SUGGESTIONS FOR ASSESSING ECONOMIC AND DEMOGRAPHIC EFFECTS OF INCOME MAINTENANCE PROGRAMS

A REPORT PREPARED UNDER A GRANT FROM THE ECONOMIC DEVELOPMENT ADMINISTRATION

JULIE DA VANZO
DAVID H. GREENBERG

R-1211-EDA
JUNE 1973

Rand
SANTA MONICA, CA. 90406
This report was prepared for the Economic Development Administration, Department of Commerce under Grant No. OER388-G-71-11. Views or conclusions contained in this study should not be interpreted as representing the official opinion or policy of the Economic Development Administration.
SUGGESTIONS FOR ASSESSING ECONOMIC AND DEMOGRAPHIC EFFECTS OF INCOME MAINTENANCE PROGRAMS

A REPORT PREPARED UNDER A GRANT FROM THE ECONOMIC DEVELOPMENT ADMINISTRATION

JULIE DA VANZO
DAVID H. GREENBERG

R-1211-EDA
JUNE 1973

Rand
SANTA MONICA, CA. 90406
PREFACE

With the support of Grant OER 388-G-71-11 (Income Supplements and Regional Economic Development) from the Office of Economic Research of the Economic Development Administration (EDA), Department of Commerce, Rand has been examining some of the fundamental research issues that must be resolved before the effects of alternative income maintenance programs can be properly assessed. Thus, the research is intended more as a guide to policy analysts than as a guide to policymakers. Nevertheless, its goal is to bring closer that time when dependable guidance can be provided to policymakers in choosing among income maintenance proposals and in preparing for their potential effects.

Rand's research for EDA has been particularly concerned with the problems of assessing the regional, as opposed to the national, effects of income maintenance programs, but the most important issues that remain unresolved are common to both. Thus, much of our research should be of general interest to those concerned with the effects of income maintenance programs on individual and family behavior.

This report summarizes the main findings of research performed under this Grant, discusses the principal issues and problems in estimating the effects of income maintenance legislation, evaluates the reasonableness of further empirical study of this issue, and provides specific recommendations for future research. The other reports prepared under this Grant are:


SUMMARY

Widespread dissatisfaction with the current welfare system has resulted in numerous proposals for reform. This report is concerned with many of the problems faced by policy analysts in attempting to provide policymakers with the information they need to choose intelligently from among these proposals and to prepare adequately for the effects of whatever programs are eventually adopted. Particular emphasis is placed on problems of assessing the potential effects of proposed income maintenance programs at the subnational level.

Although researchers have made considerable progress in the last several years, the goal of providing dependable guidance to policymakers has not yet been reached. This report summarizes recent studies at Rand for the Economic Development Administration (EDA) and other studies that provide some of the basic research needed to reach this goal and from this generates suggestions for further research. The report does not provide policy analysis of specific income maintenance programs, but rather is a precursor to that analysis.

The role of social scientists in helping policymakers assess income maintenance alternatives is divided into six tasks:

1. To enumerate a set of socioeconomic indicators that communicate to policymakers an overall impression of the potential effects of programs, both at the national and the subnational levels.

2. To list the program provisions that may induce changes in these indicators, either directly or by changing the behavioral patterns of participating families.

3. To develop an appropriate theoretical framework for analyzing potential national and regional effects of alternative income maintenance programs.

4. To obtain the data necessary to estimate these behavioral parameters.

5. To estimate the parameters using properly constructed variables, correct statistical tools, and an appropriate sample of households.
6. To incorporate program provisions and the estimated behavioral parameters into a simulation that will predict the values of the socio-economic indicators that would result from adoption of each of the income maintenance alternatives under consideration.

This report discusses the first five of these tasks.

Several main themes underlie Rand's research on the behavioral effects of welfare reform. These themes, which are elaborated on in the report, represent important principles that should govern future research on the effects of income maintenance legislation whatever the geographic unit of interest.

1. The appropriate decisionmaking unit for analyzing welfare reform effects is the family; the appropriate data to use are data on households and their members.

2. Researchers must be careful not to confuse life cycle and cross-sectional effects; longitudinal data should be used to help control for life cycle effects.

3. Available experimental and nonexperimental data each have their advantages and should be used as checks on each other.

4. Researchers should be aware of the biases associated with alternative approaches—constructions of variables and model specifications—and should, whenever possible, use several different approaches to generate reasonable bounds for the behavioral parameters of interest.

5. Short-run parameters should not be used to predict longer-range effects.

6. Models should be formulated to account for the simultaneous nature of household decisionmaking, and appropriate econometric techniques should be used to estimate them.

7. Proper econometric techniques must be used to estimate relationships determining qualitative (dichotomous, polytomous) variables or dependent variables that are otherwise limited in range.

8. Unless one econometric approach to a particular problem is clearly superior a priori to the others, appropriate alternative estimation techniques should be tried and the sensitivity of estimates to the choice of technique should be determined.
9. Statistical tests of coefficient homogeneity should be performed to determine the appropriate level of subsample composition.

The report concludes with a list of areas on which we feel future research should concentrate:

1. We recommend that future research assess the relative importance of the possible sources of regional variation in program effects. Statistical tests should be performed to determine whether or not behavioral response parameters differ significantly by region. If they do differ significantly, it will be necessary to estimate separate parameters for each region; if not, a single set of parameters based on a national sample will suffice for each demographic group. Moreover, social experiments that are conducted in one region should be applicable to other regions.

2. We recommend further work on theoretical and empirical models of asset determination so that assets can be correctly used to derive accurate estimates of income effects. Assets should be treated as endogenous in estimation, and the sample used should be as free as possible from life cycle effects.

3. We recommend an effort to improve the methodology and performances of imputed wage rates.

4. We recommend that risk aversion be incorporated into the models used to study welfare reform effects so that the effect of an income maintenance program on a family's propensity to invest, particularly in its own human capital, can be determined.
ACKNOWLEDGMENTS

The authors are indebted to Rand colleagues W. P. Butz, D. N. De Tray, K. M. Maurer, G. Nelson, J. P. Newhouse, and A. H. Pascal, and to T. P. Schultz of the University of Minnesota for their helpful comments on several earlier versions of this report.
CONTENTS

PREFACE ................................................................. iii
SUMMARY ............................................................... v
ACKNOWLEDGMENTS .................................................... ix

Section
I. INTRODUCTION ..................................................... 1
   Tasks for Social Scientists .................................... 1
   Rand's EDA Research ............................................ 2
   Subnational Effects of Income Maintenance Programs .......... 5
   Organization ...................................................... 6
II. INCOME MAINTENANCE PROGRAMS AND THEIR EFFECTS .......... 7
   Provisions of Negative Income Tax Programs ................... 7
   Socioeconomic Indicators ...................................... 10
III. PROGRAM-INDUCED CHANGES IN BEHAVIOR ................. 15
   A Conceptual Model ............................................. 15
   Changes in Labor Supply ....................................... 17
   Changes in Migration Flows .................................... 19
   Changes in Consumption Patterns ............................... 21
   Changes in Investment Patterns ................................ 24
IV. ESTIMATING THE BEHAVIORAL PARAMETERS .................... 25
   The Experimental Approach .................................... 26
   The Nonexperimental Approach .................................. 27
   The Transition from Theoretically Correct Variables to
   Measurable Ones .................................................. 29
   Estimation Techniques ......................................... 46
   Choice of the Sample and Subsamples ........................... 56
V. EXISTING NONEXPERIMENTAL DATA SETS ....................... 63
   The Basic Samples .............................................. 63
   The Time Frame ................................................ 65
   Informational Content ......................................... 66
   A Brief Comparison ............................................ 69
VI. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH .. 72
   Some Basic Principles for Welfare Reform Research ........... 72
   Recommended Topics for Future Research ....................... 76

Appendix
SUMMARIES OF RAND REPORTS PREPARED FOR EDA ........... 79
I. INTRODUCTION

Widespread dissatisfaction with the current welfare system has resulted in numerous proposals for reform. This report is concerned with many of the problems faced by policy analysts in attempting to provide policymakers with the information they need to choose intelligently from among these proposals and to prepare adequately for the effects of whatever programs are eventually adopted. Although researchers have made considerable progress in the last several years, the goal of providing dependable guidance to policymakers has not yet been reached. This report summarizes recent studies at Rand for the Economic Development Administration (EDA) and other studies that provide some of the basic research needed to reach this goal and from this generates suggestions for further research. The report does not provide policy analysis of specific income maintenance programs, but rather is a precursor to that analysis.

TASKS FOR SOCIAL SCIENTISTS

The role of social scientists in helping policymakers assess income maintenance alternatives may be divided into six tasks. Since discussion of the first five of these tasks—the five that were treated in Rand's EDA research—occupies much of the remainder of this report, they are only briefly listed here.

1. To enumerate a set of socioeconomic indicators that communicate to policymakers an overall impression of the potential effects of programs, both at the national and the subnational levels.

2. To list the program provisions that may induce changes in these indicators, either directly or by changing the behavioral patterns of participating families. Examples of spheres of individual or family behavior that income maintenance legislation may affect include labor supply, the propensity to migrate, investments in human capital (such as in schooling and on-the-job training), consumption patterns, the propensity to marry and to stay married, and fertility. Changes in any of these may have substantial effects on socioeconomic indicators.
3. To develop an appropriate theoretical framework for analyzing potential national and regional effects of alternative income maintenance programs. This framework should generate testable hypotheses about the causal relationships between the provisions of income maintenance programs and family behavior. In other words, the framework should facilitate appropriate specification and estimation of the behavioral parameters that are needed to make predictions of the magnitude and timing of potential behavioral effects of alternative income maintenance programs.

4. To obtain the data necessary to estimate these behavioral parameters.

5. To estimate the parameters using properly constructed variables, correct statistical tools, and an appropriate sample of households.

6. To incorporate program provisions and the estimated behavioral parameters into a simulation that will predict the values of the socio-economic indicators that would result from adoption of each of the income maintenance alternatives under consideration.

RAND'S EDA RESEARCH

The Rand EDA research summarized in this report especially emphasizes the third and the fifth of the tasks--development of the theoretical framework and estimation of the parameters--but explicitly or implicitly treats important aspects of all but the last of the six points. This report also covers only the first five tasks, with relative emphasis among them generally similar to that found in the earlier research. However, since the report is a synthesis of findings from a somewhat disparate set of studies that do not treat every aspect of the first five tasks listed above, we have attempted to fill in gaps by drawing upon other appropriate research.

Rand's output from the EDA project is contained in six reports\(^1\) ranging from purely theoretical studies of family behavior (Smith) and

---

\(^1\)The Rand reports for EDA are listed in the preface. Henceforth, these reports will be identified by the surname of the author. In this section, we very briefly indicate the major topics covered by each of the EDA papers. More detailed summaries of the reports appear in the Appendix.
econometric techniques (Nerlove and Press) through a relatively comprehensive study of migration (DaVanzo) that touches on most of the points listed above.

De Tray's report is mainly concerned with income maintenance provisions that influence behavior and with the appropriate conceptual framework for analyzing that behavior. He examines the legislative history of the Family Assistance Plan with special reference to those quantifiable parameters of the programs that are likely to have important effects on household behavior, and then briefly outlines the conceptual framework underlying much of Rand's work for EDA, the "household production model." Finally, he illustrates the applicability of this approach to several behavioral areas, such as migration and investments in human capital, that are likely to be affected by welfare reform.

Three of the Rand EDA reports treat what is possibly the most important behavioral implication of income maintenance legislation: its effect on the supply of labor. Greenberg's report contains a detailed discussion of many of the problems that researchers face when estimating labor supply parameters from nonexperimental data. Topics covered include defining the sample population, conceptual and operational problems in appropriately measuring the principal variables--labor supply, wage rates, and nonemployment income--and controlling for other factors. The study is based on a critical review of the methodologies used in previous cross-sectional investigations of labor supply and on a comparison of the advantages and limitations of existing cross-sectional data sets.

Schultz's focus is on the labor supply of "secondary" workers, especially married women. He uses household data from the Survey of Economic Opportunity to investigate the sensitivity of labor supply parameter estimates to alternative solutions to several technical problems, some of which are discussed by Greenberg. In particular, Schultz's work for EDA explores the effect on parameter estimates of imputed versus measured wages and of several different econometric techniques.

Most studies of the effects of income maintenance programs on labor supply have conceptually been concerned with the average level of market work during an individual's lifetime. In general, however, these studies fail to recognize that many families will be eligible for income
maintenance benefits only during certain periods in their lives. In work partly funded by EDA, Smith uses the framework provided by the household production model to explore the implications of this fact. He concludes that the traditional approach to estimating individuals' labor supply responses is inappropriate and will overestimate lifetime changes in labor supply brought about by an income maintenance program. Two other issues are also explored in a life cycle context: how an income maintenance program will affect the relative labor supply of husbands and wives, and the consequences of investments in human capital for models estimating the labor supply effect of an income maintenance program.

DaVanzo's treatment of the implications for migration of income maintenance plans covers all but the first and the last of the six tasks outlined earlier. She formulates a model of interregional migration, estimates its parameters with aggregate Census data, and then suggests how the model could be extended with individual household data to predict the effect of income maintenance programs on the volume, direction, and demographic composition of migration flows. Her discussion includes a brief survey of existing micro-data sets that might be appropriate for further work on migration, a listing of the mechanisms through which income maintenance programs may influence migration, and an examination of various methodological problems in estimating the necessary behavioral parameters.

Many of the behavioral outcomes of income maintenance legislation are of a categorical or qualitative nature. That is, the outcomes can be represented by a dichotomous or polytomous dependent variable. For example, a family either migrates or it does not (a dichotomous variable); alternatively, various types of migration can be divided into several categories, such as intracounty, intercounty-intrastate, interstate, and so on (a polytomous variable). Furthermore, many program outcomes, such as those involving migration and labor supply, may be interdependent—that is, jointly determined. The Nerlove and Press report, which was partly supported by EDA, discusses the statistical problems one encounters when estimating relationships involving jointly determined, dichotomous or polytomous dependent variables.
SUBNATIONAL EFFECTS OF INCOME MAINTENANCE PROGRAMS

In conducting the research described above, Rand was particularly concerned with problems of assessing the potential effects of proposed income maintenance programs at the subnational level. Some of the potential effects of income maintenance programs, such as the migration of people or physical capital, are essentially subnational by nature and will cancel out if examined only at the national level. In addition, for several reasons, even effects that can be measured at a national level, such as changes in labor supply, may differ substantially from one geographic area to another. First, the provisions of an income maintenance program may vary geographically. State supplements to federal payments, for example, will cause the amount of the transfer to differ from one state to another and may also be associated with special state or local taxes on nonprogram income. Moreover, unless income supplements are adjusted for regional or rural-urban differences in costs of living, there will be geographical discrepancies in real support levels, even if nominal payments are everywhere the same. Second, the proportion of persons eligible for income maintenance payments will not be the same in each location since some areas contain more poor families than others. Third, the demographic composition of various subnational populations differs, and different demographic groups—for example, various age groups—may have different behavioral parameters and, hence, respond differently to an income maintenance program. Fourth, behavioral response parameters, even for a given demographic group, may vary geographically.

Very little existing research attempts to assess the relative importance of each of these possibilities. Yet the operation of these factors will determine the subnational effects of any particular income maintenance proposal. In this report we provide suggestions on how research on these matters should proceed.

Throughout our work for EDA, we have recognized that no matter what the geographic dimension of interest there is a set of fundamental

1 In the term "subnational" we include any geographic entities—EDA regions, states, cities, rural areas, urban areas, etc.—that are of policy interest. Except when otherwise indicated, our comments apply equally to all geographic units of interest.
research issues that must be resolved before the six tasks listed above can be successfully completed. Thus, although all the topics discussed in this report are relevant to assessing the potential subnational implications of alternative income maintenance programs, many are equally pertinent to examining potential national effects.

ORGANIZATION

This report is organized in accordance with the first five of the tasks listed above—that is, the five areas covered by Rand's EDA research. It provides a summary of the major implications for future research contained in the six reports described earlier and considers certain additional topics that are not explicitly covered in these reports. In Section II, we describe the major provisions of income maintenance programs that may affect behavior and present examples of socioeconomic indicators that can be used to indicate the effects of these programs. Section III provides a brief summary of the household production model. The framework this model provides for analysis of income maintenance programs is then used to examine hypothesized relationships between program provisions and several areas of household behavior. In Section IV, we consider the major problems in estimating behavioral parameters and summarize the comparative advantages of alternative approaches to these problems. Although we discuss controlled social experiments, which are increasingly important for estimating behavioral parameters, we focus primarily on problems associated with the use of nonexperimental survey data. Existing nonexperimental data sets are compared in Section V. Section VI contains a summary of our findings and our specific recommendations for further study of the economic and demographic effects of income maintenance programs as they influence subnational economies.
II. INCOME MAINTENANCE PROGRAMS AND THEIR EFFECTS

This section begins with a summary of some of the more important features of income maintenance programs, particularly those incorporating negative income tax principles. We then consider what policymakers need to know about the potential effects of such programs and present several examples of the sorts of socioeconomic indicators that might meet their requirements.

PROVISIONS OF NEGATIVE INCOME TAX PROGRAMS

Income maintenance systems encompass one or more of the following program components: negative income tax plans, wage subsidies, government-created jobs, and income-in-kind programs such as food stamps and subsidized housing and medical care. The income maintenance programs most intensively investigated by social scientists and most seriously considered by Congress over the last few years, such as the Family Assistance Plan (FAP), are those incorporating negative income tax features; for purposes of illustration, we shall also generally refer to this type of program. However, most of the points we make are equally relevant to other types of income maintenance programs.

In discussing negative income tax programs (or other income maintenance programs), economists have tended to emphasize the two most prominent program parameters: the base level of support and the tax rate. The base level of support defines the income guarantee—that is, a floor under household income. Viewed somewhat differently, the base level of support is the maximum income transfer a household can receive under the program. In general, the transfer amount is reduced as more income is received from nonprogram sources. In effect this means that nonprogram income is taxed. The proportion of nonprogram income that is subtracted from the base level of support is commonly referred to as the program tax rate.

The values of the guarantee level and tax rate for any negative income tax plan actually implemented will probably vary among classes of households; their values for any single household are determined by
other program provisions. Through these provisions households would be assigned to various administrative categories, each of which is associated with a unique set of tax rates and income guarantees. De Tray discusses these provisions under the general headings coverage, exemptions, work regulations, and other provisions.\(^1\)

**Coverage**

The type of coverage provided a given household is usually contingent upon the household's demographic characteristics, particularly the number of persons in the household, their age and sex, and the location of their residence. For example, under the existing welfare system or under various proposed forms of FAP, the base level of support increases with the size of the household. Many negative income tax proposals would also provide higher guarantee levels (and higher tax rates) for households headed by persons of retirement age. On the other hand, childless persons who have not yet reached retirement age are entirely excluded from program benefits under many proposals. Excluded households, in effect, are placed in an administrative category where the base level of support, and consequently the tax rate, is zero.

**Exemptions**

Under most negative income tax proposals, certain types of non-program income would be exempt from taxation. Moreover, even income that is defined as "taxable" may be subject to different tax rates depending upon its source. For example, under FAP, the wage income of participants would have been taxed at a lower rate than their income from nonemployment sources.\(^2\)

**Work Regulations**

Many negative income tax proposals include a requirement that adult members of participating households who do not have child-care

\(^1\)De Tray also summarizes (Table 1) the major provisions of the five versions of the Family Assistance Plan that were considered by Congress between April 1970 and June 1971.

\(^2\)See, for example, H.R. 1, June 1971.
responsible for work or training. The usual penalty for failure to comply is a reduction in the household's base level of support.

Other Provisions

Some participants in any negative income tax program actually implemented will probably also be eligible to participate in other income maintenance programs, such as social security, unemployment compensation, health care subsidies, day-care services, food stamps, and public housing. These programs often incorporate income guarantees or tax rates of their own. Depending on various program provisions, coverage under these additional programs will vary among the households participating in a negative income tax plan. As a result, when the full set of income maintenance programs is taken into account, different classes of households will be provided with different levels of basic support and subjected to different tax rates.

Guarantee levels and tax rates for households are also likely to vary by geographic location. The precise extent of geographic variation, however, will be contingent upon program provisions for cost-of-living adjustments and for state supplementation of federal income transfers. Proposed legislation to replace the present welfare system by a program such as FAP would tend to dampen, but not eliminate, existing regional variation in program parameters.

Another program provision that may have an influence on the transfer amounts received by different households is the length of the accounting period over which program benefits are calculated. For example, if the accounting period is six months long, it is possible under some negative income tax proposals for a household that receives all of its nonprogram income during the first six months of a given year to qualify for program benefits during the last six months, whereas a similar household, receiving the same annual nonprogram income but in equal parts throughout the year, does not qualify.
SOCIOECONOMIC INDICATORS

An important requirement for analysis of income maintenance programs is the formulation of socioeconomic measures that clearly show program effects to policymakers. The goal of income maintenance research should be to predict accurately the values of such indicators under alternative programs. Comparisons with pre-program values for these indicators would then allow potential program-induced changes to be estimated. The indicators should be sufficiently flexible to satisfy the particular needs of different sets of policymakers—for example, those at the regional as well as the national level—but there must not be so many of them as to overwhelm their users in a sea of numbers.

