SIPP

SURVEY OF INCOME AND PROGRAM PARTICIPATION

Working Paper Series

Longitudinal Analysis of
Federal Survey Data

No. 9007-112

Patricia Ruggles
Joint Economic Committee

June 1990

TABLE OF CONTENTS

I. Introduction

Longitudinal panel data provide a unique opportunity to examine patterns and sources of economic and demographic change at the individual and family level. These data are relevant to a host of policy issues, from the assessment of welfare program participation to an understanding of patterns of health care usage or of the determinants of retirement. Many policy issues require some understanding of the factors that lead up to a particular event, or of the consequences that stem from it. Without repeated observations of the individuals concerned, however, such factors and consequences can only be inferred. Thus, our increasing store of longitudinal panel data holds the potential for major breakthroughs in our understanding of the basic determinants of economic and demographic change as they affect individuals and families over time.

Unfortunately, however, many of our longitudinal data sets have been somewhat under-used by researchers so far, especially compared to similar cross-sectional surveys. To some extent this under-usage may simply stem from the fact that many of these data sets are still fairly new--researchers need a chance to become familiar with the opportunities offered by these new sources of information. A more fundamental problem, however, is that to an analyst whose primary research experience is with cross-sectional microdata, a longitudinal panel of microdata on families and individuals can be rather intimidating.

A longitudinal database designed to offer a reasonably representative sample of the non-institutionalized population, for example, will be much larger and more complex than a similarly representative cross-sectional sample, since every observation will have been repeated several times. Size alone is likely to create some analysis problems, and the mechanisms used by the panel's designers to track

involving the use of longitudinal data. Further, the paper focuses almost exclusively on the application of longitudinal analysis to questions concerning patterns of family income, expenditures and/or demographic change.

The remainder of this paper is organized into three sections. The first of these addresses basic issues in designing a file for longitudinal analysis. The most crucial of these issues, in my view, is choosing the appropriate unit of analysis for the application at hand. This section discusses the pros and cons of alternative choices, and considers the implications of these choices in constructing an analysis file. Other problems in file construction--dealing with multi-wave data, handling attrition bias and longitudinal weighting, and the pros and cons of various types of imputation--are also considered very briefly.

The following section considers specific methods of making comparisons across time. The major focus of this section is on matching the outcome measures and statistical techniques chosen to the basic research question being asked. For many policy issues fairly simple outcome measures may be perfectly appropriate, but it is important to understand the measurement implications of alternative choices in order to avoid misinterpreting one's results.

The final section of the paper describes a few specific examples of current approaches to the measurement of economic and demographic phenomena using longitudinal microdata. It then concludes with a discussion of additional steps that federal statistical agencies could take to facilitate the analysis of the various longitudinal databases they produce.

II. Creating a Longitudinal Analysis File

Most longitudinal data on individuals, families and households come from surveys that consist of a series of interviews with selected sample members

for each household member. It is a fairly simple matter to analyze data across all households, across families or sub-families, or across individuals--or even to combine information from all three levels into one analysis.

Although even in the cross-sectional case the choice of a unit of analysis will have some impact on one's outcome measures--the measured poverty rate for families, for example, is different from the rate for individuals--definitions of the terms "household," "family," and "person" are by now familiar, and the implications of choosing one unit rather than another are generally clear. The appropriate unit of analysis in such a case will typically depend on the purposes of the analysis being undertaken, rather than on the constraints imposed by the data.

Unfortunately, even relatively straightforward terms such as "household" or "family" lose a great deal of their precision when they are considered longitudinally, however. Although one may think of the family income reported in the CPS as applying to essentially the same "family" over the period of a year, for example, in fact families may undergo a substantial amount of change over the course of a year. This problem is addressed in the CPS by fixing the composition of the family at the point of the interview. Retrospective data is then collected on the incomes over the previous year of all those who happen to be in the family on the interview date, regardless of whether each specific person was a member of that family for the entire year.

This approach can be duplicated using longitudinal survey data, but for a survey like the SIPP which offers actual month by month data on both family composition and income such an approach seems both cumbersome and potentially

and the use of some longitudinal definition produces results that are quite significantly different from those seen when family composition is treated as fixed.

The idea of using a longitudinal family definition is attractive, at least initially, to many policy analysts--after all, many, even most, policy issues of interest pertain to the family, not to the individual. Indeed, income considered at the person level is a fairly meaningless concept, since a major assumption of both our income support policies and of our social support system in general is that the members of families and households pool their incomes, and at least to some extent, make joint consumption decisions.

