of some aspects, were, for the most part, favorably disposed toward the operation and administration of the experiment.

Our results clearly indicate that participants in a national income maintenance program requiring monthly reporting of income can and will comply with the requirement for timely filing. On the question of the processability of those reports, however, the analysis suggests that there may be a number of problems. The extent of these problems will depend largely on the complexity of the reporting form and the family's income situation. Since reports can be simplified and still be functional, and since families with complex income circumstances are less likely to be eligible for a national program than for the experimental NIT, the incidence of reporting problems should be lower in a national program.

With regard to client-staff interaction, the variability of contact rates between sites and mixed findings preclude any clear-cut statements about the implications of such contacts for the staffing needs of a national program. However, it does appear that the level of contact activity would be substantial. At the same time, we suspect that frequent staff-participant contacts might be beneficial in terms of promoting participant compliance and satisfaction. If this turns

out to be true, our analysis of participant perceptions suggests that an income maintenance program can be devised which, despite its demands, is viewed favorably by participant families.

Finally, in addition to believing that the present results are instructive, we also think it is significant that the examination was done and was reasonably fruitful. Thus, planners and operators of both ongoing programs and experimental ventures should take note that the behaviors and attitudes of participants vis-a-vis a social program can be studied effectively and the results used to plan new programs and improve the operation and administration of existing ones. Such studies should be useful, inasmuch as these behaviors and attitudes are largely a function of the operational features of the program itself.

NOTES

1. The more thorough-going study of participant compliance and contacts in Seattle/Denver is found in George Carcagno, "Income Report Processing and Family Contacts in Income Maintenance Programs" (Mathematics Policy Research, 1976). The detailed Gary analysis is in Heather Ruth, John Hall, and Don Lara, "The Impact of Administration Change on the Income Reporting Behavior of Participants in the Gary Income Maintenance Experiment" (Mathematica Policy Research, 1978). The present examination is largely a condensation of these two studies.

2. The examination of this topic is derived from the larger study, Billy J. Tidwell, "Participant Perceptions and Evaluations of the Gary Income Maintenance Experiment: Some Initial Findings" (Mathematica Policy Research, 1976).

3. We should point out that preliminary findings from the Colorado Monthly Reporting Experiment lend further support to our own conclusion. See Alan Hershey, J. Jeffrey Morris, and Robert G. Williams, "Colorado Monthly Reporting Experiment and Pre-test: Preliminary Research Results" (Mathematica Policy Research, February 1977).

4. This judgment was based on the sample income report form developed by the IRS Task Force on the Administrative Feasibility of an Income Maintenance Program.

5. This conclusion is supported by earlier findings of a participant knowledge study, in which awareness and understanding of the rules was investigated. See Billy J. Tidwell et al., "Participants' Knowledge and Understanding of the Gary Income Maintenance Experiment" Gary, Indiana: Indiana University Northwest, 1975.

6. The comparisons are limited to female-headed families (n=419), inasmuch as male-headed or two-parent families are not eligible to receive AFDC benefits under Indiana state law.

DISCUSSION

MODERATOR: Let's take five minutes here and entertain a few questions.

AUDIENCE: It seems as though maybe I just don't know enough about what is going on here. You don't have micro data on the participants' perceptions do you?

MODERATOR: Yes, we do.

AUDIENCE: Like on the enforcement issue, did you check to see how many of the ones that perceived inequitable enforcement were also the ones that didn't understand how the rules worked?

MODERATOR: I'm glad you raised that question. There are a number of research objectives connected with the perceptions inquiry that have not been pursued. However, we have the data to do precisely what you are saying as well as to examine the extent to which perception variables influence other experimental outcomes, such as labor supply behavior. We have yet to do these kinds of investigations, but would be glad to get copies of the relevant materials to you as soon as the work is done.

AUDIENCE: Another thing of interest to me is the question that was raised yesterday about the role of community culture and values--like welfare people are automatically "bad." Did you go after anything like a community norm or opinion, as it might make a difference between, say, Gary and Denver, and try to profile that?

MODERATOR: No, not in that specific sense. In this connection, I might mention that while there has been a good deal of interest in studying "community effects" in all of the experiments, there has been very little work in that regard. What we did do as part of this perceptions inquiry was to ask the respondents about their perceptions of the community's attitude toward them and the experiment--that is, after we determined that some portion of the community knew about their participation. For the most part, the families did not believe they were perceived in a negative way by the community.

MODERATOR: We will take one more question for now.

AUDIENCE: I have a question related to administrative costs. That is whether costs would be prohibitive in terms of enough staff people to handle all of the problems that you mentioned and the family contacts that are necessary and unavoidable--despite the attractive features, prima facie, of monthly reporting.

MODERATOR: One comment about the monthly reporting system will, I believe, answer your question. This is an administrative matter that can be handled relatively easily. The reporting regimen is simply one aspect of the income maintenance experiments being imposed on an existing welfare system. It is not necessary to modify the whole system but merely to simplify the reporting procedures.

ALTERNATIVE METHODS OF SCALING WELFARE
PAYMENTS TO FAMILY COMPOSITION

by

Terry R. Johnson
Senior Economist, SRI

and

John H. Pencavel
Senior Economist, SRI

Government income transfer programs designed to raise the standard of living of the family unit--(e.g., Aid to Families with Dependent Children (AFDC), the Food Stamp Program, and some unemployment insurance programs)--generally grant benefits according to family size and composition. The index numbers used to adjust the grant go by the name of family equivalent scales.

In Table 1 we present a variety of family equivalent scales used in actual programs or suggested in the literature. Although these scales may appear to be quite similar, the differences in the benefits for differently structured families can be substantial. A family of four, with two adults and two children would receive the basic grant, as set by the legislature. Depending on the number of children and in some instances, their ages, this amount would be increased or decreased

by the index provided. For example, an income transfer program with a benefit level of \$3,800 per year for a family of two adults and two children would provide \$4,104 for a family of two adults and four children, if the program uses the index suggested by Kapteyn and van Praag (row 1), whereas the corresponding benefit level using the scale given in the AFDC example (row 7) would be \$5,510.

TABLE 1
ALTERNATIVE ADULT EQUIVALENT SCALES

<u>Source</u>	<u>Number of Children</u>				
	<u>0</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>
(1) Kapteyn and van Praag ¹	.78	.80	1.00	1.06	1.08
(2) Orshansky ²	.64	.70	1.00	1.18	1.32
(3) Jackson ³	.60	.82	1.00	1.16	1.32
(4) Bojer ⁴	.68	.84	1.00	1.16	1.32
(5) Seneca and Taussig ⁵	.63	.89	1.00	1.10	1.20
(6) Food Stamp Program	.55	.79	1.00	1.19	1.36
(7) AFDC, Denver, Colorado July, 1969	.55	.78	1.00	1.23	1.45
(8) SIME/DIME	.62	.83	1.00	1.12	1.23
(9) Mean of above 8 scales	.63	.82	1.00	1.15	1.29

The sensitivity of the benefit level to the scale used suggests an important policy question: In designing a national income maintenance program, how should the benefit level be scaled by family size and composition? Since all of the income maintenance experiments have scaled the support level by family size, it is surprising that the issue of what is an appropriate indexing scheme has not received any systematic treatment from researchers associated with the negative income tax experiments.

In this paper we outline a procedure for developing alternative family equivalent scales and provide estimates of these scales based on data from the Seattle and Denver income maintenance experiments (SIME/DIME). We report estimates of three such alternative scales for husband-wife families, leaving analysis of single-parent families for a future date.

RATIONALE BEHIND FAMILY EQUIVALENT SCALES

The logic behind the scaling of benefits to a family size and composition is rarely spelled out convincingly. Since a husband and wife are not obliged to have children, one might wonder why one should vary welfare benefits to family size any more than by other characteristics that parents can choose to acquire. One response to this takes note of the fact that the children themselves have not been a party to this choice, but then is it the welfare of the parents that is the object of scaling benefits to family size or the welfare of the children?

On this issue we know of no information on the extent to which supplementary benefits are spent on items that are largely consumed by parents as distinct from items that are more specific to the children. The tacit assumption for treating the family as the relevant unit for redistributing income seems to be that the head of the family internalizes the welfare of different family members and acts like a benevolent despot with respect to expenditures within the family.⁶

Scaling benefits to family size and composition can also have important consequences. To the extent that there are different benefits available to families according to their characteristics and these are not offset by actual costs to the family of maintaining these characteristics, the scales will pose different incentives to recipients to adopt these characteristics. For example, family equivalent scales that favor single-parent families could encourage marital dissolution. As another example, nations desiring a population increase may design transfer programs that provide for extremely generous benefits to families with more children.

Notwithstanding these issues, we observe that the prevailing sentiment appears to be that, since expenditures on necessities such as food and clothing rise with the number and perhaps the age of children,⁷ families eligible for receiving welfare benefits require additional financial assistance to meet these "costs" of satisfying the minimum

requirements of children. Assuming, then, that welfare payments are to be adjusted according to the number of children in the family, this paper addresses some of the issues that arise in determining adult equivalent scales. Our procedure in this paper, therefore, is not to challenge the notion that families with children require greater welfare benefits than child-less families so that the former can meet some of the expenditures associated with the raising of children, but to examine the logical implications of this notion and to offer some estimates of these scales that are derived from the behavior of the SIME/DIME families.

ALTERNATIVE METHODS

Essentially there are three methods for estimating family equivalent scales. The first consists simply in asking how much extra money they require to meet the minimum costs of raising a child. The Kapteyn and van Praag family equivalent scale (row 1, Table 1) was derived in this way. For well known reasons, the problems in interpreting the answers to such questions leave economists reluctant to place much weight in the derived index numbers.

A second procedure involves determining minimum or adequate nutritional requirements for different family types. This procedure was adopted by Orshansky in her determination of poverty levels and they subsequently formed the basis of the Social Security Administration's

equivalency scales (row 2, Table 1). Her procedure shares some of the characteristics of the previous method: instead of asking families to estimate the minimum cost of raising a child, Orshansky, in effect, asks a group of nutritionists to estimate it. In both cases the estimate of minimum requirements is likely to involve not merely a physiological concept of requirements, but also a social concept.

The third procedure and one that has greater appeal to economists is to rely on the observed pattern of expenditures across families of different sizes and composition. Most frequently, these procedures expand upon Ernst Engel's original idea of drawing inferences from the fraction of family income spent on food. For instance, Jackson (row 3, Table 1) estimates family equivalent scales based on the assumption (supported by considerable evidence) that the income elasticity of demand for food is 0.5, independent of family size and composition. In addition, she makes the strong assumption that families spending the same proportion of their incomes on food attain an equivalent level of total consumption.

All of these scales build exclusively on the consumption of commodities. Yet, if these scales are supposed to reflect the relative costs of meeting the minimum requirements of children, it is curious that they omit reference to one of the most important factors in attending to the welfare of children--the parental time required for child care.⁸

In forming comparisons between, say, two adults with no children on the one hand and two adults with one child on the other hand, it seems inappropriate to ignore the fact that in the second family at least one parent (more often, the wife) will have to devote some time to caring for the child. Moreover, this time has earning potential that in the child-less family could be allocated to the market place.⁹

A casual glance at the data presented in Table 2 supports the proposed association between the working behavior of each adult in husband-wife families and the number of children in the family.¹⁰ These data summarize work effort in the second year of the experiment for control families in both Seattle and Denver. Rows 1(b) and 1(c) present the fraction of husbands and wives, respectively, who worked at any time during the year (we call this the employment rate). Row 1(d) measures the fraction of families in which both the husband and wife worked some time during the year. Rows 2(b) and 2(c) measure the average hours worked by the husband and the wife for those families in which both the husband and the wife worked positive hours during the year.

The employment rate of wives with either one or two children is slightly larger than that of child-less wives, but then the employment rate falls with the number of children. The employment rate of husbands without children is almost ten percentage points lower than the rate for husbands with one child. This relationship between the hus-

TABLE 2: WORK BEHAVIOR AND NUMBER OF CHILDREN

No. of Children	0	1	2	3	4	5-6	TOTAL SAMPLE
1. Employment Rate							
(a) No. of Families	107	155	231	165	82	45	785
(b) Husband	.851	.936	.935	.976	.915	.911	.929
(c) Wife	.570	.600	.597	.576	.549	.422	.575
(d) Husband and Wife	.467	.581	.571	.558	.524	.400	.541
2. Hours Worked							
(a) No. of Families	50	90	132	92	43	18	425
(b) Husband	1584	1949	1949	2084	2027	2166	1953
(c) Wife	1413	1352	1089	1249	1069	896	1207
(d) Husband and Wife	2997	3300	3038	3333	3096	3062	3160

Notes: The employment rate measures the fraction of husbands (or wives) who worked a positive number of hours in the second year of the experiment. The row 1(d) measures the fraction of families in which both the husband and the wife worked positive hours during the year. The hours worked data measure annual hours of work for those families in which both the husband and the wife worked at all during the year. Row 2(d) is simply the sum of rows 2(b) and 2(c) in any column.

band's work behavior and number of children is also evident in the hours worked: husbands with one child work almost 400 hours more in a year than husbands without children. As for wives, hours at work fall as the number of children increase from zero to one and from one to two, but then from two children to three children the wife's hours rise considerably. Though this may appear surprising at first sight, it is consistent with the fact that child raising in some families is assisted by the presence of older children. That is, it is not merely the number of children, but the age of children that is relevant for the wife's hours of market work. In this paper we build on this apparent relationship between hours of work of the husband and wife and the number and age composition of their children in order to derive various adult equivalent scales.

