

Survey of Program Dynamics

January 28, 1999

MEMORANDUM FOR Daniel H. Weinberg
Chief, Housing and Household Economic Statistics Division

From: Donald J. Hernandez
Special Assistant

Subject: Comparing Response Rates for SPD, PSID, and NLSY

Introduction

The usefulness of data from any longitudinal or panel study that follows and interviews the same respondents over a period of years depends on the assumption that the data represent the relevant populations. Nonresponse by members of the sample is a potential source of bias that can undermine the quality of estimates derived using longitudinal data. This memorandum compares overall response rates between the initial interview and the most recent interview for three major national surveys, the Survey of Program Dynamics (SPD) conducted by the U.S. Bureau of the Census, the Panel Study of Income Dynamics (PSID) conducted by the Survey Research Center at the University of Michigan, and the National Longitudinal Survey of Youth (NLSY) conducted by the Center for Human Resource Research at Ohio State University. It also discusses major differences across the surveys in the approaches used to maximize long-run response rates, and the consequences of differential nonresponse for data quality.

Overall Response Rates and Alternative Approaches to Maximizing these Rates

Table 1 presents the current overall response rates for specified survey periods, both including and excluding deceased persons from the base (These are the most recent, final estimates published or available for each survey). The current mortality-adjusted cumulative response rate for the entire survey period between initial sample selection and the most recent interview is 50 percent for the SPD, a level substantially less than the rate of 64 percent for the NLSY, but substantially higher than the rate of 39 percent for the PSID (See two bottom lines of Table 1). Among persons designated for interview during sample selection, these rates indicate the

proportion who were successfully interviewed during each round of interviews, that is, during each interview period, without a single period of nonresponse.

There are three strategies commonly employed to recover more data in longitudinal studies:

(1) paying monetary fees to respondents, (2) returning to respondents who declined to be interviewed during a previous interview period, and (3) augmenting the second strategy by recovering data missed due to nonresponse during earlier interview periods. For a longitudinal or panel survey spanning a substantial number of interviews, it is not essential that persons be interviewed in each and every interview period to provide valuable data for an important genre of analyses, two-point-in-time before-and-after analyses that are likely to be among the most prominent applications of SPD data in assessing the effects of welfare reform. In fact, sample cases can provide valuable information even if interviews are separated by several years. For these and other reasons, the NLSY has always sought to maximize its long-term response rates by efforts to recontact respondents lost from the sample and return them to the status of interviewed respondents one or more years after an original nonresponse. These efforts are extremely successful. For example, excluding the deceased from the base, the NLSY response rate of 71.5 percent between the first and most recent interview (line 3 of Table 1) jumps to 83.0 percent when it is calculated to include persons who participated in 16 of the 17 NLSY interviews, and it jumps to 87.6 percent (nearly 90 percent) when it is calculated to include persons who participated in 15 of the 17 NLSY interviews! WOW! Drawing on the NLSY experience, the PSID recently began to experiment with and implement efforts to recontact respondents lost from the sample and return them to the status of interviewed one or more years after their initial nonresponse. Without pursuing these efforts as vigorously as it might have (because of limited funds), the PSID increased its effective sample size by 9 percent in a single year, increasing the response rate for persons interviewed in interview 1 and in the most recent interview from 54 to 59 percent.

The NLSY and the PSID also use monetary incentives to encourage responses from interviewees. Many studies have shown that response rates are improved substantially by monetary incentives.

Overall, at the present time, the SPD response rate is comparable to the other two major national longitudinal surveys focused on economic and family behavior, but the SPD does not currently employ any of these three NLYS and/or PSID strategies to maximize response. The potential problem that this poses for the SPD during the next few years is the subject of the next

section of
this memorandum.

Future Response Rates in the SPD

Although the current cumulative response rate for the SPD lies between the corresponding rates for the PSID and NLSY, the latter surveys have been in the field much longer (27 years and 17 years, respectively, compared to 6 to 7 years for the SPD), and they have involved larger numbers of field administrations (27 and 17 interviews, respectively, compared to 9 to 12 for the SPD). Many PSID and NLSY interviews occurred, however, in the 1970s and 1980s when response rates were generally higher for all surveys than has been the case in the 1990s. In the SPD, the response rate between the last SIPP interview (the 1992 and 1993 panels) and the first SPD interview (the 1997 SPD bridge) was only 79 percent, and preliminary estimates suggest that the response rate between the 1997 SPD bridge interview and the 1998 SPD interview was about 83.4 percent overall, or 85.2 percent excluding both deaths and the newly-institutionalized population from the denominator, for a preliminary estimate of a full panel SPD response rate of 48.7 percent, or an official household-based estimate of 50.0 percent (presented in Table 1). If the response rate remains at about 83 to 85 percent in subsequent annual interviews between 1999 and 2001 or 2002, the total cumulative response rate would fall to about 27 to 30 percent in 2001 and 23 to 26 percent in 2002, depending on whether or not the deceased and newly-institutionalized populations are included in the base. A response rate of 23 to 26 percent would be markedly less than the current directly-comparable rates of 39 percent for the PSID and 64 percent for the NLSY. Although as suggested in the following section, estimates from such SPD data might be unbiased, concern about the quality of SPD estimates would surely increase.