Over the past several years, however, both the work on this topic by Citro, Herriot, and others and my own experiences in analyzing longitudinal income data from the SIPP and the PSID have convinced me that the use of a longitudinal family concept can in fact be quite misleading. In general, in order for a measurement concept to be useful, minor variations in its specification should not result in major differences in the quantities being measured. When measures ar not robust in this way, it is difficult to tell whether specific outcomes are related to actual differences in behavior or some other factor across groups, or whether they are simply artifacts of the measurement method. In my experience, longitudinal concepts of the family typically fail this test. Additionally, use of such a concept can actually impede longitudinal analysis, since the very factors that are of interest for much policy research--the impacts of divorce, out of wedlock births, deaths, and so forth--also tend to change family composition, and under many definitions result in new families. Linking these new families with their predecessors in a way that facilitates our understanding of the impacts of these transitions can become very difficult, since families can combine and recombine in many different ways over a given observation period.

analysis. For the most part, these modifications involve moving additional information on the family and household in which each person resided at each observation point onto the person record.[4] In essence, this will create a personal longitudinal family history for each person in the longitudinal file.

To illustrate, consider the case of a married couple who, within the observation period contained in the longitudinal file, have a baby and then become divorced. Further, suppose that both adults were working at the beginning of the period, but that the woman left her job when the baby was born, and the man then experienced a spell of unemployment that lasted until after the divorce. Clearly, in order to have a complete picture of the events that have happened in this family it is necessary to have some information on the family as a whole--neither adult's person-record alone will contain all the necessary information (such as, for example, the other adult's work status).

On the other hand, a longitudinal family record will probably not contain all the information either--after the divorce, one or both of the adults will normally be considered part of a new family. As a result, that person (or both persons) will appear in two separate families within the observation period, making it difficult to examine issues such as the impact of the divorce on each spouse's income and poverty status. If either spouse remarries (or even moves in with a parent or other

---

[4]The slightly awkward term "observation point" is used here to refer to the unit of time over which specific longitudinal surveys collect their observations. This unit will typically be either a month (as in the SIPP, the NMCUES and the NMCES) or a year (as in the PSID and most cohorts of the NLS). A few surveys (e.g. the RHS) collect their data over longer or even irregular intervals. Note that the "observation point" is not necessarily the same as the "interview point," since some surveys collect information on several intervals of time that have occurred between interviews. As discussed briefly below, this technique is likely to lead to irregularities in the reported data series, but may facilitate using the data to simulate "event histories"--continuous records of the duration of specific states--which is helpful in applying certain analytic techniques.

by many analysts--examples include family income and family size. Such very common variables should almost certainly be constructed and appended to every adult's person-record for every observation date by the agency issuing the longitudinal data set. For family and household variables of importance to a specific analysis that are not routinely coded onto the person-record, however, the analyst must create a specific routine to produce appropriate recodes. In the longer run, it would again be helpful if agencies producing longitudinal data files could provide analysts with some help in implementing such recodes.

In addition to such simple variables, which can be extracted directly from the family record and appended to the person record, it is also very helpful for longitudinal analysis to create some specific transition flags that relate to the family changes under examination. For example, in considering the impacts of divorce on income and poverty status it would be helpful to have flags indicating changes in marital status. Appending such a flag to the person-record for the observation point in which the divorce takes place greatly simplifies the examination of the impacts of this change on other variables.

In fact, certain transitions are so important in determining income and family status that the data-issuing agencies should be encouraged to flag them routinely before issuing longitudinal data files. Such transitions might include divorce, marriage, birth of a child to a family member, death of a family member, and the loss or gain of a job by a family member.

Once family variables and transition flags have been appended to each person-record, it is a fairly straightforward matter to examine changes in family status and their impacts on income. To consider the impacts of divorce, for example, one would simply examine all the person records that contained the divorce flag. Family income and family size for each such person at each observation point could be

between interviews, although those interviews may nonetheless ask about annual income and family data.

Surveys of this type, where the time-unit observed and the interview schedule are not coincident, pose some problems for the analyst. Inevitably, respondents are most likely to report any changes in income or family status as having occurred either at the beginning or the end of the interview period. Changes in income and other transitions, therefore, appear to occur much more heavily at the "seams"--the months (or other intervals) that represent the end of one interview period and the beginning of the next. This results in a situation where the probability of a transition is overstated at the beginning and end of the interview period, and understated in the intervening interval.