Our procedure is to follow the conventional one in economics of distinguishing objects of choice from constraints (or predetermined variables) and of assuming that families make their choices so as to do the best they can with what they have. Naturally what is classified in the category of constraints and what is in the category of decision variables depends upon the particular questions under examination. Since in this paper we take our cue from current government programs and concern ourselves with the construction of family equivalent scales given the number of children, we treat the size and composition of the family as predetermined. This is in the spirit of existing welfare

programs that identify families requiring assistance on the basis of some short-term criteria (such as weekly or annual earnings) and that take such life-time decisions as the size of family as given for the purposes of the program. No doubt on the other occasions, this assumption should be relaxed and fertility decisions should be integrated with consumption and work decisions.

We assume that all families of a given size and structure have the same preferences for consumption and work that, in response to differences among families in relative prices and wages facing them, they will substitute among commodities and work time to maximize their overall welfare.¹¹ Of the many possibilities we could choose for modeling family preferences, we have selected a functional form that permits us to distinguish between family equivalent scales which compensate families of different sizes for the costs of (1) satisfying minimum requirements and (2) achieving some reference level of welfare which may be different from the welfare enjoyed at the "minimum requirements" level.

DEVELOPMENT OF ALTERNATIVE FAMILY EQUIVALENT SCALES

In this section we describe and estimate three alternative family equivalent scales that involve slightly different concepts of compensation. This discussion provides a general overview of our unpublished

paper which includes an appendix stating the procedures followed in making our estimates.¹²

Our approach to the construction of family equivalent scales appropriate to a negative income tax (NIT) program relies on the observed expenditures and working behavior of the SIME/DIME husband-wife families. We infer from the behavior of these families which expenditures can be regarded as satisfying minimum requirements and which are "discretionary" expenditures--for purchases above the minimum level. We assume that the costs required to meet minimum consumption levels are the same for all families of a given size and composition and we estimate by how much these consumption requirements change with family size.

Correspondingly, we infer from their working behavior the extent to which each adult chooses to work less than his or her maximum feasible working hours given the time needed for rest and for children. That is, income is spent not only in consuming commodities in the conventional sense, but also in consuming leisure time. In other words, some income is foregone by not working the maximum feasible number of hours. Again, we assume that all husbands in a family of the same type have the same maximum feasible hours of work, but that this earnings potential may change with family size. We also assume that all wives in a family of the same type have the same maximum feasible hours of work, but recognize that earnings potential may fall with family size as

available time is directed from the labor market to the care of children. For both the husband and the wife, we estimate the extent to which the earnings potential of each parent is affected by the presence of children. In general, we regard families as having to make certain minimum expenditures for consumption purposes, and also, as having to devote certain minimum amounts of time to the raising of and caring for children. The actual level of well-being enjoyed by a family is measured by the extent to which its consumption expenditures and leisure time exceed these minimum requirements.

The first family equivalent scale we present does no more than scale payments so as to compensate families with children for the increase in their minimum necessary consumption expenditures. This type of compensation is quite conventional and our estimates of this scale are given in row 2 of Table 3 with the designation "compensation for necessities in consumption." For comparison we also present the family equivalent scale used in SIME/DIME in row 1. As is evident, our estimates of the payments required to compensate families for necessary consumption purchases are similar to, although a little less than, the scales used to adjust the support level for SIME/DIME families.

A family equivalent scale may also be devised to take account of the loss in potential earnings that occurs when adults (principally the wife) divert time from work to the raising of children. This sort of compensation is less conventional than the first but no less real.

To illustrate, suppose that all women work the same number of full-time hours and that those women with young children would have to pay for the services of child care during their working hours. If the guiding principle of any family equivalent scale is to compensate families for children-related expenditure requirements, then it would necessarily take account of at least some of these child care expenditures. If it is appropriate to compensate families for those child care expenditures, then surely it is also appropriate to compensate in part those mothers who forego some of their earnings in order to allocate their own time (rather than buying the time of others) to children-related responsibilities. Otherwise, the scales are tacitly penalizing those mothers who reallocate either working time or leisure time to the raising of their children. Hence, the logic of a family equivalent scale that does no more than compensate families for the costs of satisfying the minimum requirements of children implies that child-related expenditures of time should be treated symmetrically with child-related consumption expenditures.

If the support levels in an NIT plan were used to compensate families for increases in minimum consumption expenditures as well as losses in foregone earnings, then the level of discretionary full income across families of different size and composition, but facing the same wage rates and prices, is left constant. We call scales derived in this way

constant discretionary full income family equivalent scales. Our estimate of this scale is presented in row 3 of Table 3. Clearly it implies a much greater per child compensation than the consumption-specific scale in row 2 since it incorporates not merely the increase in minimum consumption expenditures with respect to children, but also the loss in earnings potential. If a family of two adults and two children were granted a support level of \$3,800 per year, a support of \$4,598 would have to be provided to a family with one more child in order to compensate this family fully for the increased minimum consumption requirements plus the decreased earnings potential that is associated with this extra child. (This is \$94 more than would be needed to compensate families merely for the additional minimum consumption requirements.) Observe that the compensation required increases with the number of children at a decreasing rate, reflecting "economics of scale" in family size.

A family equivalent scale may also be directed to a somewhat different end, one which recognizes that the well-being of a family should not equal its actual income nor its discretionary income, but that it should depend upon both the goods consumed and the leisure time enjoyed by the family. With well-being gauged in this way, the well-being of, say, a family of two adults and two children on a particular NIT plan may serve as the reference point for scaling NIT payments with respect to family size. Then one may estimate the level of support a second

TABLE 3
ESTIMATES OF FAMILY EQUIVALENT SCALES
(TWO ADULT FAMILIES)

	<u>Number of Children</u>					
	<u>0</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>
(1) SIME/DIME	.62	.83	1.00	1.12	1.23	1.32
(2) Compensation for necessities in consumption	.69	.88	1.00	1.08	1.15	1.20
(3) Constant-discretionary full income	--	.65	1.00	1.21	1.37	1.50
(4) Constant utility	.76	.91	1.00	1.05	1.29	1.13

NOTE: The reference utility level in row 4 was taken to be a childless family facing mean gross wage rates on an NIT plan involving a tax rate of .50 and a support level of \$3,800. Similarly, changes in discretionary full income (the basis for the statistics in row 3) was normalized with respect to a family with two children receiving a support level of \$3,800. Since the scale in row 3 is based on changes in discretionary full income across family composition, no element can be completed for the first column. For the scales presented in rows 3 and 4, the first child is assumed to be less than 10 years of age while the second child is assumed to be between the ages of 11 and 16. The scales given in rows 1 and 2 do not distinguish between the ages of children.

family, consisting of two adults and three children, needs in order to enjoy the same well-being as the first. A family equivalent scale that leaves the well-being (or utility) of families of different sizes and composition no different from that of the reference family is designated a constant utility family equivalent scale.

Our estimate of this scale is presented in row 4 of Table 3. The reference family is given a support level of \$3,800, with a tax rate of 50%, and has two adults and two children. If this family receives \$3,800, then to maintain the same level of well-being, a family with three children should receive 5% more, or \$3,990. This compensation which maintains utility constant reveals the smallest increase with respect to family size than any of our scales. If well-being or utility depended only upon consumption expenditures, then clearly for any level of well-being above the minimum necessary consumption our constant utility index cannot be less than that in row 2 which measures the compensation required for increases in minimum necessary consumption. In fact, we have specified well-being to depend not only upon the consumption of commodities but also upon the leisure time enjoyed by the husband and by the wife. As family size increases, so the time spent by the wife at market work falls. Some of this increased home time may be regarded as necessary for the caring of children while some of this reallocated time has pure utility-augmenting effects of increased

leisure. The fact that our constant utility index increases more slowly with respect to children than the index that merely compensates for minimum necessary consumption requirements reflects the dominance of these utility-augmenting effects of the women spending more time at home.

To illustrate our results in a slightly different manner, we present in Table 4 the implied support levels for families with various numbers of children based on the family equivalent scales given in Table 3 and assuming a two-adult two-child family would receive \$3,800. As can be seen, the support levels based on the conventional notion of simply meeting the increased consumption expenditures of children are the

TABLE 4

IMPLIED SUPPORT LEVELS BASED ON FAMILY
EQUIVALENT SCALES GIVEN IN TABLE 3
(TWO-ADULT FAMILIES)

		<u>Number of Children</u>					
		<u>0</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>
(1)	SIME/DIME	\$2356	\$3154	\$3800	\$4256	\$4674	\$5016
(2)	Compensation of Necessities	2622	3344	3800	4104	4370	4560
(3)	Constant-discretionary full income	--	2470	3800	4598	5206	5700
(4)	Constant utility	2888	3458	3800	3990	4142	4294

NOTE: See Table 3 and text for further explanations.

closest of our three scales to the SIME/DIME levels. These are quite different from the support levels provided when we compensate families with children fully for their loss in total feasible work time in order to maintain discretionary full income constant. For example, a family with two adults and five children would be eligible for \$5,700 if discretionary full income is to be held constant, but only \$4,500 if families are to be compensated only for meeting the minimum consumption expenditures of children--a difference of \$1,140. Finally, the table gives the support levels assuming the utility of families of different sizes is to remain constant. This form of compensation results in support levels that increase least with respect to the number of children. The table shows how sensitive the support levels are to the method of determining how to compensate families for the extra costs of increased family size.

CONCLUSION

Adult equivalent scales are a feature of practically all welfare programs and would almost certainly be an issue in any national NIT plan. We have presented three scales in this paper, each one corresponding to a slightly different concept of compensation. Our procedure has not been to contest the tacit proposition that underlies these scales, namely, that parents require some form of compensation for the

additional costs of raising children. Instead, we have addressed the logical implications of this proposition: these children-related costs do not merely cover the increased consumption expenditures associated with larger families, but they also embrace the redirection of parental time away from leisure or from work in the labor market to the caring for and supervision of children.

Family equivalent scales based on the conventional approach--which calls for meeting only the increased minimum consumption expenditures of children--imply less compensation for children than scales that take account of the value of time reallocated to the caring for children. The constant discretionary full income scale constructed in this paper goes so far as to compensate families with children fully for their loss in total feasible work time and this yields a scale that increases faster with respect to family size than any other of which we are aware. A scale that implies the least compensation is our constant-utility family equivalent scales which indexes benefits such that the utility, or well-being, of families of different sizes is kept constant.

Our impression is that the time costs of children have not merely escaped quantification, but have not even been recognized in principle as a relevant dimension of the construction of family equivalent scales. Our purpose in this paper has been to remedy this deficiency and to

provoke a discussion of these issues by the relevant policy makers.

NOTES

*Of course, as is always the case in empirical studies of this kind, the true authors are the research assistants. In our case, this is Susan McNicoll and John Peterson. They assume full responsibility for all remaining errors. Comments by Michael Keeley, Robert Pollak, Philip Robins, and Louise Smith on earlier drafts of this paper were most useful to us and are gratefully acknowledged.