Projected changes in the indicators should be calculated for various geographic regions and for different demographic groups, as well as for the nation as a whole. This permits comparisons of the program effects on various groups and examination of the regional and demographic allocation of program transfer payments. It also allows assessment of the effects of alternative income maintenance programs on each region’s economy.

Temporal as well as spatial changes in the socioeconomic indicators must be considered. That is, if there were sufficient information on how program-induced behavioral adjustments in, say, labor supply and migratory patterns occur over time, one could project changes in the indicators for both the long and the short run. So far, however, the attempts of researchers to estimate behavioral parameters have been oriented toward short-run estimates. Since short-run estimates are likely to remain more easily obtainable than long-run estimates, policymakers should be aware of the danger of using the former to infer the latter.

The indicators must convey several different types of information. These are discussed below.

Effects on Poverty

Presumably, the raison d’être for income maintenance programs aimed at the poor is the mitigation, if not the eradication, of poverty. The effectiveness of a program as an anti-poverty device is perhaps most
naturally measured by comparing the number of persons or households who would be poor if the program were implemented with the number who would be poor were it not. One difficulty with this measure is in finding an operational definition of "poverty." A poverty line—a money income level sufficiently high to bring a family out of poverty—must be determined. The most commonly used poverty lines are federal standards, which were initially developed in 1964 by the Social Security Administration and are adjusted annually for changes in price levels. Although most users of these standards acknowledge that they are arbitrary, and considerable controversy rages over whether they are set too high or too low, they are probably adequate for comparing alternative income maintenance programs.

It may not be sufficient to compare these programs by simply counting the number of persons or families who cross a poverty line. A program that concentrated only on those whose incomes are slightly below the poverty line might then be judged superior to another program that actually did more to alleviate poverty but concentrated on those whose incomes are far below the poverty line. It may be better to assess the effects of the program on the basis of the total amount by which incomes that fall below the poverty line are in fact raised (that is, by how much the aggregate "poverty gap" is reduced).

**Budgetary Costs**

Public discussions of the "costs" of proposed income maintenance programs are usually concerned with the amount of income transfers that will actually be paid, the administrative costs, and changes in expenditures in existing programs that are associated with the introduction of the new programs. It is particularly important that such budgetary costs be projected separately for different units of government, since a budgetary increase at one level of government, say the federal, may in part be offset by a budgetary decrease at another level, say the state. Had FAP been implemented, this might have been one of its more important effects.
Market Output

Changes in transfer payment expenditures do not necessarily reflect program effects on the real resources that are used by an economy. These effects, however, will be partly reflected by program-induced changes in the earnings of participants—a measure of the market output these persons produce—and, at a more aggregate level, by changes in the Gross National Product or in measures of regional output. Relative expenditures for various market goods and services may also change. Such changes would occur if the program causes adjustments in the behavior of participants—for example, in the number of hours they supply to the labor market or in their geographic mobility. These behavioral changes and their relation to the socioeconomic indicators will be discussed in greater detail in Section III.

Demand for Government-provided Goods and Services

Program-induced changes in behavior may result in increased or decreased use by the poor of such publicly provided goods and services as education, medical care, and the State Employment Service. From one point of view, such changes will be reflected by the budgetary expenditures for these services, but from another point of view, they may appear as changes in the educational attainment and the wage rates of the poor.

Distribution of Income

Almost by definition, the income transfers provided by a new income maintenance program will change the distribution of income among income classes, demographic groups, and geographic regions. Projection of these changes will indicate which groups will be relatively better off and which will be relatively worse off. Potential program distributional effects may be estimated by projecting the total post-program disposable personal income of each income group.

Effects on Migration

In Section III, we suggest a number of reasons why income maintenance programs may influence family migration decisions. Program effects on migration should, of course, be indirectly reflected by the
socioeconomic indicators that are used to measure many of the program effects discussed above, such as those on market output within a geographic area, on interregional income distribution, and on state and local welfare budgets. Nevertheless, migration is a sufficiently important policy area that policymakers should also be provided direct measures of program-induced changes in the number and demographic composition of migrants to and from various regions.

Dissatisfaction on the Part of Nonparticipants

Income maintenance programs that cause uneasiness among nonparticipants, many of whom pay for the program, are likely either not to last very long or to place considerable stress on the social fabric; the current welfare system is a good example of the latter. Thus, it is useful to anticipate potential sources of such dissatisfaction as early as possible. Several of these, such as budgetary costs and distributional effects, are captured by some of the indicators already mentioned. Two other potential sources of dissatisfaction on the part of taxpayers may be (1) reductions in work effort by transfer payment recipients, as measured by changes in the number of such persons who participate in the labor force and in the number of hours they work, and (2) increases in the number of persons receiving transfer payments—that is, increases in those on welfare.

Nonmarket Activities

Although the indicators listed above are rather conventionally used, they tend to provide only limited perspective on program effects on the social welfare. For example, changes in the aggregate poverty gap or in family income are usually measured in dollar terms. If, after a program is implemented, a family withdraws so many hours from the labor market that its money income actually decreases, this does not mean that the family perceives itself as being worse off. Reduction of its time in the labor market indicates that it now feels this time is more valuablely spent engaging in nonmarket activities. A family's well-being includes the value of time outside the labor market as well as its value in the market. This is an important element of the "household
production model," a theory of family decisionmaking that is discussed in the next section.
III. PROGRAM-INDUCED CHANGES IN BEHAVIOR

A CONCEPTUAL MODEL

Accurate predictions of program-induced changes in economic indicators, such as those discussed in Section II, cannot be obtained without knowledge of how people change their behavior in response to various program provisions. Obtaining such knowledge depends on guidance from a conceptual framework that identifies possible causal links between specific factors in families' surroundings and specific aspects of people's behavior. Such a framework serves four functions: (1) It directs the researcher's attention toward aspects of behavior that may be significantly, though perhaps only indirectly, affected by an income maintenance program. (2) It indicates the specific characteristics of data needed to identify which of the causal links suggested are actually quantitatively important. (3) It specifies the characteristics of the empirical model that will be estimated with data. (4) It helps determine the correct statistical method to be used in estimation and inference.

The conceptual framework that Rand has used in much of its recent work on the behavioral implications of income maintenance programs is based on the "household production model." We believe that this is the most appropriate model for studying welfare reform effects because: (1) The family is the decisionmaking unit in this model; decisions about the types of behavior likely to be affected by welfare reform are made in a family context; (2) the model has been extended to study family decisionmaking over time; (3) the model measures individual and family well-being by their "full income"—the value of all their time, including time spent in productive nonmarket activities such as rearing children—and not merely by their money income.

The essential characteristics of this model are succinctly summarized in Smith's forthcoming report for EDA:¹

¹For more detailed descriptions of this model, see De Tray's EDA report and the references cited therein.
In this model, the family is viewed as if it were a small firm producing its ultimate wants within the household. In order to satisfy these wants, the family (firm) purchases market goods and services as one input in the production process.... The novelty and content of the household model come from the assumption that purchased market goods are not the sole inputs used by the family. Instead, households combine market goods with the time of various family members in order to achieve more basic desires. Families are therefore implicit demanders of their own scarce time resources. To illustrate: the commodity "enjoyment of a play" is ultimately consumed by purchasing directly such market goods as a theater ticket, travel to and from the theater, and babysitting services. Also, a considerable amount of time (which has alternative uses) of those family members who are involved is used up (in watching the play) and must properly be viewed as part of the full cost of consumption....

This approach differs from the traditional one since the price of any commodity now has two components—the money goods price and the time price. The latter has been generally neglected in the traditional approach, but both receive equal prominence in the theoretical structure of the new model. The relative empirical importance of the two components depends, of course, on their respective shares in the cost of producing an activity.

This model and other considerations indicate that behavioral changes following the introduction of an income maintenance program can be traced to three major sources:

1. By definition a new program will change the income and wealth positions of some households, consequently changing the amount and possibly the mix of "commodities" these households consume. Since these commodities are produced at home by combining market goods and the time of family members, a change in the demand for commodities will affect the derived demands for time and goods, and hence the household's allocation of time among home and labor market activities and its purchases of market inputs may change.

2. Some types of income may be taxed more heavily under the program than others. One would expect households to tend to shift their efforts toward obtaining those types of income that are subject to relatively low tax rates. Moreover, there is one particular type of "income"
that will not be taxed at all: the value a household places on goods and services that it does not purchase in the market place but produces itself (home repairs, for example). In other words, time exchanged for market goods and services—time in the labor market—is taxed under income maintenance programs, but time spent producing nonmarket goods and services or used directly in consumption activity is not. As a consequence, household commodities that require large inputs of goods and services purchased in the market become more expensive to households relative to those that require large amounts of nonmarket time. The relative consumption of the first type of commodity should, therefore, decline.

3. It was noted in Section II that income maintenance programs usually establish various categories, each associated with a different set of tax rates and income guarantee levels. To the extent possible, households will tend to tailor their behavior so that they will be placed in the category that offers them the most advantageous terms.

The ways in which these factors operate to change behavioral patterns are best illustrated by discussing the dimensions of behavior that might be affected by an income maintenance program and by indicating how changes in the socioeconomic indicators discussed in Section II will reflect these behavioral changes. Since our research for EDA concentrated on the labor supply and migration effects of income maintenance programs, our greatest emphasis throughout this report is also on these two behavioral areas.

CHANGES IN LABOR SUPPLY

Several different provisions of negative income tax programs may affect the labor supply of participating households. First is the income subsidy paid under such programs. Workers can use at least part of this subsidy to replace earnings given up by choosing to work fewer hours. In other words, the subsidy can be used to "purchase" time to engage in nonmarket activity. Second, such programs impose a tax rate on earnings, thereby changing the amount of market goods and services that can be obtained in return for an hour's work, but not the value of productive time spent outside the labor market. Third, program
provisions may stipulate that under certain conditions failure to accept a job offer will cause the amount of the subsidy to be reduced.

For reasons suggested in Section IV, the range of empirical estimates of the labor supply response to the introduction of an income maintenance program is unacceptably large. For programs similar to the Family Assistance Plan, estimates of the labor supply reduction for prime age males based on nonexperimental data have ranged from 0 to over 30 percent; the information currently available from on-going income maintenance experiments implies that little reduction in the hours of prime age males will take place.

Policymakers have at least five reasons to be concerned about the magnitude of the labor supply response:

1. The number of hours worked is a major determinant of the amount of goods and services produced by the economy.

2. The transfer costs of the program depend on the work effort of participants. For example, if participants reduce their hours of market work, thereby causing their earnings to fall, part of this loss of earnings will be replaced by program-sponsored income transfers.

3. Presumably one goal of policymakers in designing income maintenance programs for the poor is to improve participants' money incomes. Since most negative income tax programs would replace only part of any earnings loss resulting from an hours reduction, such a reduction would lessen program effectiveness in increasing the money incomes of the target population.

4. The political acceptability of the program also hinges on its effect on labor supply. The Protestant ethic implies that work is good per se and especially good for the poor. Thus, policymakers and taxpayers seem much less resistant to subsidy payments that are used by poor households to purchase goods and services than they are to payments used to replace earnings losses caused by a reduction in market work.

5. Changes in hours worked may alter the geographic allocation of the supply of low-wage labor, especially if wage and income response parameters vary geographically.¹ Thus, firms may relocate.

¹Very little is known about whether labor supply parameters vary geographically, although some evidence suggests that there may be rather
These areas of concern are to some extent accounted for by the socio-economic indicators discussed in Section II. For example, attitudes of taxpayers toward the program and its participants are likely to be most affected by the magnitude of the changes in hours itself and by changes in the number of welfare recipients who participate in the labor force. However, changes in the output of market goods and services that these hours adjustments entail are best measured by aggregate changes in earnings. Program budgetary costs are directly measured by the change in total welfare transfer payments resulting from implementation of the program, and program effectiveness in improving the money incomes of the poor is reflected by changes in the total disposable personal income of participants and by changes in the aggregate poverty gap. If knowledge of program-induced labor supply adjustments were available, the importance of these adjustments could be estimated by comparing (1) projections of the changes that will occur in total welfare transfer payments, the aggregate poverty gap, and participants' total disposable personal income with (2) what they would be if the labor supply remained unaltered.

CHANGES IN MIGRATION FLOWS

Among other reasons, people may migrate because a new area offers better employment opportunities, greater consumption returns (for example, a superior climate or the opportunity to live closer to relatives), a lower real cost of living, and, perhaps, in the case of some poor families, superior income maintenance programs. By changing the level of income supplements available to certain households and the set of tax rates that these households face, any new income maintenance program will affect the likelihood of moving for one or more of these reasons.

substantial variation among regions. David Greenberg (Income Guarantees and the Working Poor in New York City: The Effect of Income Maintenance Programs on the Hours of Work of Male Family Heads, R-658-NYC, March 1971) has estimated labor supply functions for (1) a national sample, (2) central cities, (3) large central cities, and (4) New York City. He has found the parameters to differ considerably among the four samples. As stressed later in this report, investigating whether parameters vary geographically should be a topic for future research.
The mechanisms through which these factors operate are discussed in some detail in DaVanzo's report and are briefly summarized here.

The income supplement associated with a negative income tax program will increase the gross flow of migrants by facilitating the financing of migration, subsidizing job search, and providing income insurance in a new area. On the other hand, the tax rate on earnings, if higher than the family is currently facing, will reduce a recipient family's potential earnings in all locations and thus decrease the absolute earnings differential available by moving to a new location; this is expected to weaken the incentive to search for and migrate for better job opportunities. However, a higher tax rate will also reduce any earnings forgone during the time the migrant is between jobs. This will tend to mitigate the negative effect on migration of a smaller labor market return.

At present the average monthly welfare payment per recipient varies considerably among states.\(^1\) It has been alleged that these interstate discrepancies encourage "nonproductive" migration from low-benefit states to those with more attractive programs.\(^2\) Congressionally enacted welfare reform measures are likely to establish national standards that reduce the discrepancies in interstate welfare payments, although the extent to which this happens will depend on provisions for state supplements and cost-of-living adjustments. This should reduce any existing welfare-motivated migration. In addition to the effects of income supplements and tax rates, income maintenance programs may also facilitate

---

\(^1\)For example, in 1966 the average AFDC grant per recipient was $8 per month in Mississippi, but $48 per month in New York. R. A. Cloward and F. F. Piven, "Migration, Politics, and Welfare," *Saturday Review*, November 16, 1968, point out that eligibility requirements and administrative procedures also exhibit considerable interstate variation.

\(^2\)For examples of such allegations see the list of quotes in Robert D. Reischauer, "The Impact of the Welfare System on Black Migration and Marital Stability," Columbia University, Ph.D. dissertation, 1971, p. 68, footnote 74. Available evidence does not lend very strong support to the hypothesis that (potential) welfare recipients tend to move to areas with attractive welfare programs (see DaVanzo, p. 67, footnote 1). However, most studies have considered only interregional discrepancies in benefits, neglecting differences in eligibility requirements and administrative procedures.
migration through provisions for moving allowances, job locations, and information services. Partly because of this multiplicity of effects, existing empirical studies of migration are inadequate for predicting how net migration between given pairs of regions would be affected by proposed income maintenance programs. In fact, it is not even clear at the present time whether these programs would increase or decrease the proportion of the population that annually migrates. It is important, however, that policymakers be provided with this information; along with the other program behavioral effects, the changes in the sizes and compositions of regional populations that result from the effects of income maintenance programs on the volume, direction, and demographic composition of migration flows will determine the regional distribution of program transfer payments. Program-induced population redistribution also can be expected to have important consequences for regional economic growth and development.

CHANGES IN CONSUMPTION PATTERNS

Changes in income levels and tax rates that result from the introduction of a new income maintenance program have obvious implications for consumption patterns. For example, if income increases, other things equal, a family will buy larger quantities of all goods whose income elasticities exceed zero, and will especially increase its consumption of commodities with very high income elasticities.\(^1\) Some of these increases in expenditures will be recirculated through the economy and, through "multiplier effects," should generate additional increases in expenditures, outputs, and incomes.\(^2\) These will be reflected by

\(^1\) A staff study sponsored by the President's Commission on Income Maintenance Programs tentatively concludes that "for most [poor] persons 50 to 70 percent of any increment in income will be spent on the three basics, food, clothing, and shelter" and that "substantial portions of additional expenditures would be devoted to transportation, tobacco and alcoholic beverages, and reading, recreation, and education." Chapter 4.1 on "Household Consumption," in Background Papers, President's Commission on Income Maintenance Programs, 1970, pp. 79-84.

\(^2\) For a more detailed discussion of the macroeconomic implications of a negative income tax program, see Christopher Green, "A Macroeconomic Analysis of the Economic Impact of Negative Income Taxes,"
changes in disposable personal income and in other socioeconomic indicators for the nation as a whole and for subregions.

However, the extent and direction of multiplier effects are unclear, a priori. First, any change in transfer income receipts that follows the adoption of an income maintenance program may be at least partly offset by an opposite change in earnings. This would occur, for example, if a program that increases transfer receipts also induces a reduction in hours of work and, hence, in earnings. Second, if a new program is funded through a tax increase, any increase in transfer income received by participating families may be offset by a corresponding reduction in the disposable personal income available to nonparticipating families. Of course, if the marginal propensity to consume of the low-income recipient population exceeds that of the nonrecipient population, positive multiplier effects may still be produced. Moreover, the geographic distribution of income is almost certain to change. Thus, positive multiplier effects may be produced in regions with a disproportionate number of poor persons and negative multiplier effects occur in wealthier areas.

In addition to changing the magnitude and composition of recipients' income, an income maintenance program that taxes wages will change an important price of all consumption—the price of time. As mentioned previously, such a program would tax time spent working (market time), but would not tax productive nonmarket time. Increases in the tax rate on earnings resulting from the introduction of an income maintenance program will cause commodities that use large inputs of nonmarket time to become less expensive relative to those using a large proportion of market goods and services. "Enjoyment of children" is one such time-intensive commodity. An increase in the tax rate on earnings may make children more attractive to households because the earnings that are forgone by staying home and caring for them have, in effect, been reduced.

Another feature of specific forms of income maintenance legislation that may also increase the demand for children is the provision in

many proposed programs that no family can receive benefits unless it has at least one child. In addition, even after a family qualifies for benefits, the guarantee level under many proposed income maintenance programs is positively related to family size. For example, under H.R. 1 the marginal value of the first child to a couple with no nonprogram income would have been the base level of support, $2000 a year; the marginal values of a second and third child, $400 each. Thus, if a child adds less to household costs (both psychic and monetary) than the value of welfare benefits, the family might restyle their lives—that is, have a child they might not otherwise have had—in order to move into the recipient class or receive additional benefits. Furthermore, some programs include provisions for subsidized day care. This may have a positive effect on desired family size by reducing the time cost of children.

Through its income and price effects, the imposition of an income maintenance program will affect a family's demand for publicly provided services as well as its demand for private goods. Important examples of such services are publicly funded schools and subsidized medical care. The money price of these services to recipients is either zero or very low. A major (if not the only) cost of such services to the recipient is the time that must be used in their consumption.1 In other words, the consumption of public goods tends to be very time intensive. Thus, any increase in tax rates that results from imposition of a negative income tax program should produce an increase in the demand for public goods.2

1Recent Rand empirical studies of the demand for medical services have indicated that time-price elasticities tend to rise as money prices fall. See Jan Acton, Demand for Health Care Among the Urban Poor, with Special Emphasis on the Role of Time, R-1151-OEO/NYC, April 1973; and Charles Phelps and Joseph Newhouse, Coinsurance and the Demand for Medical Services, R-964-OEO, April 1973.

2A tax on earnings will reduce the time price of consuming public goods for individuals with market options for their time; it may not necessarily increase the demand for such goods for persons who do not customarily participate in the labor force, for example, young children. An increase in tax rates on wages may encourage older children to stay in school longer but have little effect on school attendance by younger children.
CHANGES IN INVESTMENT PATTERNS

The primary types of investments made by likely participants in income maintenance programs are those in their own human capital—schooling, on-the-job training, and so forth. Since investments in human capital typically require sizable inputs of time, the effects of changes in income levels and tax rates on such investments should be similar to the effects of such changes on time-intensive commodities. For example, an increase in tax rates on earnings would be expected to increase the demand for job training and other forms of human capital by reducing the opportunity costs of such investments.\(^1\) Program-associated increases in income levels should increase the demand for investments in human capital by facilitating their internal financing.

The fact that income maintenance programs place a floor under income may also encourage investments in human capital because many forms of investments in human capital, such as migration, job search, and specialized job training, tend to be relatively risky ventures; but by reducing future variance of income, the income floor reduces the risk associated with the investment.\(^2\)

Therefore, any new income maintenance program that raises the income levels and tax rates of its participants is also likely to increase the number of households willing to invest in human capital. Assuming a corresponding increase in the supply of sources of those investments, such as training facilities, the long-run result could be an increase in the market earnings of such families and, hence, a decrease in their participation in the programs.

\(^1\)However, this incentive may be offset if post-investment earnings remain sufficiently low that the investor is still eligible to participate in the program. In this case, the tax rate will lower the eventual return expected as a result of investments in human capital.