The extent to which this is a problem for the analyst depends on the type of analysis being done. If one is primarily interested in the determinants of a particular type of transition, for example, a slight misreporting of the date of the transition may not matter much--the respondent who misremembers when a given event took place is also likely to misremember other related events, and the relationship between events may therefore not be distorted at all. To put it another way, a respondent reporting both a divorce and a job loss is likely to get them in the right sequence--"I got divorced right after I lost my job"--even if he misreports the date of one or both events.[5] Similarly, if the topic of interest is an ongoing state at the time of a particular transition--"I was unmarried and unemployed when I first went on welfare"--it is also likely to be reported correctly even if the date of the transition itself is slightly off.

---

[5]Nathan Young has examined the correlation of events on and off the seam, and has indeed found that in most cases correlations remain similar even when reported rates of transition are very different in the interview month and in other months. See Young (1989) for more discussion.

Attrition and Longitudinal Weighting. A problem related to "seam-bias" is that most longitudinal surveys lose some respondents over the course of their operation, and unfortunately these losses are often correlated with the very factors that the analyst would like to study. Indeed, some apparent transitions at the seam are in fact caused by people dropping out of the survey, since a failure to participate in the next interview is of course always discovered at the scheduled interview date. This type of seam bias can be eliminated by carefully distinguishing between those who report the termination of a given state and those who leave the sample, as discussed further in section III below, but the underlying problem of attrition is more difficult to solve.

Problems of attrition are more important for some types of analysis than for others. For comparisons of repeated cross-sections, for example, they are very important--it is easy to mistake changes in the sample for changes in the underlying population if there is differential attrition across sample sub-categories. Similarly, any analysis that attempts to describe the incidence of a given type of transition over time may be vulnerable to this problem. In many cases, it will be appropriate for the analyst to standardize across the population eligible to experience the transition in question, although sometimes that population cannot be defined narrowly enough to eliminate the problem of differential attrition, which may be correlated with unobserved variables.

The problem of attrition may have less impact on duration analyses, unless the spells being examined tend to be long relative to the observation period, in which case it may be difficult to find a reasonably representative sample of completed

---

for example, is not designed for use as a continuous survey.) As discussed further below, for such surveys analysis as repeated cross-sections is appropriate, but duration-related analyses probably are not.

using a self-weighted sample weights reflecting their relationship to the original sample universe should be used for all panel members, not just those in the sample for the full panel. For comparing repeated cross-sections, on the other hand, weights specific to the date of each cross-section should be used.

Imputation in the Longitudinal Context. Another approach sometimes used by data producers to handle the problem of attrition, as well as other reporting and response problems, is to impute variables and sometimes even whole records. Such imputations must be approached with caution. On the one hand, they can be very helpful in the context of repeated cross-sections, allowing one to carry out meaningful analyses without cumbersome changes in weights for each new cross-sectional observation. Even in this context, it is important to satisfy oneself that the imputations have been done in a way that provides some reasonable assurance that the outcome variables of major interest have not been seriously distorted, but for many types of analysis reasonably good imputations are quite possible.

Unfortunately, however, in dealing with longitudinal analyses it is often the analyst's first task to sort through the data and remove all the cross-sectional imputations put in by the data producer. Because many longitudinal panels are produced first as cross-sectional files they may contain imputations that are reasonable in the cross-sectional context but that are not designed with longitudinal applications in mind. For example, income imputations may be done for each wave without regard to the individual's income in any other wave, producing strange income patterns if the data are analyzed over time. In the longer run, it would be helpful if data-producing agencies would avoid cross-sectional imputations that distort the longitudinal record altogether, although this does require the

sectional imputation, at least for databases that are designed with longitudinal applications primarily in mind.

So far this paper has discussed the creation of a longitudinal analysis file in considerable detail (although unfortunately there remain many conceptual issues in file construction that have only been touched on briefly or that have been neglected altogether, and the practical programming problems involved have not been considered at all). As has been hinted several times, however, many of the specific decisions that must be made in putting together a longitudinal research file depend on the specific issues to be examined, and in particular, on the types of outcome measures to be used. The next section, therefore, goes on to discuss alternative approaches to comparisons over time.

III. Making Comparisons Across Time

The major purpose of a longitudinal research file is of course to facilitate the analysis of change over time. There are three major types of time-related analysis that are commonly carried out with such files, and there are some specific methodological issues that pertain to each.