1. Arie Kapteyn and Bertrand van Praag, "A New Approach to the Construction of Family Equivalence Scales," European Economic Review 7, no. 4 (May 1976): 313-36. The adults are aged 27 years for the man and 25 years for the woman.
2. Mollie Orshansky, "Counting the Poor: Another Look at the Poverty Profile," Social Security Bulletin (January 1965): 7-9.
3. Carolyn A. Jackson, "Revised Equivalence Scale for Estimating Equivalent Income or Budget Costs by Family Type," BLS Bulletin No. 1570-2 (U.S. Department of Labor, November 1968). The age of the head is 35-54 years with the older or oldest child aged 6-15 years.
4. Hilde Bojer, "The Effect on Consumption of Household Size and Composition," European Economic Review 9, no. 2 (May 1977): 169-193.
5. Joseph J. Seneca and Michael K. Taussig, "Family Equivalence Scales and Personal Income Tax Exemptions for Children," Review of Economics and Statistics 53, no. 3 (August 1971): 253-62.
6. P.A. Samuelson, "Social Indifference Curves," Quarterly Journal of Economics 70, no. 1 (February 1956): 1-22.
7. By "necessities" we mean commodities with income elasticities of demand of less than unity. For evidence and discussion of the association between estimated income elasticities and household size, see S.J. Prais, and H.S. Houthalker, "The Analysis of Family Budgets" (Monograph No. 4, University of Cambridge Department of Applied Economics,

1955); H. Bojer, "The Effect of Consumption and Household Size and Composition," European Economic Review 9, no. 2 (May 1977): 169-73; C.E.V. Leser, "Income, Household Size and Price Changes 1953-1973," Oxford Bulletin of Economics and Statistics 38, no. 1 (February 1976): 1-10.

8. See C. Vickery, "The Time-Poor: A New Look at Poverty," Journal of Human Resources 12, no. 1 (Winter 1977): 27-48; and also I. Garfinkel and R. Haveman, "Earnings Capacity, Economic Status and Poverty," Journal of Human Resources 12, no. 1 (Winter 1977): 49-70.

9. There are, of course, other options--the use of mothers-in-law or the hiring of babysitters--but the principle remains: scarce resources have to be spent in child-caring that could be allocated to other purposes.

10. Although these numbers are suggestive of an association between work behavior and family size, the reader should be reminded that such cross-tabulations ignore the influence of all other variables that affect hours or work.

11. In the empirical work we relax this strong assumption and allow for differences in preferences for consumption and for work by recognizing that individuals are creatures of habit and that current choices depend on previous consumption and work decisions.

12. Johnson and Pencavel, "Alternating Methods of Scaling Welfare Payments to Family Composition in a Negative Income Tax Program" (Stanford Research Institute mimeo, May, 1978).

DISCUSSION

CHRISTINE KISSMER: I have a little trouble understanding the constant utility family equivalent scales. What do you mean by "level of well-being?"

PENCAVEL: Let's suppose that I could attach a thermometer to you and I could measure your satisfaction, your happiness, with what you are consuming in a given year; and suppose that I measure it for another woman, who has to care for a child, while we presume that you do not. Then I can see how much that temperature is different for the woman with the child. If you are the reference person in this illustration, then I want to give this other woman just enough money to raise her temperature such that it is just equal to yours. In our paper our reference family was a family on the NIT program on a support level of \$3,800 and facing a tax rate of 50%. We try to make every family as well-off as that average family.

JEAN WRIGHT: How do you define the "minimum necessary" level of well-being?

PENCAVEL: Well, it tends to get a little more technical. In accounting for differences in consumption expenditures and in accounting for differences in work, we assumed that there is a difference between actual consumption and minimum-necessary consumption. That estimate of minimum necessary consumption falls out as an estimated parameter in the fitted equations.

DARDELL MCFARLIN: Doesn't the scale have a problem inherent in it, in that once we achieve a certain level of satisfaction, then we are not satisfied any longer, and we readjust our expectations upwards. Your scale would be continually being adjusted upwards as people approach happiness.

PENCAVEL: Well, there is a lot to that. You can ask people in different countries a question about whether they are happy or not, and it turns out that the fraction of the people who say "very happy" tends to be the same whether you go to a very, very poor country or a very, very rich country. Happiness seems to be a function of relative

income rather than absolute income. But this need not induce a revision of the scale itself since the scale adjusts payments relative to some reference level. Instead, it is the reference level which will be continually revised upwards.

DICK RITCHIE: What sort of consideration, if any, did you give to the fact that within the NIT you have a tax rate that, in fact provides for income considerably above those levels if a family has earnings?

JOHNSON: You are right that the absolute level of income would matter. In our estimates, we used both 50% and 70% tax rates--and the relative scales were virtually unaffected by what tax rate you use.

WRIGHT: Most of the scales are constructed on the basis of need, either implicit or explicit, whether it is expenditure or whatever. The question is, should there, in fact, be other considerations made when constructing these scales?

PENCAVEL: Yes. We would like to hear suggestions on that. Need itself is not a well-defined concept, and that is all most scales ever use. We used more, as you know. I don't know anybody who has even recognized the principle that perhaps compensation should be required for time. Also, in estimating our equation, we do distinguish between whether income comes from working or whether it comes from other sources. We assumed that it will have a very different effect on well-being if a person has to work 60 hours a week to get that income.

WRIGHT: How would you construct equivalent scales using some of these assumptions with respect to single-parent families?

PENCAVEL: We are doing that now.

JOHNSON: We have a similar type of model developed from a very large sample of female-headed families in SIME/DIME.

PENCAVEL: In fact, the argument that we present about child care costs presumably applies with particular force to a single woman with children.

WRIGHT: The Ullman Bill says that "Size of benefit should not be scaled to size of family, because wages do not adjust for increases in children." But I think I heard you say that the parent who is primary earner does increase working hours and, therefore, earnings when there is an additional child; is that correct?

PENCAVEL: The wage rate may not change, but total earnings will if the extra child induces the primary earner to work more or less time.

WRIGHT: The wage rate may change?

PENCAVEL: That is true, it may. In other studies they have found that--holding everything else you might want to think of constant and even accounting for various selectivity biases--married men seem to seek a higher wage rate than single men. So one has to be careful, therefore, in looking at that cross-tabulation in Table 3. It simply showed, not holding anything else constant, men with more children tend to work a little bit more. Whether that is the consequence of having children or larger family size, or whether it is associated with some other characteristics that we have not measured (such as higher wage rate), I do not know.

DAVID FORTE: I am really curious about the rationale that you used when you decided to compensate for child care time. What you are really saying is that by having a child, you are in fact performing a service of value in terms of socialization and education of that child?

PENCAVEL: Some people would indeed put it like that.

FORTE: Therefore, you solve an unemployment problem.

PENCAVEL: As Terry said in the beginning, we have taken our cue from government programs that seem to think that some families do need compensation for extra children. Given that approach, we have proposed another way to go about it.

WELFARE REFORM AND STATE FISCAL FLOWS

by

Myles Maxfield, Jr.
Senior Economist, MPR

and

David Edson
Research Associate, MPR

Much of the discussion of the variety of proposed welfare programs centers around the impacts of the reforms on federal expenditures, the national welfare caseload, distribution of income, and the behavior of recipients. Welfare reform proposals also affect state and local budgets, and the state and regional distribution of personal disposable income. One of the aims of this paper is to show how a specific welfare reform proposal can affect state budgets. A second goal of this paper is to show how the same welfare reform might affect the distribution of disposable personal income among residents of each state.

The welfare reform program evaluated in this study consists of a negative income tax (NIT) which replaced the Aid to Families with Dependent Children (AFDC), Food Stamps, and Supplemental Security Income (SSI) programs. The impacts of the reform on state welfare expenditures and on the disposable income of residents of each state are examined separately when (1) the NIT is federally financed, and (2) when the NIT is jointly financed by the federal government and the states, with states supplementing the basic

federal grant to assure that no recipient under the present programs would receive lower benefits under the proposed plan (i.e., states "hold harmless" families now receiving welfare).

The impact of the reform on the disposable income of state residents is estimated assuming that additional expenditures required by the NIT are (1) deficit financed; (2) financed through a personal income tax surcharge; and (3) financed through a surcharge on both personal and corporate income tax. State budgets are assumed to remain balanced, with any additional expenditures offset by taxes.

THE MICROSIMULATION MODEL AND DATA BASE

The questions raised in the previous section are answered using The Micro-Analysis of Transfers to Households (MATH) model maintained by Mathematica Policy Research.¹ The MATH model simulates federal and state income tax programs, the social security tax, AFDC, food stamps, SSI, and the hypothetical NIT. Program regulations are applied to a cross-sectional sample of residents in the 50 states and the District of Columbia in order to compute each family's tax liabilities and transfer receipts. The state or federal aggregates are estimated by summing the relevant family amounts for the residents of the state or nation.

The data base for the study is the Survey of Income and Education (SIE), which was fielded by the Census Bureau in Spring 1976. The SIE is similar to the monthly Current Population Survey (CPS). However, the SIE is larger, including roughly 150,000 households as compared with the CPS's 50,000. The size and design of the sample makes it possible to draw a number of statis-

tical inferences about each state, an essential quality for a study of state distribution of welfare reform impacts.

Both the SIE and CPS contain a series of questions about family demographic characteristics and structure. Respondents are also asked to list the income of each person in the family for the previous calendar year, in this case January through December 1975. Both surveys ask about each person's labor market experiences, for the previous year and for the survey month. In addition to the CPS questions, the SIE contains information pertaining to the health and education of family members.

The nature of the questionnaire imposes some limitations on the simulation model. The incomes recorded are annual, so simulating welfare programs with differing accounting periods is difficult. Simulation of several of the programs requires income, demographic, and family structure information simultaneously; however, the income amounts were accrued in the calendar year prior to that to which the demographic and family structure pertains. This discrepancy could bias some of the results.

THE NEGATIVE INCOME TAX

The type of NIT simulated for the study would replace three of the current transfer programs with a single income transfer system integrated into the federal tax structure. The program would establish a federal income floor which varies only by family size, urban/rural residency, and age and sex of family head. The amount of the grant would not fluctuate by state or region. The enactment of an NIT would establish a uniform federal assistance floor in place of the widely different benefit standards now prevailing in

the 50 states and the District of Columbia. This would mean that in many cases the amounts of grants to recipients would be different under the new system. Single individuals would be eligible for benefits for the first time, while in many states family units might incur a benefit reduction. In addition, the federal proportion of the program cost would change, affecting state budget allocations, the state aggregate income distribution, and patterns of federal support to states.

In all three of the tested NIT plans, the rate at which the payment is reduced per additional dollar of income (the benefit reduction rate, or tax) is 100% on unearned nonassistance income. The benefit reduction rate on earned income is less, in order to preserve a work incentive. The first formula, which represents an NIT which would provide moderate benefits with a medium cost and caseload, has a support level equal to 75% of the poverty line and a benefit reduction rate on earned income of 50%. The second formula, which would result in a low-cost program, provides support at 50% of the poverty line with an earned income reduction rate of 70%. Finally, a generous support plan is simulated, with benefits set at 100% of the poverty line and earned income reduced by 50%. All three formulas are applied holding harmless the families participating in the current program--that is, this study assumes that states will supplement basic grants to assure that no family will be worse off under the proposed system than under current services.

RESULTS AND IMPLICATIONS

New Federal Benefits

One criterion by which welfare reform proposals are measured is their cost to the federal government. Federal expenditures increase under two of

the simulated NIT plans but drop under the third, least generous, plan.

Under the moderate plan--with an earnings benefit reduction rate of 50% and a support level equal to 75% of the poverty line--federal expenditures will increase about 20%. The more generous plan (50% tax/100% support), with its higher breakeven and coverage, escalates federal costs by almost 170%. The increase is due to the high benefits paid to lower income units under this plan, and also to the extended coverage due to the higher breakeven point. The opposite effect occurs with the less generous plan (70% tax, 50% support). The low support and high tax rates of this program result in substantial savings to the federal government. Federal expenditures drop 53% because many units eligible for payment under the existing system are no longer eligible under this NIT, and those units which are eligible receive lower benefits.