\(^2\)For a more detailed discussion of this point, see De Tray, p. 20.
IV. ESTIMATING THE BEHAVIORAL PARAMETERS

In the last section, we identified possible implications of income maintenance programs for various dimensions of household behavior. It was seen that most potentially important behavioral responses appear to result from program effects on incomes and wage rates. To measure program effects, therefore, one must determine the relationships between changes in wages and income and the types of behavioral decisions of interest. The two principal alternative approaches to obtaining information on how wage and income changes affect behavior are through data generated by controlled social experiments and through inferences from nonexperimental data collected by household surveys.

Under either approach, a critical issue for determining the regional implications of income maintenance programs is the extent to which behavioral responses to wage and income changes vary geographically. To minimize administration costs, social experiments are almost always confined to relatively small geographic areas; thus, the question of their applicability to regions not covered by the experiment is of obvious importance. As opposed to experimental data, most of the better nonexperimental data are based on national samples of households. However, it is frequently impossible to divide such national samples into subsamples that closely correspond to the regions of interest.\(^1\) In addition, as a sample is subdivided, subsamples can become quite small. This problem is compounded because, in addition to regional variation in behavioral parameters, it is also important to know whether the parameters vary by age, race, sex, and so forth. As the number of observations in a subsample shrinks, the reliability of inferences based on that subsample is weakened.

---

\(^1\) Ideally, we would like to have data that give the household's area of residence by a relatively small geographic unit, such as state, SMSA, or county. Unfortunately, many household-level data sets suppress detailed information on residence to maintain confidentiality and provide information only on larger geographic units such as the nine Census divisions or four Census regions. The level of geographic detail available in five individual-level data sets is noted in Section V, Table 1.
Later in this section we discuss statistical techniques that can be used to examine the extent to which behavioral parameters differ among regions and among different population groups. Such an examination should begin to indicate whether existing experimental and non-experimental data can be used to generate reliable predictions of the regional implications of income maintenance programs or whether such predictions must await collection of separate, expensively obtained data for each region of interest.

THE EXPERIMENTAL APPROACH

Under the social experiment approach, a sample of selected households is randomly divided into subgroups. Wage and income conditions are then made to vary among the groups and the behavior of the groups compared. For example, households that participate in a pilot program can be compared with a control group of similar households that do not participate.

The obvious advantage of this approach is that it allows behavioral response conditions to be estimated under circumstances that are as close as possible to those that would exist under an actual program. Nevertheless, it is not really possible to duplicate program conditions accurately in all respects. Results from an experiment may be distorted by the intrusion of unforeseen changes in environmental conditions that are not subject to experimental control\(^1\) or by households that do not adjust their behavior during a temporary experiment in the same way they would if the change were permanent. For example, individuals who might withdraw from the labor force completely if an income maintenance law were passed that they expected to be permanent may not withdraw during the experimental period if they suspect this will reduce their subsequent opportunities for employment. Moreover, if the pilot

\(^1\)For example, interpretation of the results of the income maintenance experiment in Seattle was severely complicated by a drastic increase in the city's unemployment rate during the experimental period. See Mordecai Kurz and Robert G. Spregelman, "The Seattle Experiment: The Combined Effect of Income Maintenance and Manpower Investments," *The American Economic Review*, 61, No. 2, May 1971, 22-29.
program has complex rules, some time may elapse before the rules are sufficiently understood by participants for behavioral adjustments to occur. Some measure of the importance of this problem can be obtained by allowing the length of the experiment to vary for different groups of households. An additional problem in experimental data is the so-called Hawthorne Effect: Families who are being observed continuously may behave differently from other families with similar characteristics simply because they are being studied. One technique for measuring the extent of this problem is the use of an extra control group that is interviewed less frequently than the bulk of the sample.¹

Possibly the most serious drawbacks of social experimentation have to do with cost and time constraints. Such experiments are relatively costly, although it is important to recognize that much of the "cost" in an income maintenance experiment may be payments made to the low-income persons participating in the experiment. Nevertheless, administrative costs are sufficiently large that experiments tend to be carried out in fairly narrow geographic areas, bringing into question the applicability of their results to all regions of interest. In addition, experiments do not generate results until considerable time has elapsed; policymakers may not be willing to wait until relevant experimental results are available.

THE NONEXPERIMENTAL APPROACH

Because experimental data are expensive and time-consuming to generate and are subject to biases of unknown magnitude, it is important to consider estimating behavioral parameters from nonexperimental data, especially data that are already available. Particularly important is the use of such data to test whether behavioral response parameters vary across regions, thereby helping to determine whether the results of a social experiment carried out in one geographic area are applicable to other areas.

As in the case of social experiments, making inferences from non-experimental data involves comparing the behavior of households for whom wage or income conditions differ. With nonexperimental data, differences among households are not artificially induced but are observed under existing conditions; this approach is sometimes referred to as a "natural experiment." Unlike controlled experiments, differences among households in variables other than those of policy interest cannot be "held constant" by dividing a sample into randomly selected experimental and control groups but must be controlled through various statistical techniques.

Although progress has been made in learning how to use nonexperimental data to estimate behavioral parameters, substantial gaps remain. For example, even though economists have put considerable effort into estimating labor supply response parameters from nonexperimental data, their empirical conclusions remain disturbingly varied. To illustrate, it is useful to look at substitution and income elasticities calculated from three recent labor supply studies. All three studies used the

<table>
<thead>
<tr>
<th>Study</th>
<th>Estimated Substitution Elasticity</th>
<th>Estimated Income Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Garfinkel</td>
<td>Near 0 or negative</td>
<td>0</td>
</tr>
<tr>
<td>Greenberg &amp; Kosters</td>
<td>.02 to .2</td>
<td>-.15 to -.34</td>
</tr>
<tr>
<td>{Whites}</td>
<td>{Range: -.2 to .39}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>{Weighted ave.: .06}</td>
<td></td>
</tr>
<tr>
<td>{Blacks}</td>
<td>{Range: -.68 to .25}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>{Weighted ave.: -.1}</td>
<td></td>
</tr>
</tbody>
</table>

Source: Glen Cain and Harold Watts (eds.), *Income Maintenance and Labor Supply: Econometric Studies*, Markham Press, Chicago, forthcoming, Table 9.1 from the concluding chapter by Cain and Watts. All three of these studies appear in the Cain and Watts volume.

1 The income elasticity is the percentage change in hours of work that would result from a one percent change in household income (holding wage rates constant); the substitution elasticity is the percentage change in hours that would be caused by a one percent change in a worker's wage rate (holding household income and the wage rates of other family members constant).
same data base—the 1967 Survey of Economic Opportunity—and all three concentrated on roughly the same population group—married men too young to collect Social Security.

There are a number of reasons why household behavioral parameters estimated from nonexperimental data have so far been so inconsistent that they are of limited usefulness. One reason is that the conceptually correct definitions of the variables whose relationships are to be estimated are not entirely clear, and, even if they were, the transition from theoretically correct variables to measurable variables is extremely difficult. The statistical technique used for estimating the relationships among the variables, once measured, must deal correctly with the simultaneity in household decisionmaking; that is, it must take account of the fact that one type of behavioral response by an individual—say, a labor supply adjustment—is interdependent with other areas of behavior and with the characteristics of other members of the household. Finally, the relations estimated between the variables may be very sensitive to the sample of households or individuals that is drawn from the nonexperimental data.

The least tractable of the problems involved in estimating the behavioral parameters tend to derive from the same source: In a "natural experiment" there may simply be too few occasions when circumstances sufficiently approximate those that would occur upon introduction of an income maintenance program; that is, we may seldom, if ever, observe in nonexperimental data families' reactions to truly exogenous changes in wage rates and income. Changes in these variables that would result from introduction of an income maintenance program are outside the affected households' control. Wage and income differences that are observed in nonexperimental data, however, may have been strongly influenced by decisions made by the sampled households.

THE TRANSITION FROM THEORETICALLY CORRECT VARIABLES TO MEASURABLE ONES

The variables that are used in a theoretical model of household behavior are often not easily defined and constructed with available data. This subsection deals with the important transition from theoretically correct variables to measurable ones and details some of the biases
that may be introduced in the process. We first discuss the major methodological problems associated with defining the two principal explanatory variables— that is, with appropriately representing the changes income maintenance programs make in wage rates and in income levels. We then consider difficulties in constructing dependent variables to represent two types of behavior that may be influenced by such programs— labor supply and migration.

**Wage Rates**

To determine the effects of the tax on earnings that an income maintenance program imposes, we must have an accurate estimate of how individuals respond to an exogenous change in their market wage rates. Wages, which are the most important source of income for most families, are usually fairly accurately reported. However, this does not mean that they are free of measurement error or other problems. We shall discuss five major problems: (1) the actual wage rate upon which an individual's behavioral decisions are based—his marginal wage rate, net of taxes and adjusted for cost of living and fringe benefits—is seldom reported in available data; (2) wage measures available in most data sets may be subject to measurement errors that systematically bias estimated relationships; (3) it is difficult to determine the effect of an exogenous change in wage rates since many people's wage rates are jointly determined with other behavioral decisions; (4) anticipated future wages, as well as current wage rates, affect behavior; and (5) the potential wage rates of nonworkers affect family decisionmaking and should therefore be controlled for, but these are rarely reported and are difficult to infer. Imputed wage rates (that is, measures inferred from personal characteristics that are thought to affect wage rates) can be used to treat the second, third, fourth, and fifth problems, but, as discussed below, operational attempts to use imputed wages have not as yet been very successful.

The wage measure that is reported in most available data is a worker's gross average wage rate. However, to estimate the effect of a change in wages, one needs a measure of the amount of market goods and services that can be exchanged for an additional hour of market
work. Since individuals make their decisions at the margin, the marginal wage rate is the theoretically correct variable and should be used whenever possible instead of average wage rates to provide an accurate estimate of the effect of a wage change.\footnote{Some information on \textit{marginal} wage rates is available in the Income Dynamics Panel and the National Longitudinal Surveys. See Section V, Table 1.} Marginal and average wage rates will differ for many of the persons under consideration. For example, the wage rate "moonlighters" or overtime workers receive for their last, or marginal, hour of work is usually different from their average hourly compensation. Gross wages should also be adjusted for fringe benefits, cost-of-living differentials, and taxes.\footnote{Schultz adjusts his wage measures for federal income taxes and for cost-of-living differentials, but not for fringe benefits since the SEO file does not contain the information necessary for this adjustment.} For example, since marginal and average wage rates differ for all persons subject to progressive taxes, failure to take account of these taxes in the analysis will introduce a bias toward zero in the estimated relationship between wages and labor supply.

A second problem in measuring wages from available data results from the fact that the wage variable is often constructed by dividing a measure of total earnings over some period of time by the number of hours worked during that period. If, as is often the case, one or more of the data elements used in the denominator of the wage rate variable is also used to construct the labor supply measure, reporting errors may negatively bias the estimated relation between labor supply and the wage rate. In addition, the reported wage often pertains only to a single, short period of time (say, the week of the survey), and may not, therefore, be representative of the wage an individual usually receives and on which he bases his decisions.\footnote{The "Parnes" National Longitudinal Surveys ask interviewees to give their "usual" earnings and hours on their current or most recent job. Measurement error is further reduced in this survey by allowing participants to select for themselves the time unit in which to report their usual earnings (dollars per hour, week, month, and so on).} These biases, resulting from measurement error, may be avoided partly through use of imputed wage rates, an approach discussed below.
Imputed wages may also help to overcome simultaneity biases. These biases are caused by the fact that in many instances wage rates are determined jointly with other behavioral decisions. For example, wages are expected to influence the number of hours worked, but hours worked may in turn affect the wage rate the worker obtains. This is illustrated in the construction industry where it is frequently alleged that workers are compensated for the seasonal availability of work by relatively high hourly wage rates. Similarly, if families face progressive tax rates, increases in their hours may raise them to a higher tax bracket, thereby reducing each member's wage rate net of taxes. Thus, wages are not completely exogenous to much household decisionmaking. Moreover, as we increase the time horizon under consideration, they become increasingly less exogenous because they can be affected more by work experience, job training, and other investments in human capital.

Another problem with using the observed wage rate is that this measure may be systematically related to the wages people expect to receive in the future, and current behavior may be partly based on these future expectations. Smith points out in his EDA report that ignoring anticipations regarding wage profiles over the life cycle can lead to seriously biased results.\(^1\) For example, if individuals who receive relatively high wages anticipate more steeply sloped future wage profiles than low-wage individuals, the effect of current wage on labor supply is likely to be underestimated. This is because, everything else equal, individuals should attempt to work their longest hours at the point in their life cycle when wages are highest.\(^2\) To help control for the effect of differences in average or permanent wages, Smith suggests estimating

---

\(^1\)In statistical terms, this may be viewed as an omitted-variables bias. That is, whenever an explanatory variable (such as anticipated future wages) that is correlated with other explanatory variables (such as observed wages) is omitted from the model, the coefficients of the other variables in the model will be biased. In this case, of course, the omitted explanatory variable is impossible to observe.

\(^2\)To determine life cycle effects on the timing of work effort, Smith ran regressions where the observations were the mean values for age cells; thus the observed variation in wage rates was presumably variation over the life cycle rather than variation in the average (lifetime) price
regressions using cross-sectional data on narrowly defined age groups. When the interactive effect of age and wages on hours is held constant in this manner, most wage variation should be due to permanent differences in wage rates. Imputed wages might provide further control for life cycle effects.

Nonworkers present particular problems in estimating the effects of welfare reform because wage rates are seldom reported for such individuals. However, as we will later argue, it is important that nonworkers be included in the analysis. One solution is to impute a potential wage rate for nonworkers (and workers who did not report their wage rates) on the basis of the wages observed for persons of similar characteristics who did report earnings. This is done by estimating a wage function for a sample of persons reporting wages. Their reported wage rates are regressed on variables that are hypothesized to affect labor market productivity and wage levels—for example, education, age or work experience, and geographic variables. The potential wage rates of nonworkers can then be imputed by substituting their characteristics into the wage function.

One difficulty with the imputed wage rate approach is the implicit assumption that, for a given set of characteristics, those who do not work face the same potential wage rate as those who work. Since the latter group works, but the former does not, either their potential wage rates differ or they entertain different options for productive nonmarket uses of their time. The household production model tells us that an individual will not participate in the labor force if the wage he would receive is less than the value of his time in nonmarket production. Thus, a person may choose not to work because his potential market wage

---

1. However, if work experience is endogenous (correlated with the error term), ordinary least squares estimates of a wage equation including experience will be inconsistent (see p. 47).

2. Thus, it is essential that all of the explanatory variables in the wage equation be available for nonworkers as well as workers.

is low; on the other hand, he may not participate in the labor force because he is very productive in nonmarket activities. It is at least conceptually feasible that the potential wage rates of nonworkers exceed the wage rates received by workers with similar characteristics because the former are so productive in nonmarket activities that they devote no time to labor market work.\footnote{Greenberg has suggested that longitudinal data files might provide estimates of \textit{some} of the adjustment factors needed to reconcile differences between workers and nonworkers, for they permit one to compare the wages of persons who did not work in some periods but did work in others with the wage of persons who worked in all periods covered by the file.} It is not clear, \textit{a priori}, therefore, whether one understates or overstates the potential wage rates of nonworkers when using an imputation technique such as the one just described.

Reuben Gronau discusses an additional difficulty in imputing wage rates that would exist even if one were able to measure an individual's productivity in nonmarket activities—a "selectivity bias."\footnote{Reuben Gronau, "The Wage Rates of Women—A Selectivity Bias," National Bureau of Economic Research, September 1970, mimeo.} Assume we have a group of persons with the same nonmarket value of time and the same wage-determining characteristics. Because different employers offer different wages for identical bundles of characteristics, these individuals will face a \textit{distribution} of wage offers rather than a single wage offer. Some members of the group will be fortunate enough to receive relatively high wage offers and will choose to work; others, who happen to receive low offers, will discontinue job search (because it is costly) and choose not to work. What we will observe is the \textit{truncated} wage offer distribution of persons who do in fact work. Since under these assumptions the unobserved part of the distribution is the lower tail, the mean wage offer of the observed truncated distribution will exceed the true mean wage offer (the mean of the entire offer attempted to estimate the extent to which the nonmarket value of time for women who do not work exceeds the market wage of women with the same market-wage-related characteristics. He finds that the marginal value product of time in the home for nonworking wives exceeds the average market wage of wives with the same wage-related characteristics by 12 to 15 percent.
distribution). Gronau points out that the magnitude of this bias is negatively related to the labor force participation rate of the group under consideration. Thus, for groups with low labor force participation rates, such as secondary workers (married women, the young, the aged), the mean or median observed wage may considerably overstate the true mean or median of the entire wage distribution. It is very important that researchers learn to control for or at least attenuate these sources of bias so that the potential wage of nonworkers can be accurately imputed and these persons can be properly incorporated into the analysis.

To summarize, an imputed or predicted wage variable can potentially mitigate four problems that result from use of observed wage measures. It can be used to (1) correct for transitory influences and measurement error bias; (2) avoid simultaneity bias; (3) help control for life cycle influences; and (4) measure wages for nonworkers. Schultz has used both reported wages and an imputed wage rate to estimate labor supply functions for secondary workers and finds that the estimated parameters are sensitive to which wage measure is used. However, the imputed wage methodology is far from perfected. We have already discussed many of the difficulties encountered in imputing wages for nonworkers. In practice, most empirically estimated wage equations have tended to account for only a small to moderate proportion of the variation in observed wages.¹ There is a substantial risk, therefore, in using an estimated wage, that we are throwing away useful information—the unexplained residual. Some of the unexplained variation in wages is undoubtedly attributable to important factors, such as quality of education or training, that influence the behavioral decisions in which we are interested, but for which there are poor or no measures. Therefore,

¹There is a tradeoff between explanatory power and consistency. If endogenous variables such as marital status, parity, or work experience are included in the wage equation, its explanatory power may be fairly high (around 50 percent) but the estimated coefficients will be "inconsistent" (consistency is a statistical term that is defined in footnote 2 on p. 47). However, if variables correlated with the error term (i.e., endogenous variables) are left out so that the estimates are consistent, the percentage of explained variance is usually fairly low (10–15 percent).
we feel that it is most important that future research be directed toward improving the explanatory power of empirical wage equations and toward correcting the biases that result when these equations are used to predict wage rates for nonworkers. This will require richer data than are now available—for example, better measures of the quality of education, on-the-job training, and work experience. This is discussed in greater detail below.

**Nonemployment Income**

To estimate the potential behavioral effects of changing a family's income level through transfer payments it is necessary to examine exogenous variations in income that are independent of changes in the wage rate. In other words, an independent variable must be constructed from households' *nonemployment* income flows. One problem in constructing this variable correctly is that nonmonetized returns from such assets as one's home and automobile are the primary source of nonemployment income for most families, but since these returns cannot be directly observed, they are not easy to measure accurately. A more serious problem is that almost all nonemployment income is related to work decisions and, hence, is not truly exogenous.

Ignoring nonmonetized returns from assets could seriously bias estimated relationships between changes in income levels and changes in behavior. For example, if nonmonetized returns from assets are positively related to those that are observed, failure to take the unobserved flows into account will result in an overstatement of the effect of changes in income on household behavior. To incorporate unobserved income flows into the analysis, however, requires that they be imputed from available data on assets.

Greenberg suggests two different imputation approaches that might be used. (1) In the "accrual" approach, a dollar of imputed income from assets is regarded as equivalent to a dollar of currently monetized income receipts. Thus, this value is calculated as the product of an interest rate—a rate of return on assets—and the value of each household's equity in assets. (2) In the "cash flow" approach, current cash flows, rather than accruals, are emphasized. For example, one attempts
to measure the effect of owning a house, as opposed to renting one, on a household's cash-flow position.¹

Under either imputation approach, one faces the difficult problem of choosing the appropriate rate(s) of return to use, particularly since different assets vary in riskiness and liquidity. Moreover, some further adjustment may be necessary if some sources of nonemployment income are more permanent than others. Partial adjustments for this factor (and also for measurement error) might be made by using longitudinal data and averaging estimated nonemployment income flows over several years.

Even if one were able to correctly impute all nonobserved income flows from assets, however, the estimated relations between the resulting nonemployment income variable and household behavior may not accurately reflect the effects of changing income levels through an income maintenance program. The reason is that the change in family income resulting from the introduction of an income maintenance program will be an exogenous change, but almost any variations in household assets or nonwage income flows are endogenous to family decisionmaking. For example, the amount of transfer payments that families now receive from such programs as unemployment compensation, social security, welfare, and disability insurance are to a greater or lesser extent contingent on not engaging in market work. Including work-related income in the nonemployment income variable may produce a bias in the estimated relation between this variable and labor supply because some people may reduce their work effort in order to receive this type of income. Thus, work-related income should be excluded from the nonemployment income variable whenever possible.