Comparing Two Points in Time. The simplest type of time-related analysis--the comparison of data from two discrete points in time--does not actually require a complete longitudinal data file at all. The major advantage of this type of analysis is that it is relatively simple to implement and can often yield a great deal of useful information, particularly for questions that focus on rates of turnover in a specific variable. This method is very commonly used with many different longitudinal data sets--several examples of such analyses can be found for PSID data in the Institute for Social Research's volume of PSID research results entitled *Years of Poverty, Years of Plenty*, for example. Other examples include Alan Fox's study using RHS

This is helpful both in considering the effects of the transition on other variables and in estimating a causative model of the determinants of the transition itself.

To illustrate this point, let us reconsider the analysis of divorce discussed briefly above. If the analyst is interested not only in the determinants of the divorce transition, but also in its impacts, a simple comparison of two points in time may be doubly misleading. For example, family income may dip temporarily at the time of divorce as the family changes from one household to two. Eventually, however, as the two households make post-divorce adjustments in employment and living arrangements, income is likely to recover at least somewhat. Estimates of the impact of the divorce on income and poverty status for the various family members may be quite sensitive to both the unit definition used to compute income (as discussed in the last section) and to the specific timing of two income observations compared to the divorce itself.

In a case like this, examination of income or poverty status over a longer period leading up to and then following the transition will give a better picture of its actual impacts. For this type of examination it is necessary to have a longitudinally linked file with the transition flagged, but if such a file is available a descriptive analysis of this type is quite straightforward to perform.[11] Similarly, the transition flags themselves can be used as explanatory variables in a larger model of change over time as it affects some other variable. The recent paper by Suzanne Bianchi and Edie McArthur on the impacts of marital disruptions on children's economic status illustrates a transition analysis of this type.[12]

---

[11]Applications illustrating the use of this technique to analyze income change can be found in Ruggles and Williams (1986) and Williams and Ruggles (1987).

[12]See Bianchi and McArthur (1989).

shoehorn duration-related information into one's transition analysis--one could create separate dummy variables for short and long unemployment spells in the above example, for instance--this is a rather ad hoc approach that is likely to leave many unanswered questions. In addition, in many cases one is interested not only in the transition event itself, or even in its impact on other events, but also in the expected duration of the new state that it creates. One wishes to know, for example, how long someone who enters poverty may be expected to remain poor, or how long someone who loses a job may be expected to remain unemployed. Questions of this type require some type of duration analysis.

Analyzing Data on Duration. There are many possible approaches to questions of duration, and alternative approaches can produce quite different and even seemingly contradictory statistics. The confusion generally results from differences in the population to which the duration estimate applies. The two major possibilities are cohort-based estimates, which typically apply to all those observed in a given state at a point in time, and spell estimates, which apply to all those observed to enter the state within a given span of time.

To illustrate these possibilities, consider the case of welfare program participation. A point-in-time or cohort-based estimate of welfare durations will ask a question like "How long have those who are currently receiving welfare been on the program?" This question has been phrased retrospectively, but it can also be put in a prospective form: "How long are those currently on the program likely to remain on in the future?" In either case, the base population being considered is all those on the program at a given point in time. Such estimates are therefore relatively easy to line up with cross-sectional estimates of the total population on welfare, which are of necessity also point-in-time estimates. Estimates of this type are very useful for a number of purposes--for example, estimating the future costs

who are on welfare at a point in time are likely to have much longer spell durations, on average, than the typical entrant, because those with longer spell durations are more likely to be in the welfare population at any particular point in time.[16]

To see this point, consider a very simple example. Suppose the population of interest consists of 13 people, one of whom is in the state under consideration for one year, and twelve of whom are in that state for one month each. Further suppose that these twelve one-month spells are distributed so that one occurs in every month of the year. At any given point in time, therefore, the total population in the state being considered will consist of two people, one who is in a one-month spell, and one who is in a twelve month spell. A point-in-time analysis conducted any time after the first month will therefore conclude that 50 percent of the observable population reports a spell of more than one month. An analysis based on all entrants observed during the year, however, will find that only one-thirteenth of the population reports a spell of more than one month. Clearly, if the reasons for these differences in estimates are not well understood, they could lead to very different conclusions about the prevalence of long spells.