Table 1 shows the percentage change in federal benefits by state for each of the simulated NIT plans. Under the two more generous plans, most states receive more federal dollars under the NIT than under current programs. The only exceptions are the District of Columbia, Alaska, and Hawaii, which experience slight decreases in federal spending under the plan of intermediate generosity. Under the third, least generous plan, federal spending in all states decreases. In addition to some states gaining and others losing federal dollars, the percentage change of federal funds varies substantially among states. For the two more generous plans, the biggest winners are the South Atlantic, East South Central, and Mountain states, which receive larger percentage increases in federal spending than the Northeastern and Pacific states like New York, New Jersey, and California. All the states are losers

TABLE 1

FEDERAL BENEFIT AND CASELOAD UNDER EACH REFORM PLAN BY STATE

Census Division	State	.50 Tax/.75 Support		.70 Tax/.50 Support		.50 Tax/1.00 Support	
		Percent Change in Federal Benefits	Percent Change in Caseload	Percent Change in Federal Benefits	Percent Change in Caseload	Percent Change in Federal Benefits	Percent Change in Caseload
New England	Massachusetts	19	-21	-52	-66	143	46
	Rhode Island	6	-20	-60	-67	123	52
	Connecticut	34	1	-46	-58	177	96
	Maine	16	-11	-63	-68	167	76
	New Hampshire	50	12	-47	-55	249	139
	Vermont	53	19	-43	-56	222	122
Mid-Atlantic	New York	7	-21	-54	-62	120	54
	New Jersey	8	-19	-54	-63	118	54
	Pennsylvania	28	-6	-48	-58	167	83
East Central	Ohio	24	-10	-53	-65	174	81
	Indiana	39	4	-52	-63	268	119
	Illinois	16	-13	-49	-62	121	56
	Michigan	21	-12	-49	-57	135	55
	Wisconsin	22	1	-57	-62	174	99
West North Central	Minnesota	48	18	-49	-52	223	126
	Iowa	35	15	-55	-58	211	132
	Missouri	23	-12	-56	-65	161	66
	North Dakota	60	10	-49	-60	263	110
	South Dakota	56	21	-48	-56	239	118
	Nebraska	82	10	-33	-56	313	115
	Kansas	41	4	-52	-62	223	105
South Atlantic	Delaware	34	-11	-48	-61	190	71
	Maryland	25	-4	-53	-61	168	79
	District of Columbia	-6	-15	-65	-63	93	45
	Virginia	25	2	-61	-67	174	88
	West Virginia	32	10	-58	-61	189	91
	North Carolina	28	-1	-63	-63	186	83
	South Carolina	24	-4	-66	-67	167	59
	Georgia	23	-8	-59	-61	148	60
	Florida	79	25	-38	-49	261	106
	East South Central	Kentucky	32	-2	-58	-62	175
Tennessee		25	2	-62	-66	176	80
Alabama		27	6	-60	-65	174	75
Mississippi		45	-1	-54	-56	189	54
West South Central	Arkansas	29	3	-64	-68	187	74
	Louisiana	35	5	-52	-55	165	69
	Oklahoma	33	-6	-55	-58	188	72
	Texas	54	9	-53	-59	231	87
Mountain	Montana	82	24	-41	-55	300	127
	Idaho	68	25	-45	-56	292	147
	Wyoming	98	36	-37	-57	331	150
	Colorado	54	10	-45	-55	237	121
	New Mexico	52	10	-52	-56	217	88
	Arizona	104	35	-33	-50	331	142
	Utah	45	26	-54	-58	264	156
	Nevada	145	42	-18	-50	447	158
Pacific	Washington	39	-5	-45	-57	190	76
	Oregon	44	14	-47	-52	215	124
	California	22	-15	-55	-65	144	46
	Alaska	-2	-13	-63	-60	90	40
	Hawaii	-10	-20	-63	-60	69	31

SOURCE: Calculated by Mathematica Policy Research from applications of the MATH model to the SIE data base.

under the least generous plan. Under the 50/75 program, states which receive the largest increase in federal funds, the Southern and Mountain states, are those which pay the lowest benefits to recipients under the current welfare system and which pay the smallest proportion of their total state welfare bill. Conversely, states which pay high benefits under the current welfare system, and which pay a higher proportion of their total state welfare bill, receive the smallest percentage increase in federal benefits under the NIT.

The differences in the level of funding to each state are due to a combination of changes in the average payment and in the number of families receiving a payment. Under the reform plan, units now considered categorically ineligible, principally poor households that include an able-bodied male parent, may receive payments. This increases both caseload and federal expenditure. On the other hand, recipients under current services may be ruled income ineligible under a less generous reformed plan. If the 50/75 NIT plan were to be adopted without a hold-harmless provision, the number of units receiving benefits would drop in New York, California, and the other high-benefit states, but increase in the Southern and Mountain states with a current low-benefit structure. This follows from the fact that the NIT is less generous than the combined federal-state program in the former states, and more generous than the old plan in the latter.

Table 2 shows the effects of the various reform plans on caseloads. The Mid-Atlantic states experience a modest increase in federal benefits, with a decline in caseloads, under the 50/75 plan. The Mountain states, on the other hand, experience a more dramatic increase in federal benefits and a

TABLE 2

FEDERAL BENEFIT AND CASELOAD CHANGES UNDER EACH REFORM PLAN BY CENSUS DIVISION

Division	.50 Tax/.75 Support		.70 Tax/.50 Support		.50 Tax/1.00 Support	
	Percent Change in Federal Benefits	Percent Change in Caseload	Percent Change in Federal Benefits	Percent Change in Caseload	Percent Change in Federal Benefits	Percent Change in Caseload
Northeast	24	-12	-52	-63	159	68
Mid-Atlantic	13	-16	-52	-61	133	63
East Central	23	- 8	-51	-61	153	73
West North Central	39	3	-52	-60	206	99
South Atlantic	36	4	-56	-60	186	81
East South Central	31	1	-59	-62	178	68
West Central	29	5	-55	-59	204	79
Mountain	71	22	-43	-54	276	130
Pacific	23	-12	-54	-64	150	53

SOURCE: Calculated by Mathematica Policy Research from application of the MATH model to the SIE data base.

substantial growth in caseloads. This basic pattern is repeated in the 50/100 plan, with the least changes in federal benefits occurring in the Mid-Atlantic states and the greatest in the Mountain states. Caseloads decline in the Mid-Atlantic but increase in the Mountain states. In the 70/50 plan, where reform federal benefits are lower than current service levels, the proportional decrease in federal expenditures is relatively constant over states. Caseloads also decrease due to the large number of current service units which become ineligible under the 70/50 plan.

State Fiscal Relief

A major stimulus for welfare reform is to provide relief to state budgets. The impact of any reform on a state budget is a function of the state's expenditures under the existing system and the state's fiscal liabilities after the reform. Under the present system, benefits are financed jointly with state contributions determined by a formula based on per capita income and other factors. In addition, states have the option of supplementing the basic grant. Under the proposed NIT the federal government would be responsible for the entire grant, and, barring a hold-harmless provision, the amount of state relief would equal present state expenditure for assistance. Thus, high-benefit states like California, New York, and Illinois would experience the greatest fiscal relief.

Enactment of a program with a hold-harmless provision would guarantee that the total benefit received by an NIT unit under the reform is at least as large as the sum of all benefits persons in the unit would receive under the current system, regardless of the unit's benefit calculated under the

rules of the NIT. The results of our simulations show that under a reform with a hold-harmless supplement financed by state governments, not all strata would experience fiscal relief. No relief can occur where the total hold-harmless state supplement is larger than the amount of prereform state contribution.² Under the least generous plan simulated, 49 states would face increased fiscal burden; under the intermediate plan, 44 states, and under the most generous plan, 21 states would face increased fiscal burden. The Southern and Mountain states would face increased fiscal burden under each of the simulated plans. The Eastern, Central, and Pacific states, on the other hand, would face either slightly negative or positive relief depending on the generosity of the plan. The more generous the plan, the greater the fiscal relief. High-benefit states have larger total hold-harmless supplements and thus might be expected to receive less fiscal relief than low-benefit states. Instead, the analysis shows that many states which have relatively high payment levels under the current transfer programs receive greater fiscal relief than do some low-benefit states. This puzzling result is due to the fact that the state contribution to current transfer programs in high-benefit states is more than proportionally larger than the state contribution of low-benefit states. While the hold-harmless supplement in high-benefit states is larger than in low-benefit states, the amount of the prereform state contribution replaced by federal monies in high-benefit states is so much larger than it is in low-benefit states that the net fiscal relief is greater in high-benefit states.³

Table 3 shows that in the absence of any hold-harmless provision those states with high-benefit levels--that is the Northeast and Far Western states--

TABLE 3

STATE FISCAL RELIEF UNDER EACH REFORM PLAN AND HOLD-HARMLESS OPTION
(in thousands)

Census Division	State	No Hold-Harmless (all plans)	Hold Harmless		
			.50 Tax/.75 Support	.70 Tax/.50 Support	.50 Tax/1.00 Support
New England	Massachusetts	\$ 350,466.	\$127,247.	\$ 49,080.	\$185,435.
	Rhode Island	29,229.	- 457.	-10,438.	10,647.
	Connecticut	82,492.	- 1,332.	-34,686.	41,663.
	Maine	21,249.	- 14,272.	-31,987.	680.
	New Hampshire	10,881.	- 5,679.	-13,862.	3,594.
	Vermont	12,327.	- 2,025.	- 8,057.	5,762.
Mid-Atlantic	New York	919,652.	81,656.	-189,302.	479,455.
	New Jersey	235,410.	3,010.	- 97,347.	112,052.
	Pennsylvania	317,079.	- 48,243.	-229,606.	160,219.
East Central	Ohio	199,572.	-74,405.	-193,238.	19,112.
	Indiana	45,998.	-30,263.	- 82,128.	16,201.
	Illinois	411,265.	- 4,796.	-199,868.	190,884.
	Michigan	376,388.	1,415.	-151,626.	210,321.
	Wisconsin	118,678.	-24,960.	- 62,567.	32,790.
West North Central	Minnesota	71,504.	-22,958.	- 66,990.	31,397.
	Iowa	37,637.	-25,306.	- 49,064.	2,972.
	Missouri	56,243.	-87,159.	-144,888.	-15,872.
	North Dakota	4,496.	- 9,433.	- 13,622.	- 3,254.
	South Dakota	6,317.	-10,397.	- 17,002.	- 2,977.
	Nebraska	22,402.	- 1,216.	- 15,037.	8,046.
	Kansas	28,445.	-15,128.	- 39,995.	7,870.
South Atlantic	Delaware	13,181.	- 934.	- 7,767.	4,357.
	Maryland	70,792.	- 20,078.	- 73,958.	12,954.
	District of Columbia	42,847.	- 13,347.	- 31,288.	8,299.
	Virginia	56,159.	- 86,929.	-144,243.	-26,344.
	West Virginia	12,607.	- 20,040.	- 41,192.	2,558.
	North Carolina	37,847.	- 63,905.	-171,526.	-22,125.
	South Carolina	10,268.	- 64,107.	-117,678.	-31,217.
	Georgia	36,468.	-101,370.	-205,792.	-36,330.
	Florida	48,963.	- 80,737.	-180,236.	-38,288.
East South Central	Kentucky	32,306.	-70,648.	-114,562.	-34,441.
	Tennessee	25,375.	-58,556.	-132,421.	-19,715.
	Alabama	16,758.	-46,998.	-115,432.	-14,293.
	Mississippi	6,852.	-44,992.	-103,351.	-12,678.
West South Central	Arkansas	12,070.	- 50,515.	- 78,473.	- 17,491.
	Louisiana	29,506.	- 62,023.	-131,930.	- 18,768.
	Oklahoma	73,804.	831.	- 38,064.	37,577.
	Texas	39,960.	-193,058.	-373,549.	-116,398.
Mountain	Montana	4,461.	- 7,356.	-14,831.	-1,841.
	Idaho	10,728.	- 6,535.	-13,689.	921.
	Wyoming	1,749.	- 3,881.	- 6,776.	-1,764.
	Colorado	51,019.	- 9,275.	-38,496.	19,043.
	New Mexico	8,638.	-17,497.	-39,630.	-4,528.
	Arizona	14,090.	-19,902.	-39,182.	-3,869.
	Utah	10,358.	-11,653.	-27,308.	1,050.
	Nevada	8,403.	131.	- 5,930.	4,259.
Pacific	Washington	104,751.	- 11,238.	- 44,938.	39,071.
	Oregon	53,848.	- 21,870.	- 44,167.	13,307.
	California	1,612,828.	480,337.	307,940.	806,049.
	Alaska	11,868.	- 6,973.	- 9,055.	-3,119.
	Hawaii	36,847.	- 18,939.	- 27,493.	-5,255.
	U.S. Total	5,853,069.	-796,756.	-3,667,203.	2,037,977.

SOURCE: Calculated by Mathematica Policy Research from application of the MATH model to the SIE data base.

would receive the greatest relief, although this would be achieved at the expense of providing lower benefits for many recipients under any of the three NIT plans simulated. This follows from the fact that the high-benefit states are those which currently supplement federal benefits substantially.

State fiscal relief with the hold-harmless in effect is also shown in Table 3. The large number of states which experience increased expenditures rather than relief under an NIT when a hold-harmless provision is in effect suggests that, for many of those categorically eligible under current programs, current transfer payments are relatively generous. Table 3 shows that the Southern and Mountain states have large increases in expenditures while the high-benefit, more populated states receive relatively more relief.

Under the 70/50 plan, most states are subject to increased expenditures relative to the current system and only two, Massachusetts and California, continue to receive positive state fiscal relief. Under this formula state expenditures are increased and the Southern states continue to lose more than the high-benefit jurisdictions.

Under the 50/100 plan many states show a reduction in welfare expenditures, even with the hold-harmless in effect. Those states which finance the lowest proportion of current services, like those in the South, still face increased expenditures since additional outlays to replace lost federal benefits for upper income units outweigh reimbursements for state expenditures which went to lower income units under current services.