For many families, assets represent a more important source of nonemployment income flows than do transfer payments. However, James

¹The underlying economic reason for distinguishing between income flows and accruals is that the capital market transactions necessary to transform accrued income into a smooth flow of current income may be sufficiently expensive to preclude such activity when a large number of relatively small transactions would be required. Thus, behavior may be more influenced by one's current cash-flow position than by one's total wealth.
Smith, in work that is still in progress, demonstrates that assets are also likely to be endogenous to most family decisionmaking. He reaches this conclusion on the basis of a model of asset accumulation over individuals' life cycles. The model implies that since assets are generated from past labor earnings,\(^1\) life cycle patterns of assets and labor supply are simultaneously determined by similar economic forces. Therefore, a substantial proportion of the cross-sectional variation in labor supply and asset levels may merely reflect one's position in the life cycle or variations in time preferences and lifetime patterns of market wages and may have nothing to do with income effects on labor supply decisions. Hence, observed statistical relations between current levels of assets and labor supply cannot necessarily be interpreted as evidence of a causal relationship running from income flows to market work.

Smith's main point can be illustrated with an example. If an individual who has neutral time preference\(^2\) and a constant wage rate over his life cycle faces a positive rate of interest, he will tend to consume less in the present and more in the future, because the positive interest rate makes future consumption less expensive relative to current consumption. As consumption rises throughout the life cycle, so will the derived demand for time in nonmarket production. Thus, at a given wage rate, this individual will work more in the labor market in the early periods of his life cycle than in later periods. He will also have positive savings and will accumulate assets in the early stages of his life cycle because his earnings exceed his consumption. In the

---

1 Assets are accumulated from savings, which are the difference between consumption in a given period and income in that period (wage earnings and interest from assets accumulated in previous periods).

2 Technically, an individual's time preference is measured by the slope of his indifference curve between present consumption and consumption in the future when equal amounts of consumption occur in both periods. If that slope—the marginal rate of substitution between present and future consumption—is -1, the individual has "neutral" time preference. If present consumption is on the horizontal axis and the slope of the indifference curve is less than -1 (i.e., greater than 1 in absolute value) at equal amounts of consumption in both periods, the individual has time preference for the present.
early part of his working life, therefore, this individual will have few assets and work relatively many hours; at a later stage in his lifetime, he will have more assets but work fewer hours. However, this negative relationship between asset accumulation and labor supply is not due to a causal influence of assets on labor supply but rather to life cycle variations in assets and labor supply that are both caused by the positive interest rate. ¹

The implication of Smith's model is that if assets are to be used in empirical analysis, they should be treated as endogenous and the sample used should be as free as possible of life cycle effects. To some extent this can be accomplished by defining subsamples to include individuals who are at the same point in their life cycles and have the same or similar life cycle wage paths, interest rates, and other factors that may affect their allocation of time over their life cycle. For example, age should provide a partial control for the point in the life cycle and education for the shape of the life cycle wage path. Alternatively, one can attempt to hold constant the effects of differing life cycle characteristics by including age and education as explanatory variables. In the final analysis, however, adequate treatment of this problem requires further development of the theory of life cycle asset accumulation.

¹ This same argument holds for individuals who have time preference for the present (see footnote 2, p. 38) so long as the absolute value of the marginal rate of substitution between present and future consumption (at equal amounts of consumption in each period) is less than the rate of interest plus one.

Another implication of the life cycle model is that those who place a relatively higher value on goods, particularly assets, than on non-market time may tend both to have larger than average asset holdings and to work above-average hours in order to accumulate even more assets. David H. Greenberg and Marvin Kosters, Income Guarantees and the Working Poor: The Effect of Income Maintenance Programs on the Hours of Work of Male Family Heads, The Rand Corporation, R-579-OEO, December 1970, attempted to control for such "preferences" for assets by including a preference control variable, which, as expected, had a positive sign and led to a more pronounced negative relation between labor supply and nonemployment income.
Labor Supply

In his EDA report, Greenberg discusses in some detail three of the more difficult issues in correctly constructing a labor supply measure from nonexperimental data: (1) whether it is conceptually more appropriate to measure labor supply as hours offered or as actual hours worked; (2) how to treat nonparticipants in the labor force, and (3) whether the two available dimensions of annual hours—weeks per year and hours per week—should be multiplied into a single product or treated as separate dependent variables.

Because of disequilibria in labor markets, the actual hours of employment for some workers may differ substantially from their "offered hours," the number of hours they wish to work at their received wage rate. For policy purposes, we wish to predict the changes in actual hours worked resulting from implementation of an income maintenance program. Paradoxically, however, it appears that these changes may be most accurately estimated from nonexperimental data if the dependent variable is measured in terms of the number of hours offered rather than the number actually worked. Workers are unlikely to reduce actual hours until the gap between hours worked and hours offered is closed, and income levels and wage rates may be correlated with the size of this gap. In practice it is difficult to measure offered hours for individuals whose actual labor supply falls short of that desired. Greenberg has suggested that the actual amount of time worked can be used as a minimum estimate of offered labor supply and that time worked plus time spent looking for work, which is available in some data sets, can be used as a maximum estimate.

To permit estimation of the number of persons who respond to an income maintenance program by completely withdrawing from the labor force, those whose actual labor supply equals zero—that is, nonworkers—should be included in the analysis. Schultz's results indicate that, at least for some demographic groups, failure to include these persons will result in understating the sensitivity of labor supply to wage and income changes. A major problem in including nonworkers is that wage rates are seldom directly available for them and must usually be imputed on the basis of wages earned by workers with similar
characteristics. However, as discussed above, there is reason to sus-
pect that workers' wages may systematically differ from the potential
wage rates for nonworkers.

The choice of the dimension of labor supply to analyze—participa-
tion vs. nonparticipation, weeks of work per year, hours of work per
week, hours of work per year—is often constrained by the available
data. In general, an analysis of the binary decision to work or not to
work seems less useful for purposes of examining income maintenance
programs than research that focuses on the quantity of labor supplied
as a continuous variable and assigns zero to nonparticipants.¹ Rela-
tionships explaining this continuous measure of labor supply should be
estimated by a technique appropriate for handling a dependent variable
that cannot fall below zero or is otherwise limited in range; one such
method, Tobit analysis, is discussed below (p. 54).

Various continuous measures of labor supply, such as weeks per year
and hours per week, may reflect different aspects of the labor supply
decision. For example, hours per week focuses on the decision between
part-time and full-time work, whereas weeks per year reflects the more
fundamental decision of whether to work at all during a given period.
If one wishes to analyze these two aspects of the labor supply decision,

¹However, some explanatory variables may have different effects on
the participation decision and on the numbers of hours supplied. For
example, an increase in the fixed cost of working (e.g., the cost of
going to work) may decrease the probability that an individual will
participate in the labor force (a substitution effect) but increase
the number of hours he will work if he does participate (he will work
more hours in order to recoup the higher cost—an income effect).

Michael J. Boskin ("The Economics of Labor Supply," Memorandum No.
110, Research Center in Economic Growth, Stanford University, September
1970), and Edward Kalachek and F. Raines ("Labor Supply of Low Income
Workers," in President's Commission on Income Maintenance, Technical
Studies, Washington, 1970) assume that decisions on how many hours to
work occur only if a decision has already been made to participate in
the labor force. They have attempted first to estimate the probability
that a given individual will be in the labor force and then to estimate
his hours of work conditional on his participation. However, unless
one believes that labor-supply decisions really do follow such a two-
stage process, it seems more natural to integrate the participation
and hours dimensions by measuring labor supply as a continuous variable,
assigning a zero value to nonparticipants.
it is desirable to treat weeks per year and hours per week as separate dependent variables, rather than multiplying them into a single dependent variable, hours per year. As Greenberg points out, it may also be advantageous to treat these two dimensions of labor supply separately because one of the two types of information may be unavailable or unreliable for some observations; by using weeks per year and hours per week as separate variables, these observations can be incorporated into at least part of the analysis. In addition, it may be possible to minimize the influence of biases that may result from spurious correlations between each of these measures and the observed wage rate (see p. 31).

Migration

As with labor supply, there are several different dimensions of migration that may be of interest. Many studies have been concerned only with outmigration—the number of persons leaving area i (regardless of destination)—and others have focused only on immigration (regardless of origin). Other studies, such as DaVanzo's, have examined place-to-place flows—the number or percentage of persons moving from i to j (for all i, j combinations, i ≠ j). We favor this latter approach for two reasons:

1. Decisions to migrate, as DaVanzo points out, are influenced by both origin and destination characteristics. People rarely leave places with no prior sense of where they intend to move. It is important that migration models for assessing the effects of an income maintenance program on geographic mobility use place-to-place flows because such a program will affect opportunities both at origin and at all possible destinations.

2. Place-to-place flows can always be aggregated into the total flow into or out of an area, ignoring origins or destinations; the reverse, however, is not possible. Furthermore, by using place-to-place flows, one can simultaneously investigate in- and outmigration.

In Section V we argue that the effects of income maintenance programs on household behavior are most appropriately examined with data on individuals. However, most migration studies have used aggregate
data. With aggregate data, nondestination-specific outmigration rates are measured as the percentage of persons originally in area i who at the end of the period under consideration are still alive, but no longer live in area i. Destination-specific outmigration rates, on the other hand, are measured as the percent of those originally in i who at the end of the period are still alive and live in area j, for each j ≠ i. For each origin i, the sum of these destination-specific rates over all possible destinations (that is, all j ≠ i) will equal the nondestination-specific outmigration rate.

Immigration rates for area i are usually calculated with aggregate data by dividing the number of immigrants to area i by the population in area i (either at the beginning or end of the period). Unlike outmigration, though, this rate is not a probability measure, since the denominator is not the population at risk to make the move.¹

Sometimes with aggregate data in- and outmigration flows are combined into a net migration figure (population change due to migration), but this measure cannot be expressed as any meaningful kind of rate. No person is a "net migrant": people can only arrive or depart. Thus, such a dependent variable cannot be used to predict the probability that a given type of person will migrate.

Very few studies of migration have, to our knowledge, used individual-level data, and these have looked only at nonorigin-specific immigration or nondestination-specific outmigration, or sometimes just the act of moving, regardless of both origin and destination. There are several reasons for this. First, most of these studies have been done by sociologists whose main interest often is in the personal characteristics (age, previous migration history), rather than the place characteristics (job opportunities, climate), that condition the propensity to move. Second, it has not been clear how a place-to-place migration variable should be defined from micro data.

¹Actually, it is all persons not in i at the beginning of the migration period who are at risk to move to i, but, to our knowledge, no one has ever constructed an immigration rate in this manner. Actually, this is a type of destination-specific outmigration rate that does not specify origin.
DaVanzo\textsuperscript{1} has suggested a technique for constructing a place-to-place migration variable from longitudinal data on individuals using dummy variables to enter relevant information. For example, if there are \( n \) geographic areas, there will be \( n-1 \) dummies for each individual in the sample at time \( t \). These dummies will indicate for the individual whether or not he moved to area \( j \) in the period \( t \) to \( t+1 \), for all \( j \neq i \) (where \( i \) is the area of residence at time \( t \)). If the individual does not move from \( i \), all the \( n-1 \) dummies will take on a zero value; but if he does move, one of the \( n-1 \) dummies will take on the value 1, indicating the area to which he moved. The rest of the dummies will equal zero indicating that although the individual moved somewhere, he did not move to the particular areas that these dummies represent. All of the place-to-place information is contained in these \( 0-1 \) variables. Each can then be regressed on characteristics of the individual, characteristics of place \( i \), and characteristics of place \( j \) (including imputed values for the potential migrant's earnings opportunities and employment probabilities at potential destination \( j \)). In this way, one can assess how characteristics of people and places jointly condition the propensity to move.

Definitions of the migration variable, such as the one suggested by DaVanzo, require that an individual's place of residence be observed at two points in time. Ideally, these two points should be relatively close together to minimize the possibility of his moving more than once or dying during the time period. To consider an extreme example, if the migration variable is calculated by comparing the individuals' birthplaces with their residency at retirement, migration that takes place during the working life of many persons may not be registered, since many persons choose to retire near their birthplace.

It is important that the geographic units used in an analysis of migration closely approximate those of policy interest, since different types of moves (for example, intracounty moves as compared with moves between Census divisions) may be differently affected by income

\textsuperscript{1}Julie DaVanzo, "Assessing the Impact of Human Resources Development Policies on Geographic Mobility," March 1973, mimeo.
maintenance programs. Estimates of the effect of an income maintenance program on migration among Census divisions, for example, will be of limited value to someone interested in the program's influence on rural-urban flows. For EDA purposes it seems most useful to focus on movement between each EDA region and other parts of the country, and on movements such as those between urban and rural areas within each EDA region.

Once the geographic units of analysis are defined, it is important to distinguish between persons who are moving to a particular geographic unit for the first time and those who have previously lived in the unit. Such a distinction will permit separate analysis of return and nonreturn migration, which is desirable since an income maintenance program can be expected to have differential effects on the two types of moves. For example, if individuals enjoy being near friends and relatives, income maintenance payments may subsidize return migration but not have a similar effect on nonreturn migration.

More difficult than defining the geographic units and the time period that should be used in the analysis is finding data that provide the necessary information. As indicated earlier, geographic location is frequently either not specified or is available only for areas so large that they are of limited value in most studies. Moreover, in some data sets, the only information on residency, other than that on where respondents currently live, is provided by retrospective questions dealing with their location at some earlier period, such as questions on where they lived at age 16 or where they were born. Longitudinal data that follow individuals over time are best for studying geographic mobility for they enable one to determine the effect of a family's

---

1For example, DaVanzo has found that long distance moves are more sensitive to the ability to finance an investment in migration than shorter moves.

2Existing longitudinal data will give some information on an individual's residence history but will not identify all places of previous residence.

3A more detailed discussion of data deficiencies appears in Section V.
characteristics at the beginning of a period on their propensity to migrate during that period.

ESTIMATION TECHNIQUES

De Tray's report emphasizes the interrelated nature of family decisionmaking: Over individuals' lifetimes their decisions on marriage, family size, labor force participation, investments in human capital, and so forth affect one another and are influenced by the characteristics and activities of other household members as well as by their own characteristics and activities. For example, fertility decisions and decisions on whether a wife works are frequently made jointly. A decision to work may also mean that a decision has been made to postpone child bearing or vice versa. Both decisions may depend on the hours worked and the wage rates of other family members, as well as on the wife's potential wage rate. Similarly, a decision that a particular family member participate in the labor force is likely to affect the family's decision to migrate and vice versa. Thus, the appropriate decision-making unit for welfare reform analysis is the family. To take account of the interrelatedness of family decisionmaking and, hence, to generate consistent estimates of the effects of welfare reform on family behavior requires use of the appropriate econometric techniques.

Multiple Regression Analysis

Multiple regression analysis should be used to estimate the relationships of interest. This technique enables one to estimate the direction, 

---

1 Since much of this subsection is inherently more technical than the rest of the report, the less technically inclined reader may wish to skim or skip it.

magnitude, and statistical significance of the association between a
dependent variable and an explanatory variable, while holding constant
the influence of other factors that also affect the dependent variable.
Regression analysis is preferable to examining two-way cross tabulations,
which permit one to derive only separate bivariate relationships between
the dependent variable and each of the explanatory variables. Regres-
sion analysis enables one to determine the joint effects of a set of
explanatory variables on the dependent variable. These joint effects
are not equal to the sum of the separately estimated bivariate rela-
tionships, except under special restrictive conditions. In addition,
certain types of multiple regression analyses enable one to decompose
the explanatory variable's total effect into direct and indirect effects.
For example, the number of years of schooling an individual has completed
is expected to be positively related to his wage rate, which in turn
influences his likelihood of participating in the labor force; this is
an indirect effect of education on labor force participation. However,
education may also exert a direct effect on his labor force participa-
tion decision by influencing his "taste" for work in the labor force as
compared with his taste for alternative activities.

Simultaneous-Equations Estimation Techniques

The spheres of family behavior in which we are interested are inter-
related and jointly determined. Ordinary least squares (OLS), the most
frequently used multiple regression estimation technique, yields un-
biased\(^1\) estimates of the parameters in a single equation only when all
of the explanatory variables are exogenous—that is, when they are
uncorrelated with the stochastic disturbance term. If any of the ex-
planatory variables are correlated with the true error term—that is,
endogenous—OLS estimators of the parameters will be biased and inco-
sistent.\(^2\) Whether a variable is exogenous or endogenous is, in part,

\(^1\) An estimator is unbiased if its expected value is equal to the
true value of the parameter it is estimating.

\(^2\) An estimator is consistent if the probability limit of the esti-
mator (the value at which the distribution becomes concentrated as the
sample size increases) is equal to the true value of the parameter it is estimating.
a function of the time period under consideration. In the short run the family must take such factors as wage rate and family size as given—that is, exogenous. However, over the long run, these variables are subject to choice and hence endogenous.

In general, if an explanatory variable, X, is endogenous and is negatively (positively) affected by the dependent variable, Y, the disturbance term in the equation explaining Y will be negatively (positively) related to X and the OLS coefficient of X will be algebraically biased downward (upward). The coefficient will be biased because OLS attributes the entire relationship between X and Y to the effect of X on Y and ignores the causal relationship that runs from Y to X. For example, if X = number of children and Y = female labor force participation rate, an OLS equation estimating female labor force participation rates would tend to produce a negative coefficient on number of children that was biased away from zero.

There are several techniques appropriate for estimating equations that are part of a jointly determined system. Two-stage least squares (2SLS), a type of instrumental-variables estimation technique, is one such method. It is a single-equation technique that can be used to estimate one structural equation from a jointly determined system, even when the rest of the system has not been specified. As its name implies, estimation takes place in two stages. In the first stage, each endogenous explanatory variable is regressed on all of the predetermined variables (variables thought to be statistically independent of the disturbances). This first stage generates estimated values of the explanatory endogenous variables, known as "reduced-form estimates".

Instrumental-variables estimates are also used to avoid the bias toward zero that arises when there are errors in the measurement of exogenous variables (see Johnston, *Econometric Methods*, pp. 195-198).

or "instrumental variables estimates," that, unlike the endogenous variables in their original form, are not correlated with the disturbance term. In the second stage, the dependent variable is regressed on the reduced-form estimates of the explanatory endogenous variables and whichever predetermined variables are specified to directly affect the dependent variable. Since the first stage should purge the explanatory endogenous variables of their correlation with the error term, 2SLS estimators are consistent, although their variances are larger than those of corresponding OLS estimators. OLS estimators have a smaller dispersion than 2SLS, but it is around an asymptotically biased mean. 1

Other techniques for estimating jointly determined relationships are not as frequently used as 2SLS or (the inappropriate) OLS, partly because they involve much higher computational expenses and partly because they are less well known. Limited information maximum likelihood (LIML), like 2SLS, is a single-equation technique. LIML estimators have the same asymptotic properties as 2SLS, but LIML has an important advantage over 2SLS in that it treats all endogenous variables symmetrically during estimation. Therefore, it is not necessary to choose one endogenous variable arbitrarily at the beginning to be the left-hand (dependent) variable as it is in 2SLS. 2

Full-information techniques, such as three-stage least squares (3SLS) and full information maximum likelihood (FIML), can be used only when the entire system of relationships is specified. These techniques use information on the entire system to estimate each equation; all of the parameters are estimated simultaneously. Full-information estimators

---

1 In "Monte Carlo" studies of small sample properties, 2SLS estimators usually perform about as well as OLS in terms of efficiency (dispersion around the true parameter value) even though 2SLS estimators generally have a greater dispersion around their mean than OLS estimators (Johnston, Econometric Methods, Chapter 10; Goldberger, Econometric Theory, p. 362). One estimator, \( \hat{\delta} \), is more efficient than another, \( \tilde{\delta} \), if the mean square error (second moment) of \( \hat{\delta} \) is smaller than that of \( \tilde{\delta} \)---i.e., \( E(\hat{\delta} - \delta)^2 < E(\tilde{\delta} - \delta)^2 \).

2 Each possible left-hand variable will yield different 2SLS parameter estimates because residual variance is being minimized in a different direction.
are usually more efficient\(^1\) than those yielded by limited-information techniques, but are also more sensitive to specification error because parameter estimates for all of the equations are affected by specification error in any one equation.

Three-stage least squares applies generalized (Aitken's) least squares to the whole set of relations using unbiased estimates of the variance-covariance matrix based on 2SLS estimates. This technique has the same asymptotic properties\(^2\) as FIML\(^3\) but is much simpler to compute\(^4\) and therefore is more frequently used. However, 3SLS, like 2SLS, requires that the left-hand variable for each equation be chosen at the beginning of estimation, whereas FIML, like LIML, allows all included endogenous variables in each equation to be treated symmetrically during estimation. Full-information techniques have been used less frequently than limited-information (single-equation) techniques because (1) researchers are often interested in the estimated relations in only one equation and therefore do not specify the entire system; (2) it is often impossible to avoid specification error, to which full-information techniques, especially FIML, are very sensitive; and (3) full-information techniques are less well known, and computer programs to perform them are often not available.

---

1 See footnote 1, p. 49, for a definition of efficiency. Full-information techniques will yield more efficient estimates than limited information techniques whenever the disturbances from various equations in the system are correlated.

2 The asymptotic properties of an estimator refer to its limiting distribution as the sample size goes to infinity.

3 However, when a priori information on the variance-covariance matrix of the disturbances is available, FIML estimators are asymptotically more efficient than 3SLS.