Many of the most useful and interesting questions that can be addressed using a longitudinal database are questions that relate to duration. In any type of duration analysis, however, it is necessary to be sensitive to the issue of censoring. Inevitably, there will be some spells that start before the beginning of the observation period or that end after the panel has come to an end. Further, there

---

[16]Mary Jo Bane and David Ellwood's classic paper on poverty spells makes this point very well, and provides a good example of spell analysis as applied to the PSID. (See Bane and Ellwood (1986)). For a similar example using SIPP data, see Ruggles and Williams (1989). Other useful applications include the work by Pamela Farley Short and her colleagues on spells of Medicaid participation and Rebecca Blank's imaginative use of longitudinal data from the Seattle and Denver Income Maintenance Experiments to examine spells of welfare program participation. See Short et al. (1988) and Blank (1986).

some of which fell in the sample period) cannot be estimated using these data. Estimates of the proportion of those observed who experience long poverty spells will be understated, because some spells that appear short are in fact longer, but they simply haven't been completely measured. At the same time, however, because these estimates mix together people who were poor in different years, they also cannot be used to predict, say, what proportion of those poor in a given year will still be poor eight years later.

Many analysts cope with the problem of estimating spell durations when some observations are censored by using some sort of survival analysis technique. Under this methodology, a survival function for a given type of spell is estimated based on the cumulative distribution of observed spell durations. In other words, in order to compute the probability that a spell of welfare participation, for example, will end in its sixth month, conditional on its having lasted for the first five months, one must include all cases known to have lasted at least five full months, whether or not their eventual disposition is known.[18]

To put it in more technical terms, the survival function for welfare participation may be estimated by defining $F^*(t,X_t)$ as the cumulative distribution of time on welfare, with X_t defined as a vector of independent variables affecting welfare participation (which may or may not vary with time themselves) and F^* representing the results of the series of participation decisions made to time t. At any time t, then, $F^*(t,X_t)$ may be seen as representing the probability that the

[18]This discussion is aimed at the analyst trying to decide whether this approach is appropriate for the particular application he or she has in mind. Anyone attempting to implement such an analysis should of course review some of the more technical literature on this topic. Tuma and Hannan (1984) provide a good basic an overview of these methods. In addition, the treatment in Allison (1982) may be helpful to analysts who are completely unfamiliar with event history analysis techniques.

that provide a reasonably continuous record for a reasonably large sample of individuals entering the state being examined can be used with this approach, however, which limits its usefulness with smaller or less focused data sets or those in which data has been collected in an intermittent pattern.

Other Issues in Longitudinal Analysis. The most crucial decisions to make in undertaking a longitudinal analysis of family income and demographic data clearly involve the major choices concerning the basic outcome measures described above. A number of other issues also arise in longitudinal analysis, however, particularly in making income comparisons over time. Some of these issues--"seam" bias, recall and coding errors, the role of imputation--have already been discussed briefly in the section of the paper on file creation, but another set of issues--those relating to the accounting period for income measurement--can also have big impacts on one's outcome measure, especially in working with a file such as the SIPP that provides data on income over a sub-annual period.

In a cross-sectional file such as the CPS one does not have any particular choice over the accounting period that is used--income information is collected on an annual basis, and that is the only way it can be analyzed. In the SIPP, however, information is collected on monthly incomes. This allows the possibility that it can be examined over every period from one month to 32 months, the length of the panel. As work by Roberton Williams and our own more recent work has clearly demonstrated, alternative accounting period choices can have very different implications for income measures such as poverty rates.[19] More than one fourth of the panel has at least one month with an income below the poverty level during calendar year 1984, for example, while only about 6 percent are poor in every month

[19]See Williams (1985) and Ruggles and Williams (1988).

over time to some extent--save money in good months, say, to tide themselves over bad ones--longer accounting periods such as a year may give a better picture of people's real level of resources than do shorter ones. On the other hand, where detailed data are available it may be more appropriate to examine other resources such as asset holdings to judge total resources, rather than assuming that such resources are available on average over the longer period even when the family lacks income in the short run.[21]

IV. Conclusions

In summary, the many new sources of longitudinal data on incomes and family structures that have become available in the last decade offer exciting research opportunities to the policy analyst, but they bring with them their own unique measurement problems. Because these data sources are both more complex and less familiar than are cross-sectional databases covering such topics, analyzing them can present some challenges. For analysts willing to address these challenges, however, there are useful solutions, and these data can be used to provide important new insights into the processes underlying economic and demographic change.

Indeed, as discussed briefly in the various examples of measurement problems and their solutions given throughout the paper, important applications of longitudinal analysis to policy issues have already been carried out in many areas. A few examples include Bane and Ellwood's analysis of poverty spells and of AFDC participation using the PSID; the work by Bruce Vavricek and Ralph Smith of the Congressional Budget Office on spells of unemployment insurance recipiency as observed in the SIPP; several Social Security Administration-sponsored studies on

[21]See Chapters 5 and 7 in Ruggles (1990) for more discussion of this point.

and data producers are understandably anxious to get these first products to the users as fast as possible.