Redistribution of Personal Income

Another important aspect of a welfare reform is its impact on the disposable income of state residents. Net income is affected by both the benefit

level set by a new plan, and also by the manner in which the federal government finances the reform. Changes in the regional income redistribution brought about by a reform program is one important consideration for evaluating alternative proposals. We modeled two methods of financing additional federal revenues. In one (Plan 1) we applied a constant percentage surcharge to all federal personal tax liabilities, and in the other (Plan 2), a surcharge on both personal and corporate federal income taxes, with the additional tax burden distributed between the two sources in proportion to their share of total federal revenues. Table 4 shows the calculated surcharge rates under each financing scheme for the three reform plans. Differences in requirements are striking among the three benefit levels simulated. Under Plan 1, an NIT with a 50/75 benefit formula would require a surcharge of 2.89%. Under the less generous plan (70/50) federal expenditures drop below current service levels, and residents would receive a tax credit of 5.33%. The 50/100 plan increases federal expenditures so that a surcharge of 17.00% would be needed to finance the additional revenue requirement.

It was assumed that changes in state budget expenditures (the extent of state fiscal relief) would be balanced dollar for dollar by personal state tax liabilities for residents. Since state budget expenditures reflect benefit levels, a one dollar increase in state financed benefits for one state resident would be counteracted by a one dollar tax increase for another. Thus, state budget variations would have no affect on aggregate state disposable income. The extent to which personal disposable income is redistributed varies depending on benefit and tax structures in each state. Tables 5, 6, and 7 show the net gain or loss by state under three NIT formulas.

TABLE 4

FEDERAL TAX SURCHARGE RATES UNDER ALTERNATIVE FINANCING
SCHEMES BY REFORM PLAN

Financing Scheme	Reform Plan		
	.50 Tax/.75 Support	.70 Tax/.50 Support	.50 Tax/1.00 Support
Plan 1:			
Surcharge rate on positive federal individual federal income tax liabilities	.0289	-.0533	.1700
Plan 2:			
Surcharge rate on positive federal individual federal income tax liabilities	.0217	-.0401	.1277
Surcharge rate on federal corporate income tax liabilities	.0162	-.0299	.0952

SOURCE: Calculated by Mathematica Policy Research from application of the MATH model to the SIE data base.

TABLE 5

CHANGES IN AGGREGATE STATE DISPOSABLE INCOME

.50 Tax/.75 Support, Scheme 1 Financing

(in thousands of dollars)

Census Division	State	New Federal Benefits	New Federal Taxes	Net Transfer
New England	Massachusetts	63,870.	107,134.	- 43,264.
	Rhode Island	3,521.	14,971.	- 11,451.
	Connecticut	43,455.	70,528.	- 27,073.
	Maine	12,366.	13,921.	- 1,555.
	New Hampshire	15,901.	12,918.	2,983.
	Vermont	16,607.	6,112.	10,495.
Mid Atlantic	New York	94,003.	367,891.	-273,889.
	New Jersey	33,744.	156,731.	-122,988.
	Pennsylvania	197,536.	193,467.	4,069.
East Central	Ohio	135,382.	183,142.	- 47,860.
	Indiana	96,513.	94,475.	2,038.
	Illinois	124,974.	236,182.	-111,208.
	Michigan	123,742.	164,990.	- 41,218.
	Wisconsin	45,322.	78,199.	- 32,977.
West North Central	Minnesota	81,562.	67,734.	13,827.
	Iowa	41,759.	50,148.	- 8,389.
	Missouri	71,741.	78,873.	- 7,132.
	North Dakota	15,234.	10,333.	4,901.
	South Dakota	19,575.	8,510.	11,065.
	Nebraska	44,621.	27,452.	17,169.
	Kansas	38,679.	40,112.	- 1,415.
South Atlantic	Delaware	8,985.	11,697.	- 2,712.
	Maryland	49,403.	99,768.	- 50,366.
	District of Columbia	-4,812.	19,944.	- 24,756.
	Virginia	67,851.	96,763.	- 28,912.
	West Virginia	41,576.	22,835.	18,741.
	North Carolina	105,096.	77,736.	27,359.
	South Carolina	52,382.	34,989.	17,393.
	Georgia	107,329.	72,264.	35,065.
	Florida	369,478.	135,831.	233,647.
East South Central	Kentucky	88,216.	42,831.	45,385.
	Tennessee	81,713.	57,210.	24,503.
	Alabama	84,149.	47,374.	36,775.
	Mississippi	112,381.	27,473.	84,908.
West South Central	Arkansas	50,985.	24,830.	26,155.
	Louisiana	127,161.	52,101.	75,060.
	Oklahoma	59,036.	46,124.	12,913.
	Texas	408,269.	214,563.	193,706.
Mountain	Montana	24,398.	12,165.	12,233.
	Idaho	23,611.	12,306.	11,305.
	Wyoming	11,862.	7,092.	4,770.
	Colorado	59,418.	52,917.	6,500.
	New Mexico	47,432.	17,668.	29,764.
	Arizona	103,072.	39,470.	63,601.
	Utah	24,432.	19,061.	5,371.
	Nevada	23,981.	13,629.	10,351.
Pacific	Washington	65,587.	71,462.	-5,874.
	Oregon	50,414.	40,543.	9,871.
	California	312,651.	481,719.	-169,068.
	Alaska	- 321.	14,545.	- 14,866.
	Hawaii	- 6,511.	18,470.	- 24,981.
	U.S. Total	3,869,143.	3,869,170.	- 25.

SOURCE: Calculated by Mathematica Policy Research from application of the MATH model to the SIE data base.

TABLE 6

CHANGES IN AGGREGATE STATE DISPOSABLE INCOME

.70 Tax/.50 Support, Scheme 1 Financing
(in thousands)

Census Division	State	New Federal Benefit	New Federal Taxes	Net Transfer
New England	Massachusetts	- 171,107.	- 197,457.	26,349.
	Rhode Island	- 34,322.	- 27,593.	- 6,729.
	Connecticut	- 59,439.	- 129,988.	70,548.
	Maine	- 48,095.	- 25,657.	- 22,439.
	New Hampshire	- 14,951.	- 23,609.	8,858.
	Vermont	- 13,407.	- 11,265.	- 2,142.
Mid-Atlantic	New York	- 732,097.	- 678,051.	- 54,046.
	New Jersey	- 229,256.	- 288,867.	59,611.
	Pennsylvania	- 343,546.	- 356,574.	13,029.
East Central	Ohio	- 299,378.	- 337,545.	38,167.
	Indiana	- 93,257.	- 174,124.	80,867.
	Illinois	- 387,900.	- 435,301.	47,401.
	Michigan	- 280,995.	- 304,034.	23,029.
	Wisconsin	- 116,441.	- 144,127.	27,686.
West North Central	Minnesota	- 82,184.	- 124,839.	42,655.
	Iowa	- 66,051.	- 92,426.	26,375.
	Missouri	- 175,087.	- 145,369.	- 29,718.
	North Dakota	- 12,544.	- 19,044.	6,500.
	South Dakota	- 16,662.	- 15,684.	- 978.
	Nebraska	- 17,989.	- 50,596.	32,607.
	Kansas	- 48,525.	- 73,929.	25,405.
South Atlantic	Delaware	- 12,548.	- 21,599.	9,011.
	Maryland	- 103,934.	- 183,881.	79,946.
	District of Columbia	- 53,676.	- 36,758.	- 16,918.
	Virginia	- 162,514.	- 178,342.	15,828.
	West Virginia	- 75,334.	- 42,087.	- 33,247.
	North Carolina	- 237,697.	- 143,274.	- 94,423.
	South Carolina	- 142,606.	- 64,487.	- 78,119.
	Georgia	- 271,595.	- 133,188.	- 138,406.
	Florida	- 176,981.	- 250,347.	73,366.
East South Central	Kentucky	- 162,519.	- 78,942.	- 83,577.
	Tennessee	- 201,841.	- 105,442.	- 96,399.
	Alabama	- 183,483.	- 87,314.	- 96,169.
	Mississippi	- 134,087.	- 50,635.	- 83,452.
West South Central	Arkansas	- 111,844.	- 45,764.	- 66,080.
	Louisiana	- 188,879.	- 96,026.	- 92,853.
	Oklahoma	- 99,682.	- 85,009.	- 14,673.
	Texas	- 400,830.	- 395,455.	- 5,375.
Mountain	Montana	- 12,249.	- 22,421.	10,172.
	Idaho	- 15,596.	- 22,681.	7,086.
	Wyoming	- 4,443.	- 13,071.	8,628.
	Colorado	- 48,999.	- 97,531.	48,532.
	New Mexico	- 46,312.	- 32,563.	- 13,749.
	Arizona	- 32,543.	- 72,747.	40,204.
	Utah	- 28,832.	- 35,130.	6,298.
	Nevada	- 3,063.	- 25,120.	22,057.
Pacific	Washington	- 76,605.	- 131,709.	55,104.
	Oregon	- 54,547.	- 74,723.	20,176.
	California	- 791,165.	- 887,844.	96,680.
	Alaska	- 12,041.	- 26,807.	14,766.
	Hawaii	- 41,491.	- 34,041.	- 7,450.
	Total	-7,131,120.	-7,131,111.	5.

SOURCE: Calculated by Mathematica Policy Research from application of the MATH model to the SIE date base

TABLE 7

CHANGES IN AGGREGATE STATE DISPOSABLE INCOME
.50 Tax/1.00 Support, Scheme 1 Financing
(in thousands)

Census Division	State	New Federal Benefits	New Federal Taxes	Net Transfer
New England	Massachusetts	473,816.	629,242.	-155,426.
	Rhode Island	69,661.	87,933.	- 18,272.
	Connecticut	229,825.	414,236.	-184,411.
	Maine	126,778.	81,761.	45,018.
	New Hampshire	79,873.	75,873.	3,810.
	Vermont	69,404.	35,897.	33,506.
Mid-Atlantic	New York	1,630,865.	2,160,771.	-529,906.
	New Jersey	499,976.	920,543.	-420,566.
	Pennsylvania	1,180,511.	1,136,308.	44,204.
East Central	Ohio	982,529.	1,075,667.	- 93,139.
	Indiana	477,381.	554,888.	- 77,507.
	Illinois	964,814.	1,387,190.	-422,376.
	Michigan	781,791.	968,875.	-187,084.
	Wisconsin	352,909.	459,295.	-106,387.
West North Central	Minnesota	376,950.	397,829.	- 20,879.
	Iowa	253,201.	294,537.	- 41,336.
	Missouri	496,591.	463,252.	33,339.
	North Dakota	66,945.	60,688.	6,257.
	South Dakota	83,050.	49,982.	33,068.
	Nebraska	169,890.	161,236.	8,654.
	Kansas	209,352.	235,593.	- 26,241.
South Atlantic	Delaware	49,458.	68,702.	- 19,244.
	Maryland	328,019.	585,980.	-257,960.
	District of Columbia	76,922.	117,137.	- 40,215.
	Virginia	464,699.	568,330.	-103,630.
	West Virginia	246,506.	134,120.	112,388.
	North Carolina	703,703.	456,576.	247,127.
	South Carolina	362,478.	205,504.	156,974.
	Georgia	684,793.	424,436.	260,357.
	Florida	1,216,619.	797,789.	418,830.
East South Central	Kentucky	488,386.	251,566.	236,820.
	Tennessee	578,328.	336,017.	242,312.
	Alabama	535,915.	278,248.	257,667.
	Mississippi	471,710.	161,359.	310,351.
West South Central	Arkansas	326,949.	145,839.	181,110.
	Louisiana	600,190.	306,009.	294,182.
	Oklahoma	338,635.	270,902.	67,733.
	Texas	1,734,670.	1,260,210.	474,459.
Mountain	Montana	89,072.	71,451.	17,621.
	Idaho	101,004.	72,279.	28,725.
	Wyoming	40,051.	41,655.	- 1,603.
	Colorado	260,806.	310,804.	- 49,998.
	New Mexico	194,735.	103,769.	90,966.
	Arizona	328,363.	231,825.	96,538.
	Utah	142,000.	111,951.	30,049.
	Nevada	74,144.	80,050.	- 5,906.
Pacific	Washington	321,608.	419,722.	- 98,114.
	Oregon	247,941.	238,124.	9,817.
	California	2,079,251.	2,829,326.	-750,057.
	Alaska	17,051.	85,427.	- 68,375.
	Hawaii	45,157.	108,479.	- 63,322.
	U.S. Total	22,725,001.	22,725,100.	- 96.

SOURCE: Calculated by Mathematica Policy Research from application of the MATH model of the SIE data base.