4 There is a linearized version of FIML, known as "linearized maximum likelihood," which is simpler computationally than FIML, though still not as simple as 3SLS. Linearized maximum likelihood has the same asymptotic variance-covariance matrix of the disturbances as FIML when the disturbances are normally distributed (Goldberger, Econometric Theory, p. 363).
Qualitative Dependent Variables

Because the behavioral decisions likely to be affected by welfare reform are made in a family context, welfare reform effects are best estimated with data that are reported separately for each individual and household. Aggregate data do not enable us to match family members; and, in addition, they force us to summarize an entire distribution of, say, wages, with a mean or median, thus losing all variation within each group observation. At the individual level, however, many behavioral phenomena are dichotomous; for example, a person either migrates (variable = 1) or he does not (variable = 0). Standard statistical techniques, such as OLS, are usually inappropriate for estimating relationships where the dependent variable is binary, since treating a dichotomous dependent variable as if it were an ordinary continuous variable violates several of the assumptions that enable OLS regression analysis to yield best linear unbiased estimators. For example, a two-valued dependent variable is associated with a two-valued error term; thus the disturbance does not have constant variance, but is, instead, heteroscedastic. Heteroscedasticity will cause OLS coefficients estimates to be inefficient. Moreover, the probability distribution of a dichotomous variable is bounded between zero and one, but a linear regression may yield estimated values that lie outside this range. In addition, the relationship determining a dichotomous variable is usually not linear and, therefore, should not be fitted by a linear regression.

---

1 A qualitative variable is one whose outcomes are categorical and not necessarily numerical. Occupation, marital status, and labor force status are examples of qualitative variables. If the variable has two possible outcomes, such as labor force status, it is known as a "dichotomous," "binary," or "0-1" variable. If there are more than two possible categories, as in the case of marital status or occupation, the variable is "polytomous."

Some of the techniques discussed here for estimating a relationship determining a qualitative dependent variable are examined in greater detail in the forthcoming report by Nerlove and Press, which is summarized in the appendix of this report.

2 A best linear unbiased estimator ("BLUE") is one that has smaller variance than any other linear unbiased estimator.

3 Constant variance of the error term, which is known as "heteroscedasticity," is one of the assumptions that enables OLS estimators to have smaller variance (be more efficient) than any other linear unbiased estimator.
Many of the problems associated with dichotomous dependent variables persist even when individual-level data are grouped to form rates—for example, the percentage who participated in the labor force—or when existing aggregate data are analyzed.\(^1\) A linear regression may still yield estimated values of the dependent variable that lie outside the range of possible values, for example, outside the range 0–100 in the case of labor force participation rates. If many of the observations are concentrated near one of the bounds, the relationship is likely to be curvilinear and incorrectly approximated by a linear equation. Furthermore, although the error term will no longer have only two possible values, it will still be heteroscedastic because the observations are based on groups of different sizes and the variance is systematically related to the value of the dependent variable;\(^2\) however, this heteroscedasticity can be corrected by using weighted regressions.\(^3\)

A nonlinear estimation method known as "logit" can be used to provide estimators with desirable statistical properties for relationships determining dichotomous dependent variables. Logit model coefficients can be estimated by maximum likelihood, minimum chi-square, or minimum logit chi-square techniques. All three estimation methods yield consistent, asymptotically efficient, and asymptotically normally distributed estimators.\(^4\) The first two techniques can be used on ungrouped

\(^1\)Furthermore, as noted above, information on the explanatory variables may be lost through the aggregation process.

\(^2\)The variance of a proportion, \(\hat{P}_i\), is \([\hat{P}_i(1 - \hat{P}_i)]/N_i\), where \(N_i\) is the denominator of the proportion (that is, the size of the group to which it refers). Therefore, in an equation explaining \(\hat{P}_i\), \(\hat{P}_i\)'s variance, and thus the variance of the error term, will be systematically related to the value of \(\hat{P}_i\).

\(^3\)Each observation, \(\hat{P}_i\), is weighted by the inverse of an estimate of its standard deviation; i.e., \(\sqrt{N_i/[\hat{P}_i(1 - \hat{P}_i)]}\), where \(\hat{P}_i\) is the predicted value of \(P_i\) from a first-stage OLS regression. This is equivalent to weighting the moment matrix by a matrix whose diagonal elements are \(N_i/[\hat{P}_i(1 - \hat{P}_i)]\).

\(^4\)Technically these are RBAN estimators—regular best asymptotic normal (see Dhrymes, *Econometrics*, p. 115, for a definition of regularity).
data, but they usually require an iterative approach and, thus, involve relatively complicated computational procedures. The minimum logit chi-square technique, on the other hand, uses grouped data and estimates the parameters directly using generalized least squares.\(^1\) However minimum logit chi-square usually requires a larger sample than the other two techniques.\(^2\)

Another method often used to create an unbounded variable from a 0-1 variable is the probit transformation. The probit transformation, like the logit, is a monotonic transformation that is usually applied to relative frequencies for grouped data. It is the inverse of the cumulative distribution function of a standardized normal variate.\(^3\)

---

\(^1\) For grouped data the "logit" is defined as the natural logarithm of the odds ratio for a cell; that is, \(\ln([P_1(y)]/[1 - P_1(y)])\), where \(\ln = \) natural logarithm and \(P_1(y) = \) relative frequency of occurrence of event \(y\) in cell \(i\) (used as an unbiased estimate of the group mean probability of occurrence). The cells are usually defined according to values of the independent variables. When the logit is regressed on the explanatory variables, \(\ln([P_1(y)]/[1 - P_1(y)]) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots\), \(P_1(y)\) is a logistic function of the \(X_i\)s:

\[
P_1(y) = 1/[1 + e^{-\left(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots\right)}].
\]

In estimation, each observation is weighted by the square root of the inverse of its variance (see footnote 2 below).

\(^2\) An additional limitation of the minimum logit chi-square method of estimating the logit model results from the fact that the moment matrix is inverted by the inverse of the variance of the (logit) observations to correct for heteroscedasticity. Thus, the weight for each cell \(i\) is \([N_i \cdot P_1(y) \cdot (1 - P_1(y))]\), where \(N_i\) is the number of observations contained in cell \(i\). The problem with such a weighting scheme is that it deletes cells in which all observations on the dependent variable are alike, for when the estimated probability is either zero or one, the variance is zero. Thus, cells with only one observation, which in some cases could be most or all of the cells, must be excluded from the analysis, as must all cells with outcomes that are all alike.

\(^3\) If a variable \(y\) is distributed normally with mean \(\mu\) and standard deviation \(\sigma\), the standardized variable \(Y = (y - \mu)/\sigma\) is distributed normally with mean 0 and standard deviation 1 and is known as a "standardized normal variate." If \(F\) is the cumulative distribution function of a standardized normal variate and \(P_1(y)\) is the probability of occurrence, the probit, \(Z\), is defined implicitly as \(F(Z_i) = P_1(y)\), or
To estimate the coefficients in a probit model, one maximizes a logarithmic likelihood function and obtains nonlinear equations that are then solved by an iterative procedure.

For proportions between .05 and .95 in grouped data, the cumulative normal and logistic functions are very similar and, hence, probit and logit results are usually nearly the same; since logit is simpler computationally than probit, it is usually preferred for proportions in this range. However, when the observations are concentrated near one of the extremes, that is, near 0 or 1, probit and logit estimators do vary considerably, and probit is often preferred because it is based on the normal distribution.¹

Probit analysis can be extended to study situations where the dependent variable is limited in value only at one end of its range; for example, hours of work cannot fall below zero. This extension was proposed by Tobin² and has come to be known as the "Tobit" technique. Tobit is often applied to situations where researchers are concerned

\[ Z_i = F^{-1}(P_1(y)). \]

The Central Limit Theorem is used in biological application to justify the choice of the normal distribution; in bioassay, where one is interested in the critical dosage of a drug that kills a group of rats, a number of independent factors determine the critical level for each rat. A similar justification of the normality assumption applies to economic and demographic phenomena. When the distribution of critical levels is normal, the conditional probability of \( y \) given the values of the explanatory variables is

\[ P_1(y|X_{11}, X_{21}, ...) = F\left( \frac{\beta_0 + \beta_1 X_{11} + \beta_2 X_{21} + \ldots}{\sqrt{2\pi}} \right) \]

\[ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2} \, dt \]

and

\[ F^{-1}(P_1(y|X_{11}, X_{21}, \ldots)) = \beta_0 + \beta_1 X_{11} + \beta_2 X_{21} + \ldots. \]


with the determinants of the amount spent, if any, on a certain good and could be used to study the effect of income maintenance programs on consumption expenditures. Schultz uses both Tobit and OLS to estimate labor supply equations and finds that the more appropriate Tobit approach may produce quite different results from OLS.

In their forthcoming report, Marc Nerlove and S. James Press discuss a maximum likelihood logit technique for estimating relationships where the dependent variable is qualitative. This technique is related to generalized contingency table analysis and can be used to estimate relationships where some of the explanatory variables are endogenous. It is a simultaneous-equations technique for estimating relationships determining jointly dependent qualitative variables. This method of estimation can be used in studies where the dependent variable is polytomous, that is, represents events with more than two possible outcomes. For example, a polytomous variable could be used to denote different marital statuses (single, legally married, consensually married, separated, divorced, widowed) or different types of migration (intracounty mover, intercounty-intrastate migrant, migrant between contiguous states, and so forth). Polytomous variables may also result from a cross-classification of several binary variables, which may not be independent. For example, individuals may be classified according to whether or not they are participating in the labor force, and each of these two groups may be further classified by their marital status. The maximum likelihood polytomous logit model can be used to examine possible interdependencies among such classifications and to study the effect of exogenous variables on the different outcomes and on the interdependencies.

1For example, Tobit has been used to estimate the demand for health insurance (C. E. Phelps, The Demand for Health Insurance: A Theoretical and Empirical Investigation, R-1054-OEO, The Rand Corporation, forthcoming).

2However, the computational procedure becomes very complicated when there are many, say more than 10, possible categories. The computational procedure for maximum likelihood estimation of equations where the dependent variable is dichotomous has been programmed at Rand. A program for the analysis of polytomous dependent variables has been developed and is currently being tested at Rand.
Some phenomena that can be represented by a polytomous variable have an implicit ordering, such as type of migration or, perhaps more obviously, number of children. Ordering often gives one some prior information on the pattern of coefficients. Existing estimation procedures, however, including the maximum likelihood technique, do not enable one to incorporate these "priors" into the analysis. The polytomous technique may expose interesting nonlinearities that would be concealed if the ordered variable, such as the number of children, were denoted by a single discrete step variable. For example, in preliminary examination at Rand of the polytomous dependent variable estimation technique with a sample of married women in the Philippines, a positive relationship was found between education and number of children for low values of number of children, but a negative relationship was found for large numbers of children. In other words, women with little education were more likely to be childless than more highly educated women; but if they did have children, they were more likely to have large families.

**CHOICE OF THE SAMPLE AND SUBSAMPLES**

This subsection discusses two related estimation issues: (1) selecting the observations to include in the analysis and those, if any, to omit, and (2) stratifying the main sample into relatively homogeneous subgroups for which separate regressions can be estimated. The objective is to select samples allowing estimation of the behavioral parameters necessary to predict the potential effects of alternative income maintenance programs. However, as the rest of this subsection attests, this objective is very difficult to obtain.

**Choice of the Sample**

As Greenberg notes, variations in the sample population have been a major, if not the most important, source of inconsistency among cross-sectional studies of labor supply. Parameter estimates for other

---

1 Much of this subsection draws on Section III of Greenberg's report.

2 For an example of the divergent results from different labor supply studies, see p. 28.
types of behavior likely to be affected by welfare reform are undoubtedly also very sensitive to sample composition.

In estimating behavior parameters, different researchers have suggested various reasons to justify their particular choice of sample population. Some researchers, for example, have attempted to restrict their sample to those who are most likely to become participants in income maintenance programs. Samples are also, of course, limited to those for whom there is sufficient information to construct the variables. Some of the criteria that have been used to select a sample apply only to studies of particular types of behavior; for example, labor supply studies often exclude those with constraints on the number of hours they work or those who are using their time outside the labor market to engage in certain productive activities, such as schooling.

Greenberg argues that although many of the reasons researchers give for excluding certain observations from their sample are valid under certain circumstances, exclusion usually causes at least some bias in the results. One illustration of this is the bias that results when nonworkers are left out of the sample because it is not possible to make a direct observation of their potential market wage rates. Thus, as many of the obstacles to including certain observations in the sample should be removed as is feasible. Whenever obstacles cannot be surmounted, very careful judgment must be exercised in determining which observations should be included and which excluded. At the very least, the sensitivity of the estimates to excluding observations should be tested whenever possible.

There is sometimes no choice but to exclude observations with insufficient information. Nevertheless, whenever possible, variables should be redefined so that the missing information is not needed or missing values of explanatory variables should be imputed from values for persons of similar characteristics for whom the information is given. Unfortunately, while excluding observations with missing values may cause bias, so may using redefined or imputed variables. This is attested to by the problems involved in imputing wage rates discussed earlier. If feasible, these techniques should be used as sensitivity checks on one another.
It is sometimes argued that for a study of the effects of an income maintenance program on household behavior, the sample should exclude persons unlikely to be in the recipient population. This argument is especially forceful if there is reason to believe that the values of behavioral parameters differ for recipients and nonrecipients. Then nonrecipients must be excluded to generate unbiased estimates of recipient responses. Thus, one might want to exclude from the sample those individuals and families whose incomes and wage rates will not be directly affected by the program under consideration.

However, one must be careful that the variables used to distinguish the qualifying population are exogenous. Different programs, of course, would usually imply different samples. For example, under the Family Assistance Plan, only individuals or couples with a child 18 years old or younger (or a child 19-21 and attending school) were to be eligible to receive benefits. However, selecting a sample population on the basis of family structure is somewhat questionable since such programs may, at least over the long run, induce some of the poor to modify their family structure.

One must be even more careful about using an income cut-off to select the sample, since this can introduce serious biases. For example, stratification by relatively narrow income intervals will introduce a spurious negative relation between hours of work and both wage rates and nonemployment income.

Many of those who present particular estimation problems or seem to differ greatly from those who make up the bulk of the sample (and are therefore most likely to be excluded) may also be those who are most likely to participate in an income maintenance program and to exhibit strong behavioral responses to it. For example, persons who report involuntary unemployment, particularly recipients of unemployment

\footnote{DaVanzo has shown, for example, that inclusion of members of the armed forces in the sample used for studies of migration can seriously bias the results.}

\footnote{Briefly, the spurious correlation results because for individuals to remain within a narrow income interval, any variation in their hours must be offset by an opposite movement in their earnings or nonemployment income or in both (Greenberg).}
compensation, are sometimes excluded from the sample in labor supply studies;\footnote{1} but such persons are disproportionately represented among those who work a small number of hours, those who receive low wages, and those who are likely to take part in an income maintenance program. Preferably, such persons should not be excluded from the sample. Whenever possible, separate regressions should be computed for each non-homogeneous subgroup. The problem with implementing this approach is that it is sometimes difficult to distinguish appropriately among subgroups. For example, one cannot differentiate between those who receive certain types of transfer payments (such as unemployment compensation) because they are unable to work and those who reduce their work effort in order to qualify for these benefits. However, there is good reason to suspect that behavioral parameters differ for these two groups.

**Defining the Subsamples**

Whenever there is reason to believe that the underlying behavioral parameters of the model differ among various subgroups in the population—that is, the characteristics defining the subgroups interact with the explanatory variables of the model—one should stratify the main sample into subsamples so that separate regressions can then be estimated for each subgroup. However, one cannot always know \textit{a priori} whether or not certain subgroups actually are significantly different. Especially important in this connection, as indicated in Section I, is the issue of whether there is statistically significant regional variation in behavioral response parameters. An F-test that can be used to explore this issue will be discussed in some detail later in this subsection. Briefly, to conduct the test, one must estimate separate regressions for each region as well as a regression for a sample where observations for all the regions are pooled. The sums of the squared residuals from these separate and pooled regressions are then used to compute an F-ratio to test for coefficient homogeneity.

\footnote{1See Greenberg, pp. 20-22, for a discussion of the problems involved in treating these persons in labor supply studies.}
In defining subsamples, one must be careful that the variables used are exogenous and free of measurement error. Moreover, one must balance the advantage of allowing the parameters to vary for each subgroup against the loss of degrees of freedom and precision of estimates that results when there are many subsamples and each is relatively small. The F-test mentioned above can be used to determine the appropriate level of disaggregation.

Variables that might be used to define subsamples include sex, race, and age—all clearly exogenous. Behavioral parameters for those in the various categories defined by these characteristics are likely to differ because these persons are subjected to different experiences and institutional pressures. In addition, stratifying by age enables one to adjust for life cycle effects. As discussed earlier, Smith shows that very serious biases can arise when the life cycle component of variation in behavior is overlooked.

Other variables that are often used to define subsamples are education, marital status, and family size, although over the long run none of these is truly exogenous to the behavioral decisions of interest. For example, decisions on how many years to remain in school or how many hours to work after completing school may be made simultaneously.

One more variable of obvious importance is geographic location. If behavioral parameters vary geographically, separate regressions should be estimated for each region of interest. However, statistical tests should first be performed to determine whether or not the parameters do indeed differ significantly by region. It may be true that much of the regional variation in the dependent variables is due to differences in the demographic composition of the regions and to differences in the levels of the explanatory variables, rather than in responsiveness to changes in the explanatory variables. If the statistical tests indicate that behavioral parameters differ significantly by demographic groups but not by region, national estimates of the parameters for separate demographic groups, together with the demographic breakdown of the population by region (obtained, perhaps, from the 1970 Census of Population) would be sufficient to obtain estimates of the regional effects of a particular program; that is, it would be unnecessary to estimate a
separate set of parameters for each region of interest, thus saving
degrees of freedom and improving the precision of the estimates. More-
over, if existing data for estimating behavioral parameters could not
be stratified to correspond to the geographic regions of interest, it
would be unnecessary to incur the substantial cost of collecting new
data for each region.

As indicated earlier, an F-test can be used to determine whether
behavioral parameters vary significantly by region or by demographic
group. More precisely, the test indicates whether significant explana-
tory power is lost by pooling; it is a test of coefficient homogeneity
among a set of subsamples.\textsuperscript{1} If the test reveals that the coefficient
differences for a set of subgroups are not statistically significant,
those subgroups should be pooled for greater precision. If the F-test
rejects the null hypothesis of coefficient homogeneity, the source of
the nonhomogeneity should be sought. For example, if separate regres-
sions are estimated for several regions and the F-test indicates that
there is a significant difference between the equations, the specific
coefficients causing the difference should then be identified. For
example, a dummy for each region can be included in a pooled regres-
sion, and the F-test can then be repeated to determine whether coeffi-
cients other than the intercept differ significantly by region. If
the F-test no longer rejects the null hypothesis of coefficient homog-
eneity, only the intercept differs significantly with region. In this
case, degrees of freedom can be conserved and the precision of the esti-
mates increased by including the regional dummies in a pooled regres-
sion, rather than by estimating a separate regression for each region.
However, if the F-test still rejects the null hypothesis that the other
coefficients are the same, the regional dummies should be included in
interaction with different explanatory variables to determine if they
are the source of coefficient nonhomogeneity. If only one or just a
few coefficients are found to differ significantly with region,

\textsuperscript{1}The technical details of this test are described in Appendix D of DaVanzo's report.
precision will be gained by estimating a pooled regression that includes interaction terms of those variables and the regional dummies.\footnote{DaVanzo estimated separate regressions explaining migration for all sex-color-education groups of persons aged 25-29 and then performed F-tests to determine the appropriate level of disaggregation. She found that the coefficients did vary significantly by education, but usually not by color (color differences were significant only for the two college-educated groups), and never by sex. Further testing to explore the source of coefficient variation by race revealed that only the intercept and the coefficient of the income ratio differed significantly with race.}

Once the level of disaggregation has been decided, the technique for estimating "seemingly unrelated" regressions\footnote{Arnold Zellner, "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias," \textit{Journal of the American Statistical Association}, 57, June 1962, 348-368.} can be used to estimate the regressions for all subgroups together, thereby generating more efficient estimates than result from OLS. This method is a form of generalized least squares and is somewhat analogous to three-stage least squares.\footnote{In this technique, separate regressions are first estimated for each subgroup by OLS. The residuals from these equations are used to form the estimated variance-covariance matrix, which is then used to create generalized least-squares estimators. With regressions on different subgroups, the disturbances in one equation are likely to be correlated with the disturbances in another. Whenever disturbances are not independent, generalized least squares can be used to increase efficiency.} However, because, like 3SLS, the parameter estimates for all equations are affected by specification error in one equation, these estimators are sensitive to specification error.
V. EXISTING NONEXPERIMENTAL DATA SETS

In this section we briefly describe the potential usefulness of five data sets for studies of family behavior: the Survey of Economic Opportunity (SEO), the Income Dynamics Panel (IDP), the National Longitudinal Surveys (NLS), the Census Bureau's Public Use Samples from the 1960 and 1970 Census (PUB), and the Social Security Continuous Work History Sample (CWH).\footnote{Another data set that should be mentioned is the 1960 Survey of Consumer Expenditures. This Bureau of Labor Statistics survey might be useful in determining the potential effect of alternative income maintenance programs on expenditure patterns. The 1960 Survey is somewhat out of date, but a 1972 Survey of Consumer Expenditures is expected to be available in 1974. In addition, the Current Population Surveys provide up-to-date data for the study of specific types of behavior, for example, migration.} All of these files have been or are being used in Rand research.