Nonresponse and the Quality of Estimates

Response and attrition over time tends to be highly selective of persons who are concentrated among lower socioeconomic status individuals. This has been shown in analyses of the PSID (Fitzgerald, Gottschalk, and Moffit: 1997a, 1997b; van den Berg and Lindeboom, 1997; Zabel, 1997; Ziliak and Kniesner, 1979) the NLSY (Falaris and Peters, 1997; MaCurdy, Mroz, and Gritz, 1997), and the SPD and SIPP (Lamas, Tin, and Eargle, 1994; Tin, 1996; Zabel, 1997). These analyses did not, however, find evidence that sample nonresponse or attrition leads to seriously distorted estimates pertaining to a variety of topics. Although the research addressing this issue is much more extensive for the PSID and the NLSY than for the SPD and the SIPP, the similarities across surveys in the nature of nonresponse and attrition suggest that estimates

from the SPD are in general likely to be no more biased than estimates from the PSID and NLSY; that is, they are likely to be of acceptably high quality. Nevertheless, in view of the fact that the primary purpose of the SPD is to evaluate the effect of welfare reform among the low socioeconomic status population, if the SPD does not improve its year-to-year response rates, particularly among persons of low socioeconomic status, concerns about the quality of SPD estimates will surely increase during the next few years.

Table 1. Response Rates for SPD, PSID, and NLSY: Period-Specific and Total* (In percents)

Interviews	PSID				NLSY	
	<u>SPD</u>	<u>SRC</u>	<u>SEO</u>	<u>Total</u>	<u>Always</u>	
Period-Specific						
Currently						
Sample-Selection to Interview 1	90.9	77	50.8	66.5	89.2	89.2
Interview 1 to Most Recent Interview						
All Deceased Included in Base	51.6	45.2	45.2	45.2	69.6	86.7
Known Deceased Removed from Base	53.6	53.0	53.0	53.0	71.5	NA
Expected Deceased Removed from Base	NA	NA	NA	59	NA	NA
Cumulative						
Sample-Selection to Most Recent Interview						
All Deceased Included in Base	46.9	34.8	23.0	30.1	62.1	
Known Deceased Removed from Base	50.0	40.8	26.9	35.2	63.8	
Expected Deceased Removed from Base	NA	NA	NA	39.2	NA	

*See attached notes for sources and detailed derivation procedures. The estimated response rate of 50.8 percent for the SEO portion of the PSID sample may be slightly too high, and my historical investigation continues of the cumulative response rate between the point at which the Census Bureau drew the original SEO sample and at which PSID interviewing began. "Known Deceased" refers to persons specifically reported to interviewers as deceased in each survey. "Expected Deceased" is calculated by PSID staff as of the 1994 interview by applying 5-year life-table mortality rates to 1968 original sample individuals. "Expected Deceased" has not been calculated for the SPD and NLSY, but if such estimates were calculated and the expected deceased were removed from the base, the response rates for these two surveys would be slightly higher than estimates based on the known deceased. The following provides a brief description for each survey of sample selection, interview 1, and the most recent interview. NA means not available at present.

Survey of Program Dynamics -- Based on samples drawn for 1992 and 1993 SIPP panels
Sample Selection ----- Samples drawn for 1992 and 1993 SIPP panels
Interview 1----- Wave 1 interviews of 1992 and 1993 SIPP panels
Most Recent Interview -- SPD interview in 1998, based on preliminary estimate for 1997-1998 response

Panel Study of Income Dynamics -- Based on samples drawn for the Census Bureau's Survey of Economic Opportunity (SEO) and a special sample drawn by the University of Michigan Survey Research Center (SRC) specifically for the PSID
Sample Selection ----- Samples drawn for 1966/1967 SEO and 1968 PSID
Interview 1----- Wave 1 interviews of PSID combined sample in 1968
Most Recent Interview -- Wave 26 PSID interviews in 1993, "Expected deceased" is for 1994