The pattern of gainer and loser states in terms of disposable income is different from that of federal expenditures and of fiscal relief, and depends on the generosity of the plan. High-payment states like California, New York, and Massachusetts pay more taxes to support the additional federal expenditure than they receive under the two more generous plans. Southern and Mountain states receive more federal money than their residents pay in taxes for those two plans. Some of this pattern is reversed for the least generous plan, with the Southern states losing more than they gain and some high-payment states becoming net gainers.

CONCLUSION

Our simulations show that the amount of federal expenditures for transfer programs increases when AFDC, food stamps, and SSI are replaced with all but the least generous NIT simulated. The additional federal expenditures are concentrated in those states with low AFDC payment standards, the states in the Southern and Mountain regions. This concentration occurs even though the proportion of current expenditures in low-payment states that is paid by the federal government is relatively high.

If the NIT is not supplemented by the states, the state fiscal relief equals the current state expenditures. Fiscal relief is thus greatest in Northeastern and North Central states and California. If, however, states are required to supplement the NIT payments to hold harmless those families currently participating in welfare programs, the NIT does not result in fiscal relief for all states. The most generous tested NIT left 21 states with an increased fiscal burden, rather than relief, and the least generous plan left 48 states and the District of Columbia without relief.

If the additional federal expenditures are financed by an income tax surcharge, either on personal income or on both personal and corporate income, the states which are net gainers of transfer dollars are the Southern and Mountain states. The remaining states are net losers of transfer dollars. The Northeastern and Pacific states receive the most fiscal relief but are net losers of transfer dollars while the Southern and Mountain states are net gainers of transfer dollars but incur an additional fiscal burden from the welfare reform.

NOTES

1. See Harold Beebout, Microsimulation as a Policy Tool: The MATH Model (Policy Analysis Series No. 14, Washington, D.C., Mathematica Policy Research, 1977). This paper is summarized by Beebout and Mathematica Policy Research.
2. For a more technical explanation of how increased fiscal burdens come about, see Appendix B of Maxfield and Edson, Welfare Reform and State Fiscal Tools (MPR Inc., 1978).
3. See Appendix B for a more detailed explanation.

DISCUSSION

JODIE ALLEN: Under any of the NIT plans tested, some states will supplement the payment, even when income is zero. Am I right?

MAXFIELD: You are correct, there are many variations possible for figures 1 and 2.

ALLEN: But you always find some hold harmless liability in each of the three proposed programs?

MAXFIELD: Yes.

PHILIP ROBINS: You are assuming a holds harmless provision only for those people receiving benefits at the time the program replaced existing programs?

MAXFIELD: That's right. A family who did not receive welfare before the reform and was below the Federal NIT break-even point was not given a supplement.

JOHN McCOY: The states can't do that: there would be no way California could have that kind of plan. We could not have a different standard for the same type of people.

MAXFIELD: The rationale is not equity across people in the same time, but to prevent making anybody worse off.

GLADYS McCORKHILL: There is a parallel in the SSI program in the optional supplement. In fact, state legislators felt that they could not give any less of a supplement to people new to the program. It seems to me that states are going to have a different attitude about equity than your model.

MICHAEL LINN: Assuming the states would follow your proposal, which is supplementation only to hold present recipients harmless, the need for supplementation would dissipate over time; is that correct?

MAXFIELD: That's an important and unresolved question. For example, if you are trying to hold harmless a family, where the head or heads were out of work, and then return to work later, should the family then go on the new schedule or be held harmless according to the old schedule? I don't know. If you do the latter, you have to keep two welfare systems operating together.

McCOY: I think the issue really centers on political reality. Many active groups have participated in bringing up these schedules in high-benefit states. They're not going to disappear with the onset of a new plan. In addition, you are also bringing in a new group of eligibles to add to their clientele. As Ms. McCorkhill observed with respect to SSI, the pressure is real. Our SSI program in California now costs us more than the prior program which involved a 50% federal sharing.

ROBINS: Turning to another subject, do you maintain participation rates in food stamps and AFDC in a manner consistent with the aggregate special rates in each state?

MAXFIELD: No, but in a manner that is consistent with each region.

ROBINS: Participation rates could influence the results. That is, in states with generous welfare payments there could be more eligible people drawing benefit from the NIT. The lower participation rates in the AFDC and food stamp program might make these states better off.

ALLEN: That's what has happened. In the South, NIT benefits are going to poor people not previously on welfare, but that doesn't transfer into "fiscal relief" because the states weren't paying these people anyway.

LINN: Do I read figures 1 and 2 correctly, when I observe that the guarantee at zero income is somewhat higher than that presently provided by the states?

MAXFIELD: I think that's the case, yes.

LIN: That's good: the program seems to be more targeted on the very poor.

MAXFIELD: Yes. The AFDC rules impose a 67% tax rate, but with all the deductions and other provisions in the payment formula, it has an effective rate of only 25 or 30%.

McCOY: Back to your point about case loads. In addition to possible increased costs that we have already discussed, there is also the working poor. They are not currently served by our categorical assistance programs, but would be under the NIT, and your estimate of the costs for that group would be different than our estimate, because you are only paying them the base.

MAXFIELD: That's right.

FIGURE 1

PRE- AND POSTREFORM BUDGET CONSTRAINTS--HIGH-BENEFIT STATE

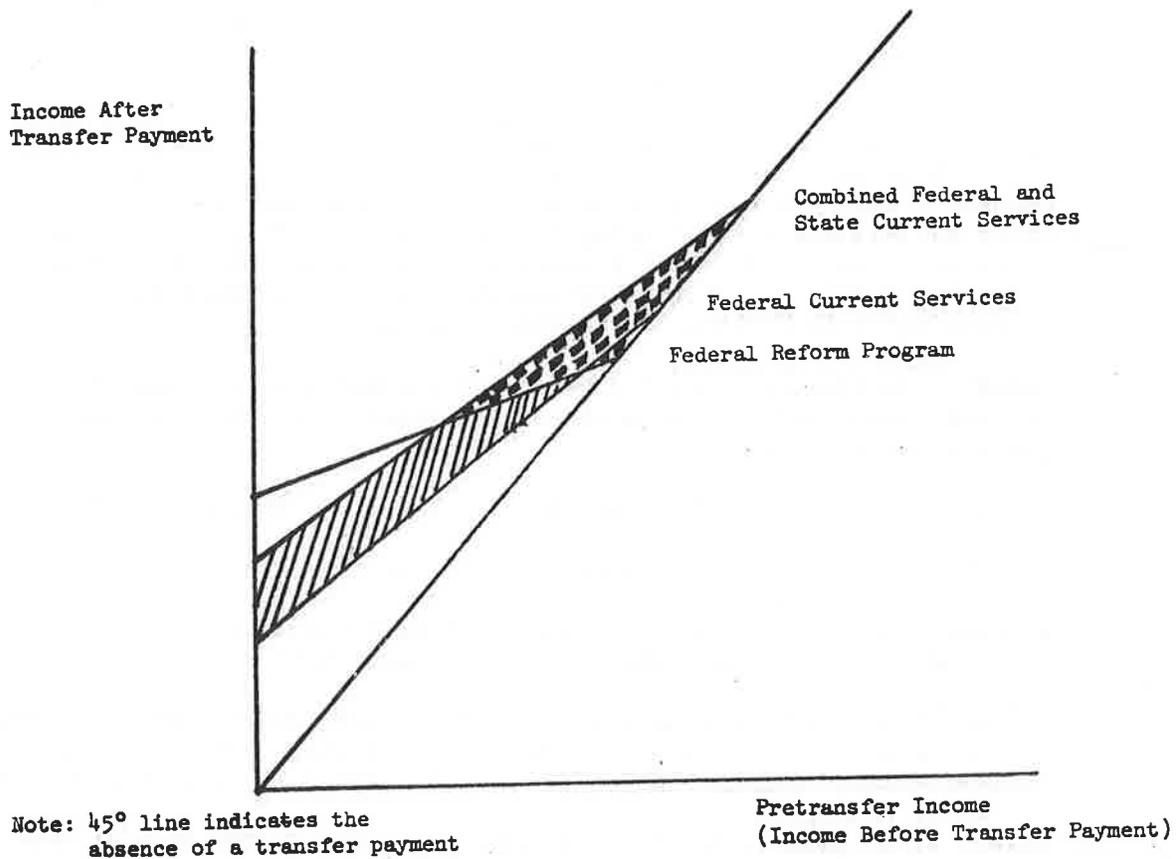
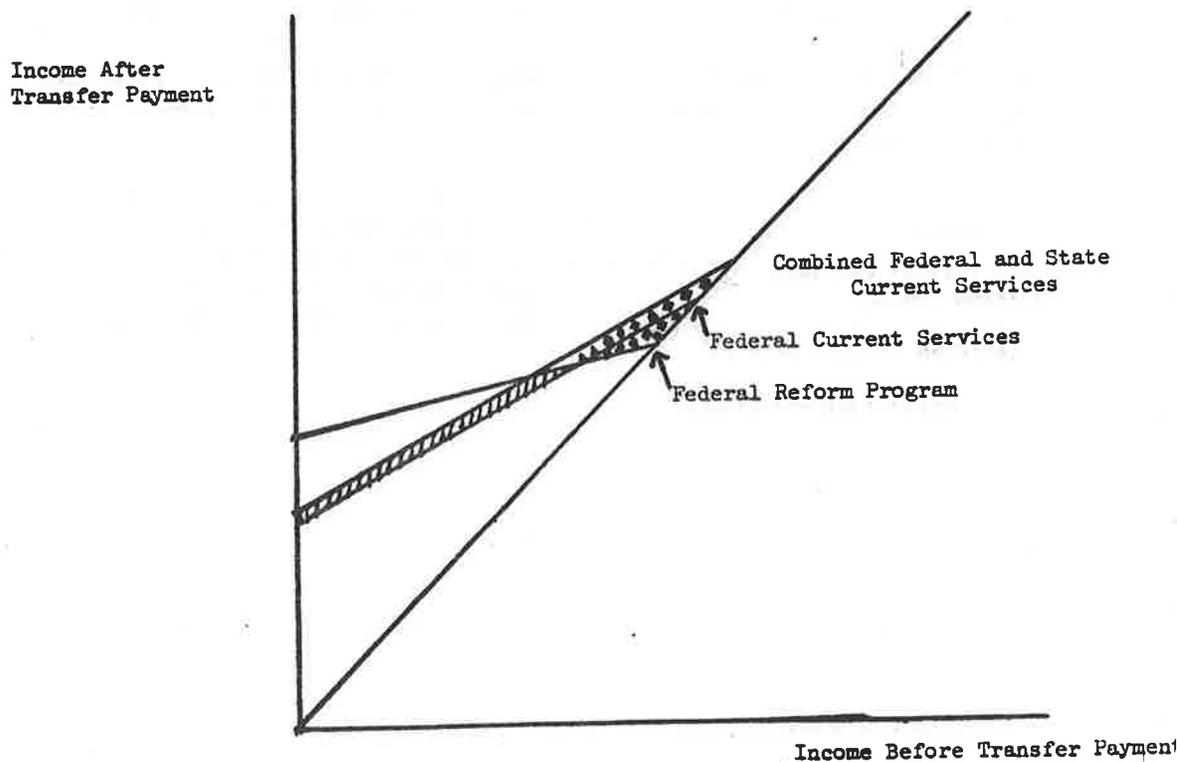


FIGURE 2

PRE- AND POSTREFORM BUDGET CONSTRAINTS--LOW-BENEFIT STATE



APPENDIX A

MONTHLY INCOME REPORT FORM

(For discussion, see the paper by Gary Christopherson and Cheri Marshall, "Data Collection," above.)

COUNCIL FOR GRANTS TO FAMILIES
INCOME AND CHANGE REPORT

Filing Status

For the Period

BEGINNING DATE

ENDING DATE

Period Number

Name(s) of Head(s) of Family

Family Number

Street Address		Telephone Number
City	State	Zip Code

Fill in your correct address.

Has your address changed in the last month? Yes No

PLEASE RETURN THIS FORM IMMEDIATELY

Forms returned after _____ will cause your check to be delayed.

If you have any questions about filling out this form, please feel free to call or visit the Field Office:

CGF FIELD OFFICE
107 Cherry Street
4th Floor, Lowman Building
Seattle, Washington 98104

622-5198

HOURS: Monday through Friday
9 a.m. - 5 p.m.

PLEASE SIGN AND DATE THIS FORM BEFORE RETURNING IT.

DATE _____

Male Head _____

Female Head _____

A.I.



PART I: EMPLOYMENT STATUS AND CHANGES IN FAMILY SIZE

List by name all family members who are 16 years of age or older. Check the box which describes their current job status.