All five of these data sets provide information at the individual or household level. This is also the level at which behavioral decisions affected by welfare reform are made, so aggregate data on larger units such as counties or states are not nearly as useful as those on households for predicting the effects of income maintenance programs. The major deficiencies of aggregate data for studying household behavior were pointed out in Section IV: (1) It is not usually possible with aggregate data to take account of the effect of the characteristics and activities of one family member on the behavior of another. (2) Since aggregate data are reported as means and medians, information is provided only about a single point on a distribution; hence, much pertinent information about the forces influencing household behavior may be lost.

THE BASIC SAMPLES

The SEO sample consists of over 27,000 household units. Many of the data, such as those pertaining to assets, liabilities, and income from nonemployment sources, are collected for the household as a whole.
Information pertaining to an individual, such as that on hours worked, earnings, and educational attainment, is generally reported separately for each member of the household at least 14 years old. Thus, the SEO may be used to obtain behavioral estimates for all adults in the household. (On the other hand, coverage by the IDP and NLS is limited to only certain members of the household.) The basic SEO sample consists of two parts. The first, about 60 percent of the total, is a national self-weighting sample. The second is drawn from areas with large concentrations of nonwhites. The reason for this supplementary sample is to increase the reliability of estimates of the characteristics of the poor nonwhite population, a group that is typically underrepresented in regular census enumerations.

The IDP is a five-year longitudinal survey conducted by the University of Michigan's Survey Research Center. At its inception, the sample consisted of 4,802 families. The focus is on the head of the household; information on the head's spouse is much more limited, and only very summary data on any additional family members are reported. Like the SEO, the IDP sample is composed of two parts. About 60 percent of the total sample is a representative cross-sectional sample of the United States. The remainder consists of households also interviewed for the 1967 SEO. In 1967, these households had heads who were under 60 years of age and incomes that did not exceed a value equal to roughly twice the federal poverty line.

The NLS is sometimes called the "Parnes" data after the director of the study, Herbert S. Parnes of the Center for Human Resource Research at Ohio State University. It is composed of five-year longitudinal surveys of four subsets of the United States population: men 45 to 59 years of age, women 30 to 44, and young men and women 14 to 24. At the study's inception, approximately 5,000 interviews—with about 3,500 whites and 1,500 blacks—were obtained for each of the four surveys. Although the survey instruments that were used for each of the four population groups are similar in overall design and many identical questions were asked, they do differ sufficiently that information especially relevant to each of the four groups could be obtained. The older males, for example, were asked about their retirement plans, and
the younger men and women were questioned in considerable detail on their educational experiences and plans. Only very summary data were obtained on other family members.

The PUB consists of data for individuals and households from the 1960 and 1970 U.S. Censuses. The primary sample is one household for every 100 households in the national population—slightly over 2 million persons in 1970—and is self-weighting. Information is provided about each housing unit and each person within the sampled household.

The CWH file is a continuous, 13-year work history of one percent of all persons covered by Social Security. The employee's date of birth, race, and sex are obtained from forms filled out when he first applied for a social security number. All other data are obtained from reports submitted to the Social Security Administration by employers. The file pertains only to individuals. No information is available on other members of the household.

THE TIME FRAME

The SEO was conducted in the spring of 1966 and again in the spring of 1967. However, it was not intended to be a longitudinal survey. The interviewers visited the same set of addresses in both years, rather than following the same set of persons. Nevertheless, more than 75 percent of the households interviewed in 1966 were re-interviewed in 1967. Weekly hours and hourly wage rate information were obtained only during the second survey year.

The IDP is a longitudinal panel survey with interviews taking place annually between March and June. The survey was initiated in 1968 and ended in 1972. The initial sample of interviewed families diminished by slightly more than one-tenth during the first survey year and by about one-seventh during the first two survey years.

The NLS is also a five-year longitudinal panel survey. Each of the four population groups is being surveyed annually. The initial surveys of the two male groups were conducted during October-December 1966, the older groups of women during mid-1967, and the younger group of women in early 1968. The young male sample, which would be
expected to diminish the fastest, decreased about one-tenth by the end of the first year and about one-sixth during the first two years of the survey. Nearly three-fifths of those lost had entered the armed forces. By way of contrast, only six percent of the initial older male sample was lost during the first survey year.

The PUB has no longitudinal dimensions. It was conducted on the basis of "long-form" Census questionnaires during the two most recent decennial censuses. Matching individuals across two censuses is not possible.

The CWH has been generated by the Social Security Administration since 1957 and is currently available through 1969. The file was constructed by selecting the records of individuals possessing certain combinations of digits in their social security numbers. Sampling of the same combination of digits year after year has assured continuity of the file—that is, retention of the same individuals in the sample as long as they continue to earn income covered by the Social Security System. However, no information is provided when persons are in noncovered employment or out of the labor force. In fact, there is no way of telling why a particular individual dropped out of the file.

INFORMATIONAL CONTENT

In Section IV, we discussed many of the problems encountered in appropriately measuring the variables one must use to estimate behavioral parameters. Table 1 lists the information included in these five data sets that can be used to construct the more important of these measures. To keep the table manageable, the NLS column pertains only to 14 to 24-year-old males. The information provided on the other population groups that were surveyed for the NLS is similar, except that the nonemployment income and assets of the two older groups are reported in somewhat greater detail, and the education data in less detail. The IDP and NLS columns are, with a few exceptions, applicable only to the third survey year. The questions asked

1 The most important exception concerns information on assets provided by the NLS. The asset questions were asked during the first
Table
SELECTED INFORMATION AVAILABLE IN THE SEO, IDP, NLS, PUB, AND CWH

<table>
<thead>
<tr>
<th>SEO</th>
<th>IDP</th>
<th>NLS</th>
<th>PUB</th>
<th>CWH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earnings last week</td>
<td>Hourly rate for regular work time on main job</td>
<td>Usual earnings on present or most recent job</td>
<td>Earnings last year</td>
<td>Annual and quarterly earnings that are taxable under the Social Security System</td>
</tr>
<tr>
<td>Earnings last year</td>
<td>Hourly rate for overtime on main job</td>
<td>Rate at which paid for overtime (i.e., time and one-half, double time, etc.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EARNINGS</td>
<td>Hourly rate for work on secondary job(s)</td>
<td>If unemployed or out of labor force and thinking about reentering</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Expected hourly wage if unemployed or out of labor force and thinking about reentering</td>
<td>Wages last year</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Money income from labor</td>
<td>Cost of travel to work last year</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cost of travel to work</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NON-EMPLOYMENT INCOME FLOWS</th>
<th>Rental income receipts</th>
<th>Asset part of farm income</th>
<th>Unemployment compensation</th>
<th>Social Security</th>
<th>No information available</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Interest and dividend income</td>
<td>Asset part of business income</td>
<td>Financial assistance from relatives</td>
<td>Welfare</td>
<td>All other non-earned income</td>
</tr>
<tr>
<td></td>
<td>Regular payments from life insurance annuities, royalties, trust funds, etc.</td>
<td>Asset part of income from roomers and boarders</td>
<td>Other income such as rental income, interest, dividends, disability payments, etc.</td>
<td>Total income of all household members</td>
<td>Total income of all household members</td>
</tr>
<tr>
<td></td>
<td>Regular contributions by persons outside the household</td>
<td>Rent, interest and dividends</td>
<td>Total income of others in family</td>
<td>Social Security</td>
<td>Retirement pay, pensions and activities</td>
</tr>
<tr>
<td></td>
<td>One-time-only money receipts (e.g., net profit from the sale of assets; lump sum payments from insurance policies, pensions, funds, etc.)</td>
<td>Spouse's income from assets</td>
<td>Total income of others in family</td>
<td></td>
<td>Unemployment or workmen's compensation</td>
</tr>
<tr>
<td></td>
<td>Social Security or Railroad Retirement Pensions</td>
<td>Welfare</td>
<td></td>
<td>Total income of others in family</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pensions from retirement programs for government employees or military personnel</td>
<td>Income from other sources</td>
<td></td>
<td>Total income of others in family</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pensions from private employers</td>
<td>Spouse's transfer income</td>
<td></td>
<td>Mortgage or rent payments last year</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Workmen's compensation, illness or accident benefits</td>
<td>Total transfer income of others in family</td>
<td></td>
<td>One-time-only money receipt (e.g., inheritance, insurance settlement)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Veterans' disability or compensation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unemployment insurance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### ASSETS

<table>
<thead>
<tr>
<th>Description</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market value of own home</td>
<td>i</td>
</tr>
<tr>
<td>Debt owed on own home</td>
<td>j</td>
</tr>
<tr>
<td>Market value of real estate holdings</td>
<td>k</td>
</tr>
<tr>
<td>Debt owed on real estate holdings</td>
<td>k</td>
</tr>
<tr>
<td>Market value of vehicles</td>
<td>k</td>
</tr>
<tr>
<td>Debt owed on vehicles</td>
<td>k</td>
</tr>
<tr>
<td>Market value of business and/or farm</td>
<td>k</td>
</tr>
<tr>
<td>Debts owed by business and/or farm</td>
<td>k</td>
</tr>
<tr>
<td>Money in interest bearing accounts</td>
<td>k</td>
</tr>
<tr>
<td>Money in checking accounts</td>
<td>k</td>
</tr>
<tr>
<td>Value of bonds, government bonds, stocks, etc.</td>
<td>k</td>
</tr>
<tr>
<td>Personal loans to others</td>
<td>k</td>
</tr>
<tr>
<td>Value of other personal property (excluding clothing, furniture, etc.)</td>
<td>k</td>
</tr>
<tr>
<td>Debts owed stores, banks and lending institutions</td>
<td>k</td>
</tr>
<tr>
<td>Non-interest debts (e.g., debts owed doctors, hospitals, other people, utility companies, etc.)</td>
<td>k</td>
</tr>
</tbody>
</table>

### LABOR SUPPLY

<table>
<thead>
<tr>
<th>Description</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hours worked last week</td>
<td>i,d</td>
</tr>
<tr>
<td>Weeks worked last year</td>
<td>i,d</td>
</tr>
<tr>
<td>Weeks unemployed last year</td>
<td>i,d</td>
</tr>
<tr>
<td>Whether usually works full-time or part-time</td>
<td>i,d</td>
</tr>
<tr>
<td>Whether part-time work voluntary or involuntary</td>
<td>i,d</td>
</tr>
<tr>
<td>The reason why if out of the labor force for at least part of last year</td>
<td>i,d</td>
</tr>
<tr>
<td>Average hours worked per week last year</td>
<td>i,p</td>
</tr>
<tr>
<td>Hours usually worked on present or (for those who have not been without work for more than a year) most recent job</td>
<td>i</td>
</tr>
<tr>
<td>Weeks worked last year</td>
<td>i,e</td>
</tr>
<tr>
<td>Weeks unemployed last year</td>
<td>i,e</td>
</tr>
<tr>
<td>Weeks of missed work because of illness</td>
<td>i,e</td>
</tr>
<tr>
<td>Whether it would have been possible to have worked more (or less) last year</td>
<td>i</td>
</tr>
<tr>
<td>Whether interviewee would have preferred working more (or less) last year</td>
<td>i</td>
</tr>
<tr>
<td>Weeks of vacation taken last year</td>
<td>i,e</td>
</tr>
<tr>
<td>Hours worked last week</td>
<td>i,d</td>
</tr>
<tr>
<td>Weeks worked last year</td>
<td>i,d</td>
</tr>
<tr>
<td>Whether working five years ago</td>
<td>i,d</td>
</tr>
<tr>
<td>Hours usually worked, during the weeks worked during the past 12 months</td>
<td>i</td>
</tr>
<tr>
<td>Weeks unemployed during past 12 months</td>
<td>i</td>
</tr>
<tr>
<td>The reasons why if out of labor force for at least part of last 12 months</td>
<td>i</td>
</tr>
<tr>
<td>If currently unemployed,</td>
<td></td>
</tr>
</tbody>
</table>
Market value of vehicles

Debt owed on vehicles

Market value of business and/or farm

Debts owed by business and/or farm

Money in interest bearing accounts

Money in checking accounts

Value of bonds, government bonds, stocks, etc.

Personal loans to others

Value of other personal property (excluding clothing, furniture, etc.)

Debts owed stores, banks and lending institutions

Non-interest debts (e.g., debts owed doctors, hospitals, other people, utility companies, etc.)

Debt owed on real estate, firm or business

Value of auto

Debt owed on auto

Other debts

---

LABOR SUPPLY

Hours worked last week

Weekly worked last year

Weeks unemployed last year

Whether usually works full-time or part-time

Whether part-time work voluntary or involuntary

The reason why if out of the labor force for at least part of last year

Average hours worked per week last year when working

Weeks worked last year

Weeks unemployed last year

Weeks of missed work because of illness

Whether it would have been possible to have worked more (or less) last year

Whether interviewee would have preferred working more (or less) last year

Weeks of vacation taken last year

Hours worked last week

Hours usually worked on present or (for those who have not been without work for more than a year) most recent job

Weeks worked during past 12 months

Hours usually worked, during the weeks worked during the past 12 months

Weeks unemployed during past 12 months

The reasons why if out of labor force for at least part of the last 12 months

If currently unemployed, hours per week interviewee would be willing to work

Whether part-time work is voluntary or involuntary

Hours worked last week

Weeks worked last year

Year when last worked

Whether working five years ago

No information available
<table>
<thead>
<tr>
<th>ALLOCATION OF NON-MARKET TIME</th>
<th>EDUCATION</th>
<th>RESIDENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whether attending school</td>
<td>Highest grade completed</td>
<td>Name of Census Region</td>
</tr>
<tr>
<td>Hours of housework done by wife, husband, other family members</td>
<td>Highest grade attended</td>
<td>For households in the 12 largest SMSAs, the SMSA is identified</td>
</tr>
<tr>
<td>Expenditures on housework help, childcare, restaurants</td>
<td>Number of years completed</td>
<td>Length of time in present area</td>
</tr>
<tr>
<td>Value of self repairs on car</td>
<td>Kinds of degrees</td>
<td>Name of Census Division</td>
</tr>
<tr>
<td>Value of self repairs on home</td>
<td>Date last enrolled in school</td>
<td>SMSAs or contingent counties having a population of over 250,000 are identified</td>
</tr>
<tr>
<td>Hours watching TV per weekend</td>
<td>Highest grade completed</td>
<td>State of residence</td>
</tr>
<tr>
<td>Most important two things done in spare time</td>
<td>Highest grade attended</td>
<td>County where employer is located</td>
</tr>
</tbody>
</table>

### Table 1 (Continued)

<table>
<thead>
<tr>
<th>SELECTED INFORMATION AVAILABLE IN THE SEO, IDP, NLS, PUB, AND CWH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SEO</strong></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Whether attending school</td>
</tr>
<tr>
<td>Hours of housework done by wife, husband, other family members</td>
</tr>
<tr>
<td>Expenditures on housework help, childcare, restaurants</td>
</tr>
<tr>
<td>Value of self repairs on car</td>
</tr>
<tr>
<td>Value of self repairs on home</td>
</tr>
<tr>
<td>Hours watching TV per weekend</td>
</tr>
<tr>
<td>Most important two things done in spare time</td>
</tr>
<tr>
<td>Highest grade completed</td>
</tr>
<tr>
<td>Highest grade attended</td>
</tr>
<tr>
<td>Number of children currently in school</td>
</tr>
<tr>
<td>Head's expectations for children's education</td>
</tr>
<tr>
<td>Number of years completed</td>
</tr>
<tr>
<td>Kinds of degrees</td>
</tr>
<tr>
<td>Date last enrolled in school</td>
</tr>
<tr>
<td>If currently enrolled, grade attending and type of curriculum</td>
</tr>
<tr>
<td>Amount of scholarship or other financial aid</td>
</tr>
<tr>
<td>Location of present and former schools</td>
</tr>
<tr>
<td>If out of school for a year or more, year out and reason for returning</td>
</tr>
<tr>
<td>Plans for future education</td>
</tr>
<tr>
<td>Highest grade attended by other family members</td>
</tr>
<tr>
<td>Name of Census Division</td>
</tr>
<tr>
<td>SMSAs or contingent counties having a population of over 250,000 are identified</td>
</tr>
<tr>
<td>State of residence</td>
</tr>
<tr>
<td>County where employer is located</td>
</tr>
<tr>
<td>Retro-</td>
</tr>
<tr>
<td>prospective</td>
</tr>
<tr>
<td>Subjective</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>FAMILY STRUCTURE</td>
</tr>
<tr>
<td>FERTILITY</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

- c Refers to data available for 14 to 24 year old males. Information available for other population groups is similar. Unless otherwise indicated the information is applicable to the respondent only.
- d Reported only for the sampled individual; only persons covered by social security are included.
- e Available for 1967 only.
- f Reported separately for each adult (persons 14 years or older) in family unit.
- g Applicable only to family head.
- h Applicable only if paid on hourly wage basis.
- i Information is also provided on how many hours must first be worked before overtime begins.
- j Applicable only if working at time of survey.
- k Applicable only to family head and spouse.
- l Includes regular wage income, income from overtime, commissions, income from professional practice or trade; and estimates by the Survey Research Center of that part of total reported income from farms, unincorporated businesses, and roomers and boarders that should be assigned to labor rather than to capital.
<table>
<thead>
<tr>
<th>FAMILY STRUCTURE</th>
<th>FERTILITY</th>
<th>SEX AND MARITAL STATUS OF HEAD</th>
<th>RELATIONSHIP OF EACH FAMILY MEMBER TO RESPONDENT IN HOUSEHOLD TO HEAD</th>
<th>NO INFORMATION AVAILABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relationship of each person in household to head of household</td>
<td>Number of children ever born to head</td>
<td>Number of children in family unit under 25</td>
<td>Age of each person in household</td>
<td>No information available</td>
</tr>
<tr>
<td>Age of each person in the household</td>
<td>Number of children in family unit under 17</td>
<td>Age of youngest child</td>
<td>Age of each person in the household</td>
<td>No information available</td>
</tr>
<tr>
<td>Number of children ever born</td>
<td>Number of dependent children</td>
<td>If an additional child is desired, the length of time</td>
<td>Number of children ever born</td>
<td>No information available</td>
</tr>
<tr>
<td>Ages of children c, s</td>
<td>Number of all family members</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average time between births</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of additional children expected</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of children expected in next 5 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

aRefers to data available for 14 to 24 year old males. Information available for other population groups is similar. Unless otherwise indicated the information is applicable to the respondent only.
bReported only for the sampled individual; only persons covered by social security are included.
cAvailable for 1967 only.
dReported separately for each adult (persons 14 years or older) in family unit.
eApplicable only to family head.
fApplicable only if paid on hourly wage basis.
gInformation is also provided on how many hours must first be worked before overtime begins.
hApplicable only if working at time of survey.
iApplicable only to family head and spouse.
jIncludes regular wage income, income from overtime, commissions, income from professional practice or trade; and estimates by the Survey Research Center of that part of total reported income from farms, unincorporated businesses, and roomers and boarders that should be assigned to labor rather than to capital.
kReported for the family unit as a whole.
lnEstimates by the Survey Research Center of that part of the total income from this source attributable to capital rather than labor.
moReported separately for main job and secondary job(s) for family heads who were employed at the time of the interview.
qThe SEO, IDP, NLS and PUB all provide considerable descriptive information about a household's geographic location (rural, urban, suburban; size of SMSA). Only data that pertain to the identity rather than the characteristics of the geographic location are listed.
rtThe United States is divided into 4 Census Regions and 9 Census Divisions.
sReported for all adult women.
during the earlier two years were similar to those asked during the third; nevertheless, a few variables listed in Table 1 are not available for the first or second survey years. Identical information was obtained for each year covered by the CWH.

A BRIEF COMPARISON

A researcher's choice of data obviously depends on the issues he is studying. For certain types of studies each of the data files examined in this section has comparative advantages. In many ways, however, the SEO is the most versatile of the five data sets and provides a useful basis of comparison. For example, the nonemployment income and asset data in the SEO are probably better than those provided by any of the other four data files; and SEO data on hours, earnings, family structure, and fertility are overwhelmingly superior to those in the PUB and the CWH (see Table 1). Since only the SEO and PUB obtain answers to their full set of questions from all members of the household, these two surveys are probably the most appropriate of the five for studies that are set in a family context. Moreover, these and the IDP are the only data sets that can be used to obtain behavioral estimates for all population groups.

Although the SEO does have important advantages over each of the other four data sets listed in Table 1, it is not the superior data source for all potential income maintenance issues. The IDP and the NLS, for example, provide important information on interviewees that is not available in the SEO. In particular, they offer better measured hours and wage variables than the SEO. This is especially true of the NLS,¹ although this data set is limited to four age-sex groups.

¹For example, participants in the NLS were asked what their usual wages and hours are or were on their present or most recent job. Information on hours worked per week and on hourly wages for those interviewed for the IDP is based on the employment experience over the entire year preceding the interview. There was also an attempt in the IDP to collect hours and wage information on the self-employed. Furthermore, participants in both the NLS and IDP were asked a series of questions providing somewhat greater information than the questions in
Two major limitations of the SEO are that it provides very little longitudinal information and it does not allow the sample to be broken into small geographic units. Therefore, its usefulness for migration studies is severely limited. Moreover, the limited geographic detail means that only restricted testing for geographic variation in behavioral parameters is permitted.