National Longitudinal Survey of Youth -- Based on sample drawn for NLSY
Sample Selection ----- Sample drawn for 1979 NLSY
Interview 1----- Wave 1 interviews of NLSY in 1979
Most Recent Interview -- Wave 17 interviews of NLSY in 1996

NOTES TO TABLE 1:

SPD -- Household response rates between sample-selection and the first interview were calculated for the SIPP 1992 and 1993 panels combined, based on results presented in U.S. Bureau of the Census internal memoranda for The Record from Michael F. McMahon on November 27, 1995 for Final 1992 Panel Type A and Type D Sample Loss Report and Final 1993 Panel Type A and Type D Sample Loss Report.

Persons interviewed in the first nine waves of the SIPP 1992 and 1993 samples became the SPD sample. Individual response rates between the first SIPP 1992 and 1993 panel interviews and the SPD "bridge" interview (1997) were derived by Donald J. Hernandez using the SIPP 1992 Panel Waves 1-10 Longitudinal File, the SIPP 1993 Panel Waves 1-7 Longitudinal File, and the U.S.

Bureau of the Census internal SPD 1997 file available on the Housing and Household Economic Statistics Division server on December 4, 1998. Deceased are identified from the SIPP data for the period between interview 1 and the final SIPP interview prior to the SPD interview. The response rate between SPD 1997 and SPD 1998 is preliminary, and removed from the base was both deceased and newly institutionalized populations. The estimate of 48.7 percent for the cumulative person-based estimate with known deceased removed from the base is slightly less than the estimate of 50.0 percent officially derived by the Bureau of the Census at the household level. The estimate of 50.0 percent is presented in Table 1.

PSID -- The PSID user's guide (Hill, 1992) notes that the original PSID sample actually consisted of two independent samples, one drawn by the Survey Research Center and referred to as the SRC sample, the other selected from the Survey of Economic Opportunity (SEO) conducted in 1966 and 1967 by the Bureau of the Census for the Office of Economic Opportunity (OEO). The initial response rate for the SEO sample (Survey Research Center, 1972) was calculated to be 74 percent, based on the sample of households provided to SRC by the Bureau of the Census and the Office of Economic Opportunity. This does not include the effect of (1) attrition between sample selection by the Census Bureau and the first interview by the Census Bureau of respondents in 1967, which led to a response rate was 91.6 percent (Office of Economic Opportunity, 1970), (2) sample loss through subsequent refusals to remain in the sample that became the SEO component of the PSID, since about 25 percent refused to allow their names to be passed to SRC (Hill 1992), and (3) the failure of some sampled addresses to be transmitted from OEO to SRC (Hill, 1992). To calculate the PSID SEO sample-selection to interview response rate of 50.8 percent, I multiplied the initial response rate of 91.6 percent, by the 75 percent rate of willingness to have names transmitted from the Bureau of the Census to SRC, by the 74 percent response rate obtained by SRC in seeking to interview households provided by the Bureau of the Census and the OEO. This does not take into account that address information for some sample persons who were willing to participate was not transmitted from the Bureau of the Census and OEO to SRC. I am conducting historical research to ascertain the magnitude of this sample loss. Introducing this source of sample loss into the calculations will reduce the current estimate of 50.8 percent. Of course, as in all the surveys discussed here, weighting procedures were designed to take various factors, including sample attrition, into account. The response rate for the SRC sample was 76 percent. The SRC sample constituted about 60 percent of the initial PSID sample, while the SEO sample constituted about 40 percent of the initial PSID sample.

The response rates for "interview 1 to the most recent interview" were obtained Survey Research Center (1972) and from Table 2a of the documentation provided by PSID via internet and titled "A Panel Study of Income Dynamics: Procedures and Code Books, Guide to the 1993 Interviewing Year, Procedures, Wave XXVI, A Supplement" obtained December 1998. The first and most recent interview years were 1968 and 1993. Techla Loup of the PSID was very helpful in identifying needed estimates and confirming the interpretation of specific estimates. Deceased are identified by PSID staff from the PSID data between the first and most recent interviews.

Sandra Hofferth provided the estimated response rate for 1994 with expected deceased removed from the base.