Once a survey has been in regular production for some period of time, however, it would make sense to lessen the emphasis on cross-sectional files and to increase efforts to produce reasonable longitudinal data in a reasonably timely fashion. We already have excellent cross-sectional data on family incomes and labor force status, and unless the survey in question is clearly adding to our store of available cross-sectional data on a particular topic, cross-sectional applications should receive less attention. In particular, the level of effort devoted to activities such as cross-sectional imputation that have no application in the longitudinal context should be reduced. Instead, greater research efforts should be devoted to continuing problems like longitudinal editing and the development of reasonable longitudinal imputation procedures.

The second way in which statistical agencies could support longitudinal analysis would be to undertake more of it themselves. Data producers typically publish at least some cross-sectional information from the files they produce, and in some cases--the CPS publications in the Census P-60 series, for example, come to mind--these tables themselves provide important information on which policy-makers come to rely. It ought to be possible for the Bureau of the Census and other data producers to publish similar information, but of a longitudinal nature, using the longitudinal databases that they now produce.

The assumptions underlying survival analyses might be difficult to explain in such a context, but basic information on the experience of a given cohort, for example, is fairly easy to explain and to interpret. For instance, one could look at how many of those becoming unemployed in a given period were still unemployed

could also increase substantially the useful information that we are able to obtain from these surveys.

REFERENCES

Allison, P.D. "Discrete-Time Methods for the Analysis of Event Histories," in S. Leinhardt (ed.), *Sociological Methodology 1982*, San Francisco: Jossey-Bass, 1982.

Bane, Mary Jo and David T. Ellwood. "The Dynamics of Dependence: the Routes to Self-Sufficiency." Report prepared for the U.S. Department of Health and Human Services. Cambridge, Mass.: Harvard University, 1983.

Bane, Mary Jo and David T. Ellwood. "Slipping Into and Out of Poverty: The Dynamics of Spells." *Journal of Human Resources*, Winter 1986, 21(1), pp. 1-23.

Bianchi, Suzanne, and Edith McArthur. "Family Disruption and Economic Hardship: The Short-Run Picture for Children." Paper presented at the annual meeting of the Population Association of America, May 1989.

Blank, Rebecca. "How Important is Welfare Dependence?" Working Paper No. 2026. Cambridge, Mass.: National Bureau of Economic Research, Sept. 1986.

Citro, Constance F., Donald J. Hernandez, and Roger A. Herriot. "Longitudinal Household Concepts in SIPP: Preliminary Results." SIPP Working Paper Series No. 8611. Washington D.C.: U. S. Bureau of the Census, 1986.

Cox, B. and S. Cohen. *Methodological Issues for Health Care Surveys.* New York: Marcel Dekker, 1985.

Duncan, Greg J. (ed.). *Years of Poverty, Years of Plenty.* Ann Arbor, Mich.: Institute for Social Research, 1984.

Duncan, Greg J., Richard D. Coe, and Martha S. Hill. "The Dynamics of Poverty," in G. Duncan, ed., (op. cit.) 1984, pp. 33-70.

Ernst, L., D. Hubble, and D. Judkins. "Longitudinal Family and Household Estimation in SIPP." Proceedings of the Survey Research Methods Section. Washington D.C.: American Statistical Association, 1984.

Fox, Alan. "Work Status and Income Change, 1968-72: Retirement History Study Preview." *Social Security Bulletin*, 1976.

Gottschalk, Peter. "The Intergenerational Transmission of Welfare Participation: Facts and Possible Causes." Paper presented at the annual meeting of the Association for Public Policy Analysis and Management, November 1989.

Gustman, Alan L. and Thomas L. Steinmeier. "A Structural Retirement Model." *Econometrica*, May 1986, pp. 555-584.

Jones, B. "Development of Sample Weights for the National Household Component of the NMCUES." Research Triangle Park, NC: Research Triangle Institute, 1982.

Williams, Roberton and Patricia Ruggles. "Determinants of Changes in Income Status and Welfare Program Participation." Paper presented at the annual meeting of the American Statistical Association, August 1987.

Young, Nathan. "Wave Seam Effects in the SIPP." Paper presented at the annual meeting of the American Statistical Association, August 1989.