The best data for migration studies seem to be those that follow individuals over time, for these data offer the best opportunities for obtaining necessary information on economic conditions at each stop along the migratory path. Furthermore, no matter what dimension of behavior is being examined, longitudinal information may allow some control over transitory and life-cycle influences and facilitate the study of the timing of the effects of income maintenance programs. ¹

Of the three longitudinal data sets we are considering, the CWH covers the longest period of time. However, although the IDP and NLS cover only five years (the CWH covers 13) and do not provide nearly as large a sample or as much geographic detail as the CWH, they offer much more information on each included observation. Although the CWH may be somewhat useful for examining migration, it cannot be used at all in estimating labor supply, fertility, and most other behavioral parameters.

Data from the PUB seem clearly inferior to the SEO, IDP, or NLS for estimating behavioral parameters. A measure of the hourly wage rate, for example, can be obtained only by multiplying hours during the week before the interview by weeks worked during the year before the interview and then dividing the product into earnings during the SEO provided on hours constraints. Some information was also obtained on the reservation wages corresponding to hours offered but not actually worked. The necessarily hypothetical nature of these questions, however, leaves their answers highly suspect.

¹There are special statistical techniques that should be used when the data are a time series of cross-sections—that is, when one pools cross-sectional data from several time periods. See, for example, Marc Nerlove, "Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross-Sections," Econometrica, March 1971. Nerlove and Schultz, Love and Life Between the Censuses, have used this technique to estimate equations explaining family behavior in Puerto Rico.
year before the interview. Because of its vast size, however, the PUB provides probably the only data base that can be satisfactorily used to test for possible differences in behavioral parameters across relatively small geographic areas. As Table 1 indicates, the number of observations contained in these files permits relatively small geographic units to be identified. Possibly even more important, the PUB is potentially very useful for simulating the effects of alternative income maintenance programs in particular geographic areas once the basic behavioral parameters have been estimated from other data sources.
VI. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

SOME BASIC PRINCIPLES FOR WELFARE REFORM RESEARCH

Several main themes underlie Rand's research on the behavioral effects of welfare reform. These do not pertain specifically to assessing regional effects; rather, they represent important principles that should govern future research on the effects of income maintenance legislation whatever the geographic unit of interest.

1. The appropriate decisionmaking unit for analyzing welfare reform effects is the family; the appropriate data to use are data on households and their members (De Tray, DaVanzo, Greenberg, Schultz, Smith). Most individuals who will be affected by welfare reform do not live alone and do not make their decisions regarding the number of hours they participate in the labor market, whether or not they migrate, and so forth in isolation from other members of their families. These decisions depend not only on the characteristics of the individuals under consideration but also on the characteristics and activities of other household members. Ignoring other family members in the analysis may result in seriously biased estimates of the potential effects of welfare reform. Because welfare reform should be analyzed within a family context, aggregate data on larger units, say counties or states, are not appropriate for estimating the behavioral effects of income maintenance legislation, for aggregate data do not enable one to match family members. Data on individuals and households, on the other hand, do permit determination of how a family's behavioral decisions are influenced by the characteristics of each of its members.

2. Researchers must be careful not to confuse life cycle and cross-sectional effects; longitudinal data should be used to help control for life cycle effects (Smith). To attribute observed differences across individuals to differences in current conditions alone may lead to biased estimates of the true response to the exogenous changes that will be brought about by the introduction of an income maintenance program. At a single point in time, a substantial portion of the variation across households in such activities as consumption,
labor force participation, and asset accumulation may be due to these households being at different points in their life cycles and to their having different lifetime patterns of consumption, labor force participation, and asset accumulation. One should attempt, whenever possible, to control in empirical analysis for changes in behavior related to the life cycle by (1) including age and other time-related variables in the regressions (or stratifying subsamples by these variables), (2) using simultaneous-equations estimation techniques when explanatory variables are endogenously determined over the life cycle, and (3) using longitudinal data, which allow one to observe the same individual at several points in his life cycle. (Longitudinal data also permit one to follow individuals over time and, thus, analyze the timing of their decisions.)

3. Available experimental and nonexperimental data each have their advantages and should be used as checks on each other. We have argued above (points 1 and 2) that longitudinal data on individuals and their households should be used to study welfare reform effects. The two main sources of such information are nonexperimental data generated through survey interviews or Census questionnaires and data gathered during social experiments. This report has been mainly concerned with problems of obtaining usable results from the former. The major comparative advantage of the latter, however, is clear: Experimental data permit analysis of the behavioral responses to conditions as close as possible to those that would exist under a given actual program. Nevertheless, results from nonexperimental data provide an important check on and supplement to those from experiments. Results from experimental data may be subject to certain biases (see pp. 26–27) that nonexperimental results are not. Moreover, compared with obtaining results from nonexperimental data, particularly from existing data files, experiments are likely to be expensive, to involve long time delays before producing usable results, and to be conducted in narrow geographic areas. In fact, one of the more important ways in which nonexperimental data can be used to supplement experiments is in determining the applicability of experimental results for one area to all geographic regions of interest. This can be accomplished by using
nonexperimental data to test whether behavioral parameters vary geographically.

4. Researchers should be aware of the biases associated with alternative approaches—constructions of variables and model specifications—and should, whenever possible, use several different approaches to generate reasonable bounds for the behavioral parameters of interest (Greenberg). Greenberg has provided numerous examples of the implications of making one methodological choice rather than another in formulating and estimating models of labor supply and has illustrated how alternative approaches can be used to bound a parameter estimate.

5. Short-run parameters should not be used to predict longer-range effects. The distinction between long-run and short-run effects has often been overlooked in studies of family behavior. The implicit assumptions underlying most previous studies of welfare reform indicate that their results apply mainly to the shorter-run effects of income maintenance legislation. To estimate these effects one can assume that many household characteristics, such as family size and composition, are predetermined. However, in the longer run many of these household characteristics are subject to choice and are therefore likely to be affected by welfare reform; as the time horizon lengthens, more variables become endogenous to household decisionmaking.

6. Models should be formulated to account for the simultaneous nature of household decisionmaking, and appropriate econometric techniques should be used to estimate them (De Tray, Schultz). The spheres of family behavior that will be affected by welfare reform are interrelated and jointly determined. When relationships are jointly determined, special statistical methods—simultaneous equations estimation techniques—are required to generate estimators with desirable properties. Use of inappropriate econometric techniques may result in misleading policy implications.

7. Proper econometric techniques must be used to estimate relationships determining qualitative (dichotomous, polytomous) variables or dependent variables that are otherwise limited in range (Nerlove and Press, Schultz). Often the behavioral phenomena one
wishes to explain can be represented by a variable that has only a small number of possible values (for example, a 0-1 variable indicating whether or not a family migrated) or a variable otherwise limited in range (for example, one that cannot fall below zero, such as hours of work). The most commonly used regression technique, ordinary least squares, is inappropriate for estimating relationships determining such variables; therefore, special estimation techniques are required. For binary dependent variables, appropriate estimation techniques include logit and probit analysis; logit can also be used to estimate relationships determining polytomous variables. For dependent variables limited at one end of their range, Tobit analysis should be used. These techniques can be used on individual-level or aggregate data. Moreover, one type of logit analysis has been extended by Nerlove and Press under Rand's EDA grant for use in simultaneous equations systems.

8. Unless one econometric approach to a particular problem is clearly superior a priori to the others, appropriate alternative estimation techniques should be tried and the sensitivity of estimates to the choice of technique should be determined. In this report we have discussed a number of statistical techniques. Usually several methods are available to treat a particular problem. For example, two-stage least squares and limited information maximum likelihood are both single-equation/simultaneous-equations estimation techniques; logit and probit analysis are alternative methods for estimating a relationship determining a dichotomous variable.

9. Statistical tests of coefficient homogeneity should be performed to determine the appropriate level of subsample composition (DaVanzo). Subsamples should be as homogeneous as possible; that is, the underlying behavioral parameters should be the same for all members of each subsample. If there is reason to believe that the underlying parameters differ among geographic areas or among age, racial, or other demographic groups, separate regressions should be estimated for each subgroup. A F-test will determine whether the explanatory power gained by estimating separate regressions for each subgroup is sufficient to compensate for the loss in degrees of freedom and precision of estimates.
RECOMMENDED TOPICS FOR FUTURE RESEARCH

As discussed in the introduction to this report, economists are not yet able to provide dependable estimates of the potential regional and national effects of welfare reform. The range of uncertainty is still unacceptably large. In this report we have identified some of the major methodological issues and problems that we feel must be resolved if the range of uncertainty is to be satisfactorily reduced.\footnote{It is interesting to note that only a few years ago some of these (for example, the appropriate definitions of the wage and nonemployment income variables) were not even perceived to be problems.} We believe that the state of the art is sufficiently advanced that with adequate effort these problems are solvable; in fact, progress has already been made on most of them at Rand and elsewhere.

The following list identifies those areas on which we feel future research should concentrate. Although only the first recommendation deals exclusively with assessing regional effects, the others also identify research areas where further development is required before the potential regional effects of welfare reform can be adequately predicted.

1. We have mentioned several possible sources of regional differences in program effects: (1) the provisions of an income maintenance program may vary geographically; (2) the proportion of persons eligible to receive benefits may not be the same in each area; (3) the demographic compositions of regional populations differ; different groups may respond differently to an income maintenance program (see point 9 above); and (4) behavioral response parameters, even for a given demographic group, may vary geographically. We recommend that future research assess the relative importances of the possible sources of regional variation in program effects. Statistical tests should be performed to determine whether or not behavioral response parameters differ significantly by region. If they do differ significantly, it will be necessary to estimate separate parameters for each region; if not, a single set of parameters based on a national sample will suffice for each demographic group. Moreover, social experiments that are conducted in one region should be applicable to other regions.
2. Over the life cycle, assets—the primary source of nonemployment income flows for many families—are endogenous to most family decisionmaking. This causes an important problem in estimating the income effects associated with an income maintenance program because life-cycle patterns of assets and, say, labor supply are simultaneously determined by similar economic forces and are not necessarily causally related (see pp. 38-39). We recommend further work on theoretical and empirical models of asset determination so that this variable can be correctly used to derive accurate estimates of income effects. Assets should be treated as endogenous in estimation, and the sample used should be as free as possible from life cycle effects.

3. Although the arguments for using an imputed wage are convincing—they can correct for transitory influences and measurement error, avoid simultaneity bias, help control for life-cycle influences, and measure wages for nonworkers—actual efforts to do so have proved less than satisfactory. The explanatory power of estimated wage equations is often fairly low. Moreover, as the discussion of selectivity biases illustrates, the difficult problem must still be resolved of adjusting for the biases that arise when wage rates for nonworkers are imputed on the basis of the wages of workers with similar characteristics. We recommend an effort to improve the methodology and performances of imputed wage rates.

4. By providing a higher floor under household income than was previously available, an income maintenance program may increase the expected values of the return to investments and reduce the variance of possible outcomes. The latter implies a decrease in the "risk" of investing, encouraging "risk averse" people to make investments. Perhaps the most important avenue of investment open to poor people is investment in their own human capital—education, job training, job search, migration, and so forth. We recommend that risk aversion be incorporated into the models used to study welfare reform effects so that the effect of an income maintenance program on a family's propensity to invest, particularly in its own human capital, can be determined.
One reason progress on the topics mentioned above has not been more rapid is that available data sets do not contain all the information researchers need. Some of this information is very difficult to obtain. However, several types of information that have not usually been available can relatively easily be incorporated into data sets that are generated in the future.

Sufficient information on geographic location is seldom provided by existing data sets. Frequently only large, heterogeneous areas, such as the four Census regions or nine Census divisions, are identified. Information on smaller, more homogeneous areas is required to analyze interregional migration and interregional variation in program parameters. Simulation of program effects within a given region also requires, of course, that the region be identified.

Household income flows and equity in assets tend to be endogenous, making it very difficult to use existing asset and income information to estimate income effects. Although considerable additional theoretical work is necessary before there is very precise knowledge of the kinds of data required to overcome this problem, one type of information that would be helpful is on inheritance, for this income (at least its timing) should be nearly independent of work decisions. However, such data have seldom been collected. It would also be useful if more information were collected on how individual households have varied their consumption and savings patterns over time.

To estimate substitution effects one needs a measure of the wage rates on which individuals' behavioral decisions are based—the amount of goods and services that can be exchanged for an hour of work. Wage information provided in recent data sets, such as the NLS, represent a substantial improvement over earlier available data. Nevertheless, additional information is needed on such subjects as fringe benefits, overtime wages, wages obtained on second jobs, tax rates, and variations in costs of living. Moreover, to improve imputed wage estimates, better information is needed for both workers and nonworkers on such wage determinants as the quality of schooling and the quantity, quality, and timing of all work experience, both during school and since leaving school.
APPENDIX

Summaries of Rand Reports Prepared for EDA
A CONCEPTUAL BASIS FOR THE STUDY OF
WELFARE REFORM EFFECTS

R-1066-DOC, Dennis N. De Tray, September 1972

This study develops and explores a framework in which the potential regional effects of welfare reform can be assessed. It does not, however, deal directly with either inter- or intraregional effects but rather argues for a methodology that should underlie an analysis of these effects. Previous research efforts in this area have tended to concentrate on labor supply effects of welfare reform, usually at the national level. In the approach outlined here, emphasis is on the nature of influences on family behavior of such legislation as the Family Assistance Plan and the need for a general framework within which these related effects may be analyzed.

A review of the legislative history of the different Family Assistance Plans (FAP) illustrates the complexity of these plans and their potential for affecting all aspects of family behavior. Four of the five forms of the bill considered have come before Congress. These are the basic FAP bill, HR 16311 of April 1970, two subsequent revisions of that bill (June 1970 and October 1970), and H.R. 1 of June 1971. The last alternative, the so-called Ribicoff Amendment, is included for the sake of comparison.

Proponents of FAP and related plans that incorporate negative income tax principles argue that current welfare programs have a number of undesirable characteristics. In particular, existing welfare programs apply only to a certain subgroup of the population, creating pressure for nonrecipients to "restyle" their lives so as to move into the recipient class. A negative income tax plan mitigates this shortcoming to some extent in that all those classified as poor—that is, below some given income level—are eligible for payments under the program. A negative income tax also partly removes the undesirable incentives for families to break up (or look as if they have broken up) and reduces the work disincentives of existing welfare programs.
In reality, FAP programs currently under legislative review fall short of these goals. The legislation now contains distortions that are similar to those of the existing welfare system.

To discuss the consequences of various forms of FAP on the allocation of time of household members, the number of factors to be considered must first be reduced to a manageable level and the language of legislators translated into economic concepts. This report groups the numerous provisions of FAP into six classifications—the extent of coverage as determined by the subpopulation eligible for FAP payments, the basic support level, the implicit marginal tax rate on earnings, exemptions to the income and wealth calculations, registration and work requirements, and other important provisions.

Major criticized features of the proposed legislation are (1) in the tradition of existing welfare programs, single individuals and childless couples are not covered, (2) recipient families can still gain considerably if they look as if they have broken up, (3) the actual tax rate on earned income is often very high, ranging in excess of 90 percent in some cases, (4) the tax rate on unearned income is even higher (100 percent), providing little incentive for those eligible for FAP payments to apply for pensions or unemployment compensation, to maintain rental properties, and so on.

In analyzing welfare reform one must also be concerned with the political feasibility of alternative FAP bills. A review of the debate, both public and Congressional, on the pros and cons of FAP legislation leads to the conclusion that the three major constraints on the bill are the cost of the program, the percent of the population on welfare, and the effect of the program on work incentives.

Evaluation of an income maintenance program's potentially important side effects on marriage behavior, fertility, schooling, migration, savings, and even labor supply requires an analytical framework that identifies the important links among these phenomena in families' decisionmaking processes. One such framework, known as the household production model, is proving a useful source of specific hypotheses concerning families' responses to changes in their surroundings.
In this model, households are assumed to derive satisfaction not directly from purchased market goods and services but from more basic objects of consumption called "household commodities." These commodities are not exchanged in the market place, and their production in the household requires not only market goods and services, but "home time" as well (that is, time of one or more members of the household).

Description of the model begins with a discussion of the economic incentives for marriage. The household production model points to several reasons that might account for the fact that most people marry: First, marriage permits the household to take better advantage of gains from specialization; second, marriage reduces the severity of the time constraint within which household production must take place. In turn, these reasons suggest several factors that affect the household's allocation of time. In general, this theory argues that the relative market wages of the husband and wife and their relative home productivities determine time allocation within the family. Analysis of households' time allocation decisions are significantly complicated, however, by the fact that families can alter the value of a member's time by increasing the amount of human capital invested in that individual, or by allowing existing human capital to depreciate.

As in the more traditional labor-leisure model of time allocation, another factor affecting family decisions is the level of household income. In this model, however, income refers not to monetary (market) earnings of household members, but to a broader "full wealth" concept: the lifetime sum total of all household resources, including the value of all time whether or not that time is spent in the market place working. The distinction between monetary income and full wealth is that monetary earnings will depend on the amount of time that family members choose to allocate to market work over their life cycle, but full wealth will be independent of that choice.

The broad applicability of this approach can be demonstrated in four areas that are likely to be affected by welfare reform. Recent work at Rand on migration as a human capital investment has emphasized that the decision to migrate is often made in a family rather than an individual context and therefore that the wife's employment opportunities
enter her husband's migration decision and vice versa. Decisions on
nonmarket production in general, and fertility in particular, are likely
to be affected by welfare reform. Even though some of the major criti-
cisms of existing welfare systems fall within this area, little has been
written and less hard evidence accumulated on this topic. For some
recipients, income maintenance programs will reduce the variance asso-
ciated with a given expected level of future income, which may in turn
affect some of the investment decisions (in both human and non-human
forms of capital) that families make. Finally, the implications of the
household production model underscore the importance of life-cycle
effects and the potential usefulness of simultaneous-equation estimation
techniques in the study of labor supply behavior.

One cannot possibly hope to assess regional effects of welfare
reform until one can, with some degree of accuracy, estimate the effects
of those reforms on individual families. In the model of household
behavior outlined above, emphasis is on two features of our society
that may influence the choices individuals make when faced by changes
in their economic environment. These are, first, the notion that as
a general empirical rule most individuals do not live alone but in multi-
person (husband-wife) households, and, second, that the term "leisure"
is inadequate to describe the diverse set of activities individuals
engage in when not working in the market.

Economic studies of the effects of welfare reform (or, in fact,
of household behavior in general) should recognize these characteristics
and their implications for parameter estimation. Not recognizing that
most men are married, for example, will lead to a misspecification of the
male labor supply function and to the possibility of biased response
estimators. The same holds a fortiori for estimates of the female
labor supply function. Wives work less in the market than their husbands
not simply because their wages are low but also because husbands' wages
are usually higher than wives' wages.

The concept of "nonmarket production" as it applies in the house-
hold production model may also have important consequences for welfare
reform analysis. The effect of welfare reform on activities other
than market work—for example, migration and education—may be both
predictable and substantial, especially over the longer run. The more immediate regional effect of an income maintenance program may in turn depend on the types of nonmarket activities that families are engaged in when the plan is enacted. An obvious example is the effect of the presence of small children in the home on the responsiveness of wives to changes in their market wages. Since the mix of these nonmarket activities may vary considerably among regions, the initial response to the program may also differ systematically from one area to another.
PROBLEMS OF MODEL SPECIFICATION AND MEASUREMENT:
THE LABOR SUPPLY FUNCTION

R-1085-EDA, David Greenberg, December 1972

For the past several years, there has apparently been a public consensus that the nation's existing welfare system is no longer viable. A continuing debate has gone on, however, on the dimensions of an income maintenance program (IMP) to replace it. Inability to choose among IMP alternatives is partly (though significantly) caused by uncertainty over the potential effect that various proposed programs will have on labor supply. The extent to which a given program alters the number of hours worked by participants will ultimately determine the program's cost, its effectiveness in reducing poverty, and its national and regional effects.

Income maintenance programs may influence the work effort of participants in two ways. First, IMPs pay a subsidy to participating families. Although workers can use this subsidy to purchase more market goods, they could also use it to replace earnings given up by choosing to work fewer hours. Second, the total subsidy a family receives is usually contingent upon the family's earnings. Thus, an IMP in effect imposes a tax on earnings, which produces a change in the amount of market goods and services that can be obtained in return for an hour's work. Presumably, this could produce adjustments in the number of hours worked. Estimates of the labor supply response to both of these program dimensions are necessary in performing an empirical analysis of the effect of alternative IMPs on labor supply.

One way to obtain such estimates is through inferential studies based on cross-sectional data. These studies use existing interfamily variations in wages and income, particularly income not directly received as the result of employment, to predict families' labor supply responses to changes in their wage rates and income. It is assumed that by comparing individuals, systematic relationships between work choices and wage rates and between work choices and income levels will be obtained. These relationships can then be inferred for a single typical worker.
To do this, other differences between individuals—such as family structure, education, tastes, and geographic location—must be controlled for.