Fill in the NAME of each family member 16 or over include both heads of household.	Has job status changed this period?		CHECK ONE BOX FOR EACH PERSON		
	YES	NO	Unemployed	Working Part-Time	Working Full-Time
Male Head Name _____					
Female Head Name _____					
Other Member Name _____					
Other Member Name _____					
Other Member Name _____					

Has anyone come to live with your family or left your family during this reporting period? (Include births and deaths)

CHECK BOX YES NO

If the answer is yes, please fill out all the following information about them.

NAME	Relationship
Date joined	Date Left
Birth Date	How long do you expect him to stay or be away?
Reason for coming to live with or leaving family	
New address, if any, for departing members	

NAME	Relationship
Date joined	Date Left
Birth Date	How long do you expect him to stay or be away?
Reason for coming to live with or leaving family	
New address, if any, for departing members	

PART 2: GROSS EARNINGS AND WITHHOLDINGS FOR ALL JOBS

Fill in the NAME of each family member who has held a job during the reporting period in the spaces provided below. Enter all wages, tips, salaries, and commissions from all jobs in Gross Earnings Column. Gross earnings is your pay before any deductions are made. Enter Social Security (FICA) and federal income taxes withheld in Withholding Column. Enter the number of hours worked for each paycheck listed in Hours Worked Column. If you received more than one paycheck in a week, add them together.

NOTE: YOU MUST SEND IN ALL PAY STUBS OR PAY ENVELOPES WITH THIS FORM (THESE WILL BE RETURNED TO YOU)

The dates paychecks are received must fall on or between the BEGINNING and ENDING DATES listed on page 1 of this form.

NAME OF MALE HEAD			
Date Paycheck Received	Gross Earnings (Before Taxes)	State Tax, Federal Withholding Tax plus FICA	Hours Worked
	\$	\$	hrs

NAME OF FEMALE HEAD			
Date Paycheck Received	Gross Earnings (Before Taxes)	State Tax, Federal Withholding Tax plus FICA	Hours Worked
	\$	\$	hrs

NAME OF OTHER MEMBER			
Date Paycheck Received	Gross Earnings (Before Taxes)	State Tax, Federal Withholding Tax plus FICA	Hours Worked
	\$	\$	hrs

NAME OF OTHER MEMBER			
Date Paycheck Received	Gross Earnings (Before Taxes)	State Tax, Federal Withholding Tax plus FICA	Hours Worked
	\$	\$	hrs

Did you send all Pay Stubs? CHECK BOX: YES NO

If you marked NO, please explain _____.

NOTE: GROSS EARNINGS ARE ADDED TO BASIC INCOME

PART 3: OTHER INCOME RECEIVED DURING THE LAST MONTH

Check whether or not you received income from each of the sources listed. If you check YES, enter the amount in the box provided.

DO NOT INCLUDE business or self-employment income or income from the sale of capital assets here. Separate pages are provided for this purpose.

		CHECK BOX		AMOUNT
		YES	NO	RECEIVED
				\$
1. Welfare (AFDC or ADC)				\$
Items which reduce SUPPORT by their full amount	2. Other Welfare (Old Age, Disability or General Assistance)			\$
	3. Unemployment Insurance (total for month)			\$
	4. Veteran's Survivors or Disability Benefits			\$
	5. Workman's Compensation			\$
	6. Training stipends (INCLUDE scholarships, fellowships, G.I. Bill SUBTRACT tuition, books and fees)			\$
	7. Social Security Benefits (Old Age, Disability or Survivors Benefits)			\$
	Items which are added to BASIC INCOME	8. Life Insurance Benefits		
9. Other Insurance Settlements (DO NOT INCLUDE amounts received to cover expenses.)				\$
10. Pensions or Annuities (INCLUDE Veteran's Retirement Pensions)				\$
11. Payments from Private Disability Plans				\$
Items which reduce SUPPORT by half their amount	12. Alimony and/or Child Support or Other Support - in cash, merchandise, groceries, etc.			\$
An assigned value is added to BASIC INCOME	14. Rent from roomers or boarders if under \$150 (If more than one unit, fill in as business income.)			\$
	15. Rent from income property of one family unit (If more than one unit, fill in as business income)			\$
Miscellaneous	17. Other Income--from: prizes, inheritances; earnings of children under 16; odd jobs, etc. PLEASE DESCRIBE			\$
Item which reduces SUPPORT by full amount	20. IF YOU BOUGHT FOOD STAMPS DURING THIS PERIOD, PLEASE LIST: How much did they cost How much are they worth			\$
				\$

PUBLIC HOUSING:

If you live in Public Housing, enter monthly rent
Has there been an increase in your rent from last month?

YES

PART 4: EXPENSES OF THE LAST MONTH

DO NOT INCLUDE here business or self-employment expenses or losses from the sale of capital assets. Separate pages are provided for this purpose.

Medical and day care expenditures must be accompanied by receipts.

		AMOUNT SPENT
Items which are deducted from BASIC INCOME	1. Child Care - because of a job DO NOT INCLUDE expenses reimbursed by another agency	\$
	If you paid for child care, how many children were cared for: _____	XXXXXXXXXX
	2. Care for Aged or Invalid Member of Family because of a job	\$
	If you paid for care of an aged or invalid family member, how many were cared for? _____	XXXXXXXXXX
	3. Cost of Medical Expenses - NOT covered by insurance - Please include receipts	\$

Items which increase SUPPORT by half their amount	5. Alimony and/or Child Support or Other Support Paid to a Relative	\$
	How many persons were supported? _____	XXXXXXXXXX
	For each person supported, fill out the following information:	XXXXXXXXXX
	Name _____ Name _____ Address _____ Address _____	
Relationship _____ Relationship _____		

PART 5: BUSINESS OR SELF-EMPLOYMENT INCOME AND THE SALE OF PROPERTY OR OTHER CAPITAL ASSETS

1. Is anyone in the family currently self-employed or receiving income from a business?

CHECK BOX YES NO

2. Has anyone in this family sold any property, buildings, equipment, stocks or bonds during this period?

CHECK BOX YES NO

If the answer is YES, please fill out page 7 of this form.

NOTE: If you answered the last two questions NO, please disregard pages 6 and 7.

ADJ

CALC.
WITH.

411	415	413

SALE OF CAPITAL ASSETS

List below any income you have received during the reporting period from the sale of property, buildings, equipment, stocks or bonds or any other capital assets.

List each asset sold separately; enter the cost at the time of purchase and the date of purchase. Then enter the date of sale and the selling price.

Asset Description	Date of Purchase	Cost of Purchase Incl. Commissions	Date of Sale	Selling Price
		\$		\$
		\$		\$
		\$		\$
		\$		\$

Have you paid any federal taxes on capital gains from these sales of assets?

YES

NO

If YES, how much tax have you paid? \$ _____

If there has been any gain from your transaction, how has it been used:

YES NO 1. entered into a Savings Account

YES NO 2. Re-invested

YES NO 3. General expenses

4. Other

(The last page was reserved for questions and comments--Ed.)

APPENDIX B

Authors and Editors

BIOGRAPHICAL NOTES

Miriam Aiken is a researcher at Mathematica Policy Research. She has authored or co-authored several papers on welfare and education while at MPR.

Jodie Allen is Special Assistant to the Secretary of Labor for Welfare Reform. She was previously Senior Vice President of Mathematica Policy Research, and founder and director of the firm's Washington office. She is currently responsible for development and management of DOL's program of applied research particularly in the areas of income distribution and employment. In this and in her previous work, both in and out of government, Ms. Allen has been an intensive user of governmental statistics employed in the measurement of labor market conditions and has been active in efforts to improve both their quality and their usefulness to governmental and non-governmental users. She is the author of numerous popular articles and technical studies on public policy questions.

Marcy Elkind Avrin (Ph.D., Economics, Stanford University) is currently a visiting assistant professor at Stanford University. She has been an economist with SRI International, and is the author of several papers in the housing area.

Bonnie Bayes (J.D. Candidate, University of Puget Sound) was an assistant editor of this book, and aided in the difficult task of reducing the transcript and papers into a more manageable package.

Harold Beebout (Ph.D., Economics, University of Wisconsin) has played a key role in completing two important microsimulation models that have been used widely as policy analysis tools--the Transfer Income Model (TRIM) and the Micro Analysis of Transfers to Households Model (MATH). He is currently Vice President at Mathematica Policy Research, and Director of the Policy Studies Division.

Joseph Bell (Ph.D., Education, Washington State University) is the director of the income maintenance experiment for the State of Washington. As editor of this book, he was responsible for the commissioning and selection of the papers, as well as general oversight of the editing process. He has also contributed chapters on welfare policy to other books.

David M. Betson (M.A., Economics, University of Wisconsin) is presently affiliated with the Department of Health, Education and Welfare, Office of Income Security Policy/Research. He is a co-author of Earning Capacity, Poverty and Inequality (1977) and has written several articles on the application of microsimulation to welfare and jobs programs.

Matthew Black (Ph.D., Economics and Social Work, University of Michigan) is with Mathematica Policy Research. He has published papers at MPR and presented some at various conferences on subjects such as income uncertainty, changes in job satisfaction, and other job-related behavior.

Gary Christopherson (M.A., Economics, University of Washington) is currently a consultant to Mathematica Policy Research. He was the project director and chief administrator of Seattle operations in SIME.

Constance F. Citro (Ph.D., Political Science, Yale University) is presently a deputy director, Policy Studies Division, Mathematic Policy Research, Washington, D.C. She has authored numerous papers on various demographic problems.

Virgil Davis (B.S., Engineering Science, Washington University) is a data processing manager, Center for the Study of Welfare Policy, and has worked extensively in the later stages of SIME/DIME data processing.

Barbara Devaney (Ph.D., Economics, University of Michigan) is with Mathematica Policy Research, and is currently principal investigator for a study of the determinants of adolescent pregnancy and childbearing, and a co-principal investigator for a project on variations in U.S. Fertility trends. Dr. Deveney is also an instructor of labor economics and econometrics at Johns Hopkins University.

Linda Drazga (Ph.D., Sociology, Stanford University) is a sociologist at Mathematica Policy Research. She is currently developing evaluation research strategies and methodological techniques for selected government nutrition programs.

David Edson (M.A., Economics, University of Washington) is a research associate with the Policy Studies Division of Mathematica Policy Research in Washington, D.C. He has worked on several SIME/DIME research tasks and prepared a number of papers on this work.

Henry Felder (Ph.D., Economics, Stanford University) is a senior economist at SRI International in Washington, D.C. He has prepared a number of SRI reports for SIME/DIME and other agencies and has contributed a chapter to a book on welfare reform.

Charles Froland (Ph.D. Candidate, Public Health, University of California at Berkeley) is a consultant to SIME. He has authored numerous articles on income maintenance, problems related to drug and alcohol abuse and similar matters appearing in a variety of media, including the Christian Science Monitor and the International Journal of Evaluation and Planning. As an assistant to the editorial process in preparing this book, he participated heavily in the selection of authors and topics, and the organization of the conference.

Margaret Ann Grady (Ph.D., Economics, University of North Carolina) has been with Mathematica Policy Research where she prepared several papers on the effects of the income maintenance experiments on women and marriage.

David H. Greenberg (Ph.D., Economics, Massachusetts Institute of Technology) is presently affiliated with the Department of Health, Education and Welfare, Office of Income Security Policy/Research. He has authored and co-authored numerous articles on a wide variety of topics relating generally to welfare, job training and development, and poverty. They appear in journals such as the Journal of Human Resources, Policy Analysis, the Review of Economics and Statistics, Industrial and Labor Relations Review, and in several books.

Lyle P. Groeneveld (Ph.D., Sociology, Indiana University) is currently at SRI International. He has co-authored numerous papers in the SIME/DIME study, several of which appear in the American Journal of Sociology.

Arden Hall (Ph.D., Economics, University of California at Berkeley; M.A., Statistics, University of California at Berkeley) is a senior economist at SRI International. He has contributed articles for publication in Child Care and Public Policy and has prepared a number of SRI reports. He also presented a paper for the Annual Meeting of the Econometric Society in Toronto, Canada.

John Hall (J.D., University of Denver), is currently a senior survey researcher at Mathematica Policy Research, and served as project manager of the SIME/DIME School Performance Study. His research includes the use of administrative records as sources of research data, the impact of administrative change on the participants of the income maintenance experiments, and the legal problems associated with social science research.

Harlan Halsey (Ph.D., Economics, Stanford University) is an economist at SRI International, where he has prepared numerous papers relating to the income maintenance experiments.