NLSY -- The source for these estimates is "NLSY79 Users' Guide, A Guide to the 1979-1996

National Longitudinal Survey of Youth Data" prepared for the U.S. Department of Labor by the

Center for Human Resources Research, the Ohio State University, August 1997. The response

rate of 89.2 percent between sample selection and first interview is obtained from Table 3.3.1

based on the cross-sectional and supplemental (sub)samples. The response rate of 69.6% between the initial interview (1979) and the most recent interview (1996) is obtained from Table 3.7.1. Deceased who number 224 by 1994 according to Table 3.6.1 were removed from the base. An additional 39 deaths (personal communication, Randall J. Olsen) for years 1995 and 1996 were also removed from the base. The "always" interviewed column includes in the numerator those persons who were interviewed in each of the 17 interviews. The "currently" interviewed column includes in the numerator persons who were interviewed in at least the first interview and the current interview. If the base is limited to those not dropped from the survey or deceased, the proportion of those interviewed in the first interview who missed no more than one interview out of 17 was 83.0 percent, and combined with the 4.6 percent who missed only two out of 17 interviews, the response rate was 87.6 percent.

Technical Note: Accounting for "Expected" Mortality

Although each of the response rates in Table 1 for the period between sample-selection and interview 1 would be slightly higher if mortality were taken into account by removing deceased

from the base, and although the response rates for the period between interview 1 and the most

recent interview for the SPD and NLSY would be slightly higher if "expected" mortality beyond

that reported to interviewers were taken into account as it is in the corresponding PSID response

rate of 59 percent, the general pattern of differences across the surveys would be unchanged for

rates that were fully adjusted for mortality.

REFERENCES

Center for Human Resource Research, Ohio State University. 1997.

NLSY79 Users' Guide: A Guide to the 1979-1996 National Longitudinal Survey of Youth

Data. Columbus, OH: Center for Human Resources Research.

Hill, Martha, 1992.

The Panel Study of Income Dynamics: A User's Guide. Newbury CA: Sage

Publications,
Inc.

Fitzgerald, John, Peter Gottschalk, and Robert Moffit. 1997a.
An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics. *The Journal of Human Resources* XXXIII: 251-299.

Fitzgerald, John; Peter Gottschalk, and Robert Moffit. 1997a.
An Analysis of the Impact of Sample attrition on the Second Generation of Respondents in the Michigan Panel Study of Income Dynamics. *The Journal of Human Resources* XXXIII: 300-344.

Lamas, Enrique, Jan. Tin, and Judith Eargle. 1994.
The Effect of Attrition on Income and Poverty Estimates from the survey of Income and Program Participation (SIPP), U.S. Bureau of the Census Working, SIPP Working Paper 9402.

Lillard, Lee A. And Constatijn W.A. Panis. 1979.
Panel Attrition from the Panel Study of Income Dynamics: Household Income, Marital Status, and Mortality. *The Journal of Human Resources* XXXIII: 437-457.

MaCurdy, Thomas; Thomas Mroz, and R. Mark Gritz. 1979.
An Evaluation of the National Longitudinal Survey on Youth. *The Journal of Human Resources* XXXIII: 346-436.

Office of Economic Opportunity, 1970
A comparison of Selected Economic and Demographic Characteristics from the 1966 and
1967 Surveys of Economic Opportunity and the Comparable Current Population Survey. Transmitted on January 20, 1970, by Carol B. Kalish, Office of Economic Opportunity, Executive Office of the President to Distribution List of Persons at Bureau of the
Census,
Bureau of Labor Statistics, Department of Commerce, Department of Health, Education and Welfare, Brookings Institution, Pennsylvania State University, and Urban Institute.

Survey Research Center, University of Michigan, 1972.
A Panel Study of Income Dynamics: Study Design, Procedures, Available Data, 1968-1972
Interviewing Years (Waves I-V), Volume 1. Ann Arbor, MI: Institute for Social Research.

Survey Research Center, University of Michigan. 1998
A Panel Study of Income Dynamics: Procedures and Codebooks (Documentation),
Guide
to the 1993 Interviewing Year, Procedures, Wave XXVI, A Supplement. Ann Arbor, MI,
Institute for Social Research.

van den Berg, Gerard J. And Maarten Lindeboom. 1979.

Attrition in Panel Survey Data and the Estimation of Multi-State labor Market Models. The Journal of Human Resources XXXIII: 458-478.

Ziliak, James P. and Thomas J. Kniesner. 1997.

The Importance of sample Attrition in Life Cycle Labor Supply Estimation. The Journal of Human Resources XXXIII: 507-530.

Zabel, Jeffrey E. 1997

An Analysis of Attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation with an Application to a Model of Labor Market Behavior. The Journal of Human Resources XXXIII: 479-506.