So far, studies that have used cross-sectional data have failed to make sufficient reductions in the range of uncertainty about the labor supply effects of alternative IMPs to be very helpful to policymakers; the studies' conclusions have been disturbingly varied. Improving the usefulness of such studies depends on defining the sample population, resolving conceptual and operational problems in measuring the principal variables—labor supply, wage rates, and nonemployment income—and controlling for other factors.

This report focuses on the advantages and disadvantages of alternative approaches to these problems. The study is based on a detailed review of the methodologies in previous cross-sectional investigations of labor supply and is intended for use in conducting and evaluating future empirical research on the effects of alternative IMPs on national and regional labor supply. This report discusses the implications of making one methodological choice rather than another and points to areas where improved data or theory can make the greatest contribution in improving labor supply estimates. One important step in obtaining improved labor supply estimates is careful construction of sensitivity tests to examine the empirical significance of making one methodological choice rather than another. Existing cross-sectional data sets are examined, and there is a discussion of testing for geographic differences in labor supply parameters.
AN ANALYTICAL FRAMEWORK FOR STUDYING THE POTENTIAL EFFECTS OF AN INCOME MAINTENANCE PROGRAM ON U.S. INTERREGIONAL MIGRATION

R-1081-EDA, Julie DaVanzo, December 1972

It is often alleged that welfare reform will have important repercussions on interregional migration in general and on migration to and from depressed areas in particular. Existing analytical frameworks are inadequate to answer the question from a theoretical standpoint alone. Empirical studies of U.S. migration patterns have been limited by available data and have not shed much light on this issue. In this report a model is proposed that is more appropriate for predicting the potential effects of an income maintenance program (IMP) on interregional migration.

Following are listed the unique features of this study:

1. The ability to finance investments in migration plays an important role. The supply of funds (ability to finance) as well as the demand for funds (derived from the rate of return available on the migration investment) are considered.

2. For married people the decisionmaking unit is the family rather than the individual; hence, the wife's employment opportunities affect the husband's migration decision and vice versa.

3. A wide range of demographic groups, including women and non-whites, is studied in the empirical analysis. Statistical tests are performed to determine whether the differences in the parameters estimated for various demographic groups are statistically significant.

4. Some new explanatory variables (military indexes, an income-distance interaction term) are included in the econometric model.

Migration is viewed as an investment in human capital—that is, an investment in the productivity of the human agent—because it involves costs (direct, opportunity, information, and psychic costs, and decreases in the value of capital that is costly to transfer to a new location) and yields returns (changes in real income) over some subsequent period of time and because it is embodied in the human being. An individual
or family will invest in migration so long as the marginal net benefit, or rate of return, from an additional dollar invested is greater than the marginal cost of financing that dollar of investment. Wealthy individuals and families will tend to have lower financing costs: (1) they can borrow at lower rates of interest in the capital market because they can offer acceptable collateral, and (2) they can more easily self-finance an investment out of past savings or current forgone consumption. Thus, other things the same, wealthy individuals and families can be expected to migrate more often than poorer ones. The rate of return to migration will depend on the change in earnings and in non-earning returns (such as welfare benefits) and the costs associated with the move.

An econometric model explaining 1955-1960 gross migration rates among the nine U.S. Census divisions is estimated for all sex-color-education cohorts of persons aged 25-29. Aggregate data from the 1960 Census are used; shortcomings of these data are discussed. For each cohort the migration rate is regressed on six explanatory variables and a constant:

1. the ratio of income at destination to income at origin—a proxy for the rate of return to migration,
2. income at origin—a proxy for ability to finance the migration investment,
3. the logarithm of the distance between origin and destination—a proxy for many of the costs of moving,
4. the multiplicative product of income at origin and the logarithm of distance—an interaction term to allow for the possibility that ability to finance becomes a more important consideration the costlier (longer distance) the move,
5. and (6) military indexes for origin and destination—to hold constant for military moves (inductions, separations, transfers) included in the migration data studied here.

The difference between income at destination and income at origin has been tried as an alternative measure of the return to migration. In addition, other explanatory variables—for example, cost-of-living
indexes, unemployment rates, welfare benefits—have also been tried and their performances are discussed.

The regression coefficients and partial derivatives of migration with respect to income and with respect to distance usually have the signs expected and are statistically significant (at the 5 percent level or better). The income ratio and the origin and destination military indexes generally have positive coefficients. The partial derivative of migration with respect to distance is usually negative and decreases in absolute magnitude as income at origin increases; thus distance appears to be less of a deterrent to migration as income increases. The partial derivative of migration with respect to income at origin, holding constant the income ratio, is usually positive and increases with distance, indicating that long-distance moves are more sensitive to income than shorter moves. The partial derivative of migration with respect to income increases with education; this, along with the increase in income with education, explains in part why migration rates increase with education.

Because members of a family usually move together, the family is the appropriate decisionmaking unit for studies of migration. "Family" measures of income, which are weighted averages of male and female incomes, have been tried in the econometric model. They sometimes perform better than measures that refer only to the sex whose migration behavior is being studied or only to the opposite sex. However, the performances of these "family" measures are undoubtedly hindered by the many fairly strong assumptions necessary to construct them from aggregate data. Data on individuals in households, matching husbands and wives, will permit a more satisfactory study of this family decision-making process.

Estimated coefficients of the explanatory variables tend to vary considerably, often systematically, by sex, color, and education. To determine whether these differences are statistically significant, tests of coefficient homogeneity (F-tests) were performed. Educational differences in coefficients were statistically significant (at 5 percent) for all sex-color groups. Differences in coefficients by sex were never significant. Coefficient differences by color were significant
only for persons who had attended college, but further testing revealed that for these highly educated cohorts only the intercept and the coefficient of the income ratio differed significantly by color.

Although these empirical results provide some insights into possible effects of an income maintenance program on migration, further research is needed for accurate prediction of the potential effects of an income transfer program on the volume, direction, and composition of migration flows. Six hypotheses are proposed that should be explored in future work:

1. Income maintenance payments should facilitate the financing of migration.

2. Income maintenance payments should reduce the riskiness of investments in migration.

3. If the migration process itself is a normal or superior good, an IMP is likely to increase migration in general for families whose incomes increase after the program is instituted. If the consumption returns to migration are normal or superior goods, an IMP may in particular increase migration to areas where these consumption returns are highest, for example, areas with pleasant climates.

4. For many persons covered by an IMP, the high marginal tax rates on earnings implicit in the program will probably reduce the labor market return available on migration investments and reduce the incentive to search for and migrate for better employment opportunities. This will be mitigated by a reduction in the opportunity costs of migrating.

5. A federally administered IMP is likely to reduce interstate welfare discrepancies. If persons are currently migrating to areas with attractive welfare programs to receive the greater benefits, a national IMP is likely to reduce this type of migration.

6. Supplementary programs and provisions of an IMP—for example, moving allowances—or short- or long-run side effects of the program—for example, changes in fertility—may influence migration decisions.

Several data sets that could be used to test these hypotheses are discussed. Longitudinal data on individuals and households are preferred: Data on individuals and families enable us to study the migration decisionmaking process at a micro-level and to determine how the
family's decision to migrate is influenced by the economic opportunities of each of its members; longitudinal data permit us to follow individuals and families when they move and to investigate the effect of a family's economic and demographic characteristics at the beginning of a period on their propensity to migrate during that period.
FAMILY DECISIONMAKING OVER THE LIFE CYCLE: SOME IMPLICATIONS FOR ESTIMATING LABOR SUPPLY

R-1121-EDA, James P. Smith, forthcoming

Perhaps the most important piece of social legislation to be considered recently by the Congress was the Family Assistance Plan (FAP) proposed by the Nixon administration. Although this particular bill was defeated, it is likely that some form of FAP will be submitted to future legislatures. In some ways a FAP is a radical departure from existing public policy, and scholars from many disciplines have been called upon to analyze and predict the behavioral consequences of this legislation. One behavioral aspect that has received considerable attention is the labor supply effect. Because a FAP contains negative income tax elements, it was felt that it might seriously disrupt work incentives and lead to a large reduction in work effort among new welfare recipients. Since economists could offer a well-developed theory dealing with the labor supply aspects, it was natural that they would play a central role in designing and evaluating alternative proposals.

The model that economists offered was the standard textbook problem of an individual choosing between labor and leisure. An individual will select that combination of goods and leisure that maximizes his utility subject to the constraint that his income is limited and the prices of goods and leisure are positive. Conceptually, the opportunity cost of consuming an hour of leisure is the amount of market goods forgone, so the price of leisure equals the real wage rate. Because there are income and substitution effects operating in conflicting directions, an increase in the real wage has an ambiguous effect on the number of hours worked. As long as leisure is a normal good, the larger income associated with a higher wage would tend to reduce market work. The higher wage has also made the consumption of leisure more expensive and, on this substitution effect, market hours will increase.

A FAP includes two dimensions that could alter an individual's labor supply: First, there is an income subsidy to recipients that increases real wealth; and second, the program imposes a tax on market
earnings. Since the FAP simultaneously increases real income and lowers the net wage received from working, using the model described above economists have predicted that enactment of such a proposal would lower the amount of market work of those individuals participating in the program. (This statement is a comparison between a FAP and no welfare system, and not a comparison between a FAP and the existing welfare system.) An impressive number of empirical studies have already been completed by economists, but the range of estimates on the labor supply effects of a FAP based on these studies is extremely wide. Many of the criticisms and suggestions for improving these estimates have centered either on increasing the quality of the data or on using a more sophisticated econometric technique. There has been almost no questioning of the appropriateness of the theoretical model. In this study, I will argue that the standard model is in some ways an inadequate tool to evaluate a FAP.

The principal difficulty is that a FAP will have important inter-period or life-cycle effects. In the standard model, which contains just one time period, interperiod effects are ignored. The one-period model is appropriate only when the proposal being investigated does not alter the incentives to substitute economic activity between time periods. But a FAP will typically change an individual's wages by different percentage amounts at different points in his life cycle. This provides individuals with incentives to alter the timing of their market participation. Observing the change in labor supplied in only one period can give a misleading indication of the total effect of a FAP. For the purpose of studying a FAP, a complete model of labor supply must incorporate the effects on the timing of market responses. Recent contributions by Ghez and Becker have enabled us to include the timing aspects in an economic model of choice. By extending Becker's original one-period model to a lifetime context, they were able to place in sharp focus the previously neglected influence of cyclical, seasonal, and life-cycle movements in wage rates and other variables.

First, the pure life cycle model, an extension of the work of Becker and Ghez, is derived without reference to its implications for a FAP. The timing of market participation is shown to depend upon the
life cycle wage pattern of men and women, the rate of interest and the rate of time preference, and any age-related changes in the productivity of nonmarket uses of time. The predictions of the pure life cycle and pure one-period models are contrasted. This comparison attempts to clarify circumstances under which the life cycle model should be used and those under which the single-period model is appropriate.

Next the theoretical model is used to predict and analyze the expected labor supply effects of a FAP. It is shown that the effects of a FAP partly reflect life cycle considerations and partly the more standard one-period model. The appropriate model to use, a marriage of the two pure special cases, is examined and justified. Particular emphasis and attention is given to those instances in which the predictions of this new model diverge from those obtainable in the model currently used.

This new model shows that it is essential to identify those periods in the family life cycle when the family is eligible for benefits and those in which it is not. The use of the one-period model has probably led researchers to underestimate the magnitude of the labor market withdrawals in those years in which the family is eligible for benefits.

Finally, human capital investments are included in the model. This generalization leads to a number of predictions concerning which groups in society are most likely to have the largest labor supply reactions to a FAP. For example, economic theory suggests that young married women and individuals in older families will exhibit larger reductions in their market work than other groups in society.

Another aspect of the model should be mentioned—the use of the family context. As recent theoretical contributions have stressed, an individual's decision about the number of hours to exchange for market dollars is often made in a family context. Hence, the hours of work of any family member depend not only on his wage and other variables specific to him but also on similar variables of other members and on those variables common to the family unit. In this model, a fundamental assumption is that the division of labor within a family is determined in large part by economic forces. Relative male-female productivity differences in both the market and household sectors supply the
incentives for families to concentrate the time of a member in that sector where he has a comparative advantage.

Ignoring this family context has led to serious errors in predicting the effects of a FAP. Typically, to predict the effect of a FAP on the hours worked by male heads of household, others have shown the effect on male hours worked of a reduction in male wages, holding female wages constant. In a family that participates in a FAP, however, both the male and female true market wages (net of FAP taxes) are reduced simultaneously. The model in this report is developed on the assumption that both market wages are changed, and the resulting expected changes in market hours for male and female members of the household are derived. Part of the incentive to substitute in favor of male time in the household is offset because the earnings females can receive by additional market work have been reduced. Thus, economic theory would lead us to expect that the main avenue of the reduction in a family's work effort will be a large decrease in hours worked by female household members. Also, I show that a FAP generally will have a larger "effect" on the "true" wage of female members of the household than on the wage of males. One of the main conclusions of this report is that those studies that have concentrated on the reallocation of market time of either males or females separately are giving policymakers seriously biased estimates of the labor supply effects of a FAP.
NOTES ON THE LOG-LINEAR OR LOGISTIC MODEL FOR THE
ANALYSIS OF QUALITATIVE SOCIO-ECONOMIC DATA

R-1306-EDA, Marc Nerlove and S. James Press, forthcoming

Although the problem of estimating relationships that determine qualitative variables\(^1\) has been studied for many years in such fields as bioassay, economic analysis has concentrated almost entirely on quantitative problems. Whenever problems were considered that were actually qualitative, the data were generally forced into a quantitative framework and the fundamental differences between the two types of problems were either obscured or ignored entirely. The increasing availability of survey data for individuals, households, and firms calls for the study of qualitative relations in an appropriate framework in order to take full advantage of such data.

To study such relationships, one must first distinguish between two types of qualitative problems: those in which the relevant dependent variable takes on one of two distinct values (for example, true or false, success or failure, migrate or not) and the more general problem in which any number of outcomes may occur (for example, individuals choose one from among several migration destinations, or take one of a variety of jobs, or fall into one of several marital status categories). The former are referred to as "dichotomous" or "binary" variables, and the latter are known as "polytomous" variables.

Although the analysis of binary data has been considered for some time in econometrics, the techniques used have been largely linear regression techniques that were designed for the analysis of quantitative problems, but modified to take into account the special nature of binary variables. Two distinct types of problems arise when a dependent variable is dichotomous. Since the variable of interest is actually a probability, its value must lie in the range between zero and one. There is nothing in standard regression techniques, however, that constrains the predicted values of the probability to this range. Moreover,

\(^1\)Defined in footnote 1, p. 51.
the variance of the error for a dichotomous dependent variable will be smaller in extreme cases when the probability is near zero or near one than in cases when the probability is around one-half. This non-constant variance (known as heteroscedasticity) will produce estimators that are inefficient—that is, estimators with fairly large variances. Although modifications of linear regression techniques can reduce the influence of these problems, these modifications also result in a loss of information in many cases because some observations may be dropped from the sample or grouped with other observations for which the explanatory variables are slightly different.

Certain nonlinear techniques, known as maximum likelihood methods, retain all the sample information. These nonlinear models include the probit, normit, and logit models. All three are computationally more complex and expensive than most linear estimation procedures. The first two are based on the normal distribution, and the last uses the logistic distribution. Although computer programs have been written to estimate all three of these models, the logit model is used most often for two reasons. First, although both the normal and logistic models give similar results for most problems, the logistic function is computationally much simpler and, hence, less expensive to use. Second, the logistic model can easily be generalized to treat problems in which the dependent variable represents an event with any number of qualitative outcomes. It is with this polytomous or multivariate dichotomous problem that the study is primarily concerned.

Polytomous (or many-valued) qualitative variables are of two conceptually distinct types. Appropriate estimation techniques for treating both cases, however, are identical. The first type results from the classification of an observation into one of several categories according to a single criterion. The choice from among several migration destinations is one such classification. The second type of polytomous variable results from the cross-classification of several binary variables, which may or may not be independent. Individuals, for example, may be classified according to whether or not they are participating in the labor force, and each of these two groups may then be classified as to whether or not they are high school graduates. A
determination of whether or not such classifications are interdependent can be of great interest. Such questions are commonly analyzed through contingency tables. It can be shown, however, that the normal analysis of contingency tables can be treated as a special case of the multivariate logistic model. By estimating the relationships in a logit model, the interdependencies among various cross-classifications can be studied. Moreover, a logit model can also be used to examine the effects of various exogenous variables on different classifications and on interdependencies among classifications.

A computer program for estimating a polytomous logit model by maximum likelihood has been written and is currently being tested. A complete description of this program and its uses, together with a summary of work completed on a simple binary logit model, will be included in the report.
THE ESTIMATION OF LABOR SUPPLY FUNCTIONS
FOR SECONDARY WORKERS

R-1265-NIH/EDA, T. Paul Schultz, forthcoming

The regional influence of envisioned changes in income maintenance programs is likely to be most noticeable in its effects on the labor market behavior of secondary workers, in particular that of married women. The analytical difficulties of predicting the responsiveness of this group to reform in welfare programs is the topic of this exploratory report. Understanding labor market behavior of secondary workers depends on inferring the nature of two structural relations: The first determines the market wage offered individuals, and the second determines the shadow value individuals attach to their time in non-market activities. Very strong assumptions are currently required for rigorous identification and estimation of these relationships from available survey information. This report argues that several sources of bias in specifying and estimating labor supply functions for secondary workers can be isolated and at least reduced, if not eliminated altogether, by improved analysis of existing empirical materials.

It is widely believed that observed wages of secondary workers embody the effects of past on-the-job training and experience, which are likely to be related to how many hours the individual currently works. If one considers only labor force participants, the observed wage rate can be replaced as an explanatory variable in the labor supply equation by a predicted wage rate based on an equation for a "permanent" market derived demand wage, as suggested from the human capital literature. This simultaneous equation approach should purge the wage rate of the influence of such factors as past participation decisions that are likely to influence both observed wages and current hours worked and bias upward (positively) the estimated labor supply response to permanent changes in market wage opportunities. It is commonly observed that nonparticipants are disproportionately concentrated in those groups that receive relatively low market wages when they enter the labor force. If the sample under consideration includes nonparticipants for
whom a market wage offer is attributed according to the same structural equation, the estimated responsiveness of secondary workers to changes in their market wage rate is increased.

A maximum likelihood estimation procedure for estimating a relationship where the dependent variable takes on the value of zero for a substantial proportion of the sample has been proposed by Tobin. This procedure seems preferable to ordinary regression techniques when labor supply functions are estimated for secondary workers. This nonlinear method yields estimates that do not differ radically from those obtained by ordinary regression techniques evaluated for individuals with average characteristics. But the "Tobit" model estimates do imply different responsiveness for low-wage segments of the population who are likely to be the chief beneficiaries of welfare reform. The report restates, but does not resolve, the difficult problems of distinguishing how specific characteristics of individuals shift their derived supply of labor (to the market) and also influence their market wage offers. Several approaches to this problem, as recently developed by Gronau and Heckman, are discussed.

To illustrate the quantitative importance of the alternative model specifications and estimation techniques discussed in the report, labor supply functions are estimated for currently married women aged 14 to 64, based on the 1967 Survey of Economic Opportunity. The first model of labor supply is estimated for currently married women for whom a wage rate was observed. Both the observed hourly wage rate and the instrumental variable estimate of the hourly wage rate, each adjusted for taxes and price level, are alternatively used. The own-wage effects based on imputed wages tended to be smaller and less statistically significant than those based on observed wage rates for all of the younger women, but larger for older black women. Women outside of the labor force are added to the sample by imputing a wage rate based on their characteristics; the model is then reestimated for all currently married women. In every age and race group both the own-wage effect and the husband-wage effect are increased in absolute magnitude and usually in statistical significance. Income effects, however, are weak and vacillate in sign when a "nonemployment income" variable is included
in the labor supply model, which implies that the \textit{uncompensated} wage effects are parametrically given. On the other hand, when a Beckerian "full income" variable is included in the model, which implies that the \textit{compensated} wage effects are fixed, the negative income effect is statistically significant for most white groups, but not consistently for the black samples. Only when the reduced form model of the labor market is estimated are the regression coefficients for the "full income" variable consistently negative and greater than their standard errors in every age and race group; however, the statistical significance of the implied income effects is still greater for whites than for blacks. Application of the Tobit nonlinear framework strengthens these results further, increasing the apparent responsiveness of the labor supply decision to own wages, husband's wages, and full income.

The major unresolved puzzle that emerges from these preliminary estimates of labor supply functions for currently married women is the strong effect of secondary and higher education on the labor supply decision that is not accounted for by schooling's effect on women's market wage offers. In other words, married women with secondary and higher education consistently supply more hours per year to the labor market than is explained by variation in their own permanent wage offers, their husband's permanent wages, and their household's full income. Although this result is not difficult to rationalize, it suggests that the power of labor market conditions to explain differences in labor force participation by women of different educational attainment has been overstated, and perhaps the labor supply responsiveness of women to their own permanent wage offers has also been overestimated by traditional formulations of the labor supply model. Improved specification and estimation of labor supply functions for secondary workers will considerably alter estimates of the effects of welfare reform on labor force behavior.