Michael T. Hannan (Ph.D., Sociology, University of North Carolina) is an associate professor of sociology, Stanford University. He is an author or co-author of several papers on income maintenance issues as well as several books, including, most recently, Studies on Organization and Environment (1978); National Development and the World System: Educational, Economic and Political Change (1979).

Alan M. Hershey (M.P.A., Princeton) is currently associate director of Mathematica Policy Research in Denver, and director of the Vermont Integrated Reporting and Eligibility System project. He was the principal designer of the Colorado Monthly Reporting Experiment and Pretest, and Project Manager for three years during the design, implementation and analysis phases. He is also the co-author of One Nation, So Many Governments (1977).

Terry R. Johnson (Ph.D., Economics, University of Washington) is a senior economist at SRI International. In addition to his work on SIME/DIME he has written several major works appearing in, e.g., the Review of Economic Studies and Economic Inquiry.

Richard A. Kasten (Ph.D., Economics, Massachusetts Institute of Technology) is at the Department of Health, Education and Welfare, Office of Income Security Policy/Research. He is responsible for several articles on micro-simulation as an analytical tool.

Michael C. Keeley (Ph.D., Economics, University of Chicago) is a senior economist at SRI International. At SRI he has developed numerous papers and memoranda on effects of income maintenance on labor supply, migration, and fertility. He is author of The Economics of Labor Supply - A Critical Review, forthcoming, and the editor of Population, Public Policy and Economic Development. He has published articles dealing with the economics of marriage, fertility, labor supply, and economic development in several professional journals including the American Economic Review, the Journal of Economic Literature, the International Economic Review, and Economic Inquiry.

Barbara H. Kehrer (Ph.D., Economics, Yale University) is a health economist on the senior research staff at Mathematica Policy Research. In addition to the SIME/DIME work appearing here, Dr. Kehrer also worked on the impact of national health insurance on physician behavior, determinants of physicians' incomes, and alcohol abuse. She has published in such journals as Journal of Human Resources, Inquiry, and Medical Care.

Kenneth Kehrer (Ph.D., Economics, Yale University) is director of the Survey Division and Senior Vice President at Mathematica Policy Research. He has contributed to the Evaluation Studies Review Annual, and prepared chapters for books, prepared numerous papers published by MPR, and numerous papers on the Gary income maintenance experiment, published by Indiana University.

David Kershaw (M.P.A. Woodrow Wilson School of Public and International Affairs, Princeton University) is President of Mathematica Policy Research. He has authored or co-authored several books on income maintenance and has contributed chapters to other books, and his articles appear in, e.g., Public Policy and Scientific American.

Patricia M. Lines (J.D., University of Minnesota) is the editor of this book and responsible for abridging the some 3,000 pages of transcripts and papers from the May 1978 SIME/DIME conference. She is an assistant professor at the Graduate School of Public Affairs, University of Washington, and has contributed to and edited two other books, and has written numerous scholarly articles appearing as chapters to books and in, e.g., the Texas Law Review, the Journal of Family Law, Inequality and Personnel Journal.

Michael Linn (M.S.W., Social Work, University of Washington) has been a consultant to SIME and has contributed to the editing of several papers in this book, as well as the general editing task. He was the first project director of SIME for the State of Washington and is presently a senior researcher for Mathematica Policy Research.

Larry Manheim (Ph.D., Economics, University of California at Berkeley) is a senior economist at Mathematica Policy Research and the author or co-author of numerous papers on health, education and welfare issues, one of which has appeared in the Journal of Economics and Business.

Cheri Marshall (M.B.A. Candidate, Harvard Graduate School of Business Administration) was Vice President and Deputy Director of the Survey Division of Mathematica Policy Research. She was responsible for development of questionnaires and survey methodology for SIME/DIME.

Myles Maxfield, Jr. (Ph.D., Economics, University of Maryland) is a senior economist at Mathematica Policy Research. He has been involved in the development of microsimulation methodology as a policy analysis tool. He has also worked in the areas of labor force turnover, the decision to participate in welfare programs, and monthly income reporting in transfer programs.

Charles E. Metcalf (Ph.D., Economics, Massachusetts Institute of Technology) is Director of Research and Senior Vice President at Mathematica Policy Research. His work appears in the Journal of Human Resources, American Economic Review, and Public Policy. He is the author of An Econometric Model of the Income Distribution (1972), and co-author of Industrial Location in the United States: an Econometric Analysis (1971).

Mary E. Minchella (M.S., Statistics, Rutgers University) is a statistician for Mathematica Policy Research. She has authored or co-authored several papers on income maintenance issues and other topics.

Robert A. Moffitt (Ph.D., Economics, Brown University) is an assistant professor of economics at Rutgers University and a research economist at Mathematica Policy Research. In addition to his SIME/DIME contributions, he has worked on the labor supply effects of the Gary income maintenance experiment, the AFDC program, and positive tax programs. His research has appeared in the American Economic Review, Southern Economic Journal, and Journal of Human Resources.

William Morrill (M.A., Public Administration, Syracuse University) is a senior fellow at Mathematica Policy Research. His work appears in Toward an Effective Support System: Problems, Prospects and Choices and Evaluation Magazine. He has also authored several papers on welfare reform and white collar crime for MPR. He is contributing a chapter to a book on welfare reform which will be published in the near future.

Arnold H. Packer (Ph.D., Economics, University of North Carolina) is currently the Assistant Secretary for Policy, Evaluation and Research at the U.S. Department of Labor. He has been the chief economist for the U.S. Senate Committee on the Budget, and for the Committee for Economic Development. He is the author of Models of Economic Systems: A Theory for Their Development and Use (1972) and has authored and co-authored numerous articles on various economic and social issues appearing in national journals.

John H. Pencavel (Ph.D., Economics, Princeton University) is presently affiliated with the Department of Economics, Stanford University, and Center for the Study of Welfare Policy, SRI International. He has made scholarly contributions to the general field of labor economics, and in particular to the economics of trade unions, relative wages, labor supply, and real wage movements.

Kristen Puckett is a researcher at Mathematica Policy Research and has contributed to papers on welfare and education while at MPR.

Philip K. Robins (Ph.D., Economics, University of Wisconsin) is a senior economist at the Center for the Study of Welfare Policy, SRI International. For the past several years he has been associated with the Seattle and Denver Income Maintenance Experiments as a principal researcher. He has authored numerous articles in the fields of labor supply, child care, and housing that are published in journals such as the American Economic Review, Journal of Human Resources, Economic Inquiry, and Journal of the American Statistical Association. He is also co-editor of a book entitled Child Care and Public Policy (1978).

Robert G. Spiegelman (Ph.D., Economics, Columbia University) is director of the Center for the Study of Welfare Policy at SRI International and has been involved in economic analysis of government programs at SRI for the past 19 years. During the past seven years, he has served as project leader of SIME/DIME. He was responsible for the design and implementation of these experiments, and for the determination of principal evaluation components. He has authored 11 reports of SIME/DIME evidence and impacts, and has contributed numerous articles to national journals, including the American Economic Review, Economic Inquiry, and the Journal of Human Resources, as well as one to a book by Robins and Weiner.

Ricardo Springs (M.A., Economics, University of Maryland) is at the Policy Studies Division of Mathematica Policy Research in Washington, D.C. He has prepared numerous papers for MPR on welfare issues, with an emphasis on analysis of alternative accounting problems.

Michael Stern (M.A., Columbia University) is the staff director for the Senate Finance Committee, and has had an opportunity to observe the political reception of the research analysis of SIME/DIME.

Manika Sukhatme (M.A., Economics, University of Southern California; M.A., Economics, University of Calcutta) is a research associate at Mathematica Policy Research in Washington D.C. and has prepared a number of papers for MPR on low-income women and related issues.

Peggy Thoits (Ph.D., Sociology, Stanford University) is presently assistant professor of sociology at Washington State University. She has worked as a sociologist for SIME/DIME at SRI International and prepared several papers on the subject while employed at SRI International.

Cynthia Thomas (Ph.D., Political Science, University of Rochester) is a project director and senior researcher at Mathematica Policy Research. She has written on housing, welfare, and other policy related topics. Before coming to MPR, she worked at the Urban Institute, taught political science, and served as a consultant to a variety of organizations.

Billy J. Tidwell (Ph.D., Social Welfare, University of Wisconsin) is senior sociologist with Mathematica Policy Research in Princeton. Dr. Tidwell was formerly a senior researcher with the Gary Income Maintenance Experiment. His research on income maintenance includes studies of various social consequences of negative income taxation; the knowledgeability of income maintenance participants; and participants' perceptions of income maintenance, both as an operating program and an experimental project.

Nancy Brandon Tuma (Ph.D., Sociology, Michigan State University) is presently an assistant professor at Stanford University and a consultant for the Center for the Study of Welfare Policy, SRI International. Her recent work is as a co-author with Hannan and Groeneveld on various "side effects" of income maintenance programs.

Samuel Weiner (Ph.D. Candidate, Economics, Stanford University) is a senior economist at SRI International. He has prepared several papers for SIME/DIME on issues relating to child care, and is co-editor of Child Care and Public Policy: Studies of the Economic Issues (1978), and The Federal Reserve System and the Supply of Money (1964).

Richard West (Ph.D. Candidate, Economics, Massachusetts Institute of Technology), is a research associate at SRI International. He has published several scholarly articles in, e.g., Journal of Human Resources, the American Economic Review, 1976 Proceedings of the American Statistical Association, and the Journal of the American Statistical Association.

Charles Wolin (Ph.D. Candidate, Economics, University of California at Berkeley) is an economic demographer and statistician at Mathematica Policy Research. His work on the income maintenance experiments includes analysis of the impact of the Gary experiment on low birth weight (with Kehrer) and on the fertility of participants.

APPENDIX C

SIME/DIME CONFERENCE PARTICIPANTS

May 13, 14, 15 and 16, 1978

Orcas Island, Washington

Curtis Aller	Professor of Economics San Francisco State University
Frank Anderson	Policy Analyst Office of Financial Management State of Washington
J. M. Anderson	Acting Director, Bureau of Social Services, Department of Social and Health Services (DSHS), State of Washington
Sally Anderson	Kansas Social and Rehabilitation Services
Burt Barnow	Office of the Assistant Secretary for Policy, Evaluation and Research, U.S. Department of Labor
Diane Barton	AFDC Program Specialist Texas Department of Human Resources
Rikki Baum	American Public Welfare Association Washington, D.C.
Lee Bawden	Director, Human Resources and Income Security, Urban Institute Washington, D.C.
Michelle Bell	School of Social Work University of Washington
R. Emmick	Senior Operations Analyst SRI International
Gordon Fisher	Senior Fiscal Analyst, Idaho Legislature, Joint Finance Appropriations Committee
David Forte	Executive Director, Colorado, Planning and Budgeting

Curt Funk	Conference Coordinator State of Washington, DSHS
J. Spencer Hammond	Staff Director, House Appropriations Committee, Washington State House of Representatives
Thomas Harper	DIME Project Manager
Beth Harris	Child Care Advocate Olympia, Washington
Ethel B. Harris	Planning Officer, Louisiana Health and Human Resources Administration
Wendy Holden	Assistant Administrator, Employment and Training, Employment Security Department, State of Washington
Joyce M. Hopson	Acting Regional Administrator, Region 5, State of Washington, DSHS
Catherine M. Lloyd	Deputy Commissioner for Administrative Management, Department of Social and Health Services, Alaska
Robert Lolcama	Administrator, State of Washington, DSHS
Al Loyd	Contracting Officer State of Washington, DSHS
Gladys McCarkhill	Chief, Office of Income Maintenance, State of Washington, DSHS
John M. McCoy	Program Director, Assembly Office Research, California State Assembly
Dardell McFarlin	Director, Department of Social Service, Monterrey County, California
Carol Mahoney	SRI International
Susanne E. Marten	School of Social Work University of Washington
John W. Mount	Systems Analyst/Manager Micro- Simulation, SRI International

C. Eric Munson	Economist, SRI International
Henry Neil	Staff Assistant, Committee on Appropriations, U.S. House of Representatives
Gary Smith	Assistant Director for Program Planning, Evaluation, and Budget, Department of Social Services, Colorado
Sidney E. Smith	Seattle
Kay Stevenson	School of Social Work University of Washington
Bernie Stumbras	Administrator of Economic Assistance, Wisconsin Department of Health and Social Services
Chris Swift	Seattle Urban League
David B. Swoap	Professional Staff Member U.S. Senate Committee on Finance
Irving Tallman	Chairman, Sociology Department Washington State University
Kay Thode	Director of Health and Welfare, Seattle Urban League
Arlene Wasberg	S/D Data Quality Supervisor, SRI International
James E. Walsh	Social Systems Design, Mercer Island, Washington
Frank Waynewood	Project Coordinator, SIME

