4

Effects on Panel Attrition and Fieldwork Outcomes from Selection for a Supplemental Study: Evidence from the Panel Study of Income Dynamics

Narayan Sastry, Paula Fomby and Katherine A. McGonagle

Survey Research Center, University of Michigan, Ann Arbor, MI, USA

4.1 Introduction

A key issue for panel surveys is the relationship between changes in respondent burden and resistance or attrition in future waves. In particular, does asking respondents to participate in longer or more frequent interviews, additional study components, or more demanding activities have negative effects on panel attrition, fieldwork effort, and survey costs? Of course, any negative effects may be balanced by the benefits of collecting valuable additional data, and hence may be worth pursuing. Among the potentially most burdensome activities for panel respondents is participation in a supplemental study, which may require a separate interview, interaction with other family members (such as children or a spouse), a home visit or other new mode, or providing new types of data (such as biological samples). Respondents may perceive the burden to be larger if the topic, content, and nature of the supplemental study diverge significantly from that of the main study. On the other hand, some respondents may enjoy the opportunity to participate in a supplement – or, at least, may appreciate the additional incentives.

An understanding of the effects of a supplemental study invitation on panel survey outcomes can be useful in several ways. First, knowledge of the potential effects of supplemental studies permits an informed assessment of the trade-off between the benefits and costs of these supplemental studies and can help guide decision-making about whether to launch a supplemental study. Second, comparing outcomes for cases that participated in a supplemental study with those for cases that were not invited may provide some insight into the effect of participating, though such evidence would only be suggestive because of the endogenous nature of the decision to participate among those invited to the supplemental study.

Advances in Longitudinal Survey Methodology, First Edition. Edited by Peter Lynn.

Companion website: www.wiley.com/go/lynn/advancesinlongitudinalsurvey

A major challenge in analysing the effects of supplemental studies on panel outcomes is that the offer to participate is rarely randomised to provide a clear comparison group (however, see Chapter 5 in this book for an exception). Although this presents a challenge for our analysis, we use several complementary approaches to create appropriate comparison groups to those selected for the supplement.

This study uses data from multiple waves of the Panel Study of Income Dynamics (PSID) from 1997 to 2015 to examine the effects on attrition and on various other measures of respondent cooperation of being invited to take part in a major supplemental study to PSID, namely the 1997 PSID Child Development Supplement (CDS). In the next two sections we describe our conceptual framework and previous research. We then describe the data and methods. We present our results next, and then end with our conclusions.

4.2 Conceptual Framework

Groves and Couper (1998) describe a generalised model of survey participation that focuses on the initial decision to participate in a survey but is nevertheless relevant to the process of deciding whether to participate in each wave of a panel study. The choice to participate in a survey interview is characterised by Groves and Couper (1998) as being based on heuristic decision-making, rather than by deep, thoughtful consideration of the pros and cons of participation. The heuristics include *reciprocation* (related to the heuristic of social exchange); *authority* (counterbalanced by social isolation, whereby people who feel socially excluded may have less regard for authority); *consistency* (doing what you did before); *scarcity* (perceiving participation as a rare opportunity); *social validation* (being more likely to participate if you think others like you are also participating); and *liking* (connecting with the interviewer). In addition, salience, relevance, and interest in the study topic are likely to influence participation. Each of these factors is likely to be relevant to: (i) the decision to continue to participate in subsequent waves of PSID after having been asked to participate in CDS; (ii) the level of effort required by interviewers to complete the interviews with respondents; and (iii) respondents' behaviour in response to the interview request even if they ultimately participate in the interview.

Guided by these heuristics, Groves and Couper (1998) develop three hypotheses. The first is that people are less likely to participate in a survey when opportunity costs are higher. In our analysis, we control for the effects of variables that reflected higher opportunity costs associated with participating in CDS and subsequent waves of PSID, such as larger family size, higher income, and being unmarried. The second hypothesis, based on social exchange, suggests that an equitable relationship between respondents and the study sponsor or representative is

likely to lead to higher response rates; in addition, social exchange is influenced by respondent incentive payments. In our analysis, we expect that families of low socioeconomic status are more likely to value the social exchange of participating in CDS and, therefore, have a greater likelihood of participating in future waves of PSID. The third hypothesis is that individuals who are socially isolated are less likely to be persuaded to participate in a study out of a sense of obligation, duty, or belonging. We use several respondent characteristics as indicators of social isolation, including being unpartnered and not having children.

The leverage-saliency theory (Groves et al. 2000), which extends Groves and Couper's original conceptual framework, suggests that survey design features have different leverage on the decision to cooperate for different individuals. For instance, respondents' interest in the survey topic might increase their likelihood of participation. One example is provided by Barber et al. (2016), who found that respondents in a longitudinal study using weekly web surveys who experienced the behaviours measured by the study maintained higher participation levels than respondents who did not experience those behaviours. We consider the effects of survey-related variables that reveal otherwise-hidden propensity to participate. These variables indicate whether a person was previously a non-respondent and include the sampling weight as an (inverse) indicator of cumulative attrition in the past of individuals with similar characteristics. More generally, previous research shows that characteristics of respondents affect subsequent survey behaviour and outcomes in panel studies (e.g. Fitzgerald et al. 1998; Groves 2006; Lugtig 2014; Lugtig et al. 2014).

The relevance of the extended Groves-Couper framework for our analysis is reflected in the crucial need for us to adequately control for these various types of respondent characteristics when comparing subsequent panel outcomes between individuals who were and were not asked to participate in the supplemental study. To the extent that the invitation to participate – which was not randomised – reflects individuals' behaviour and characteristics, it is important that our analysis adjusts for or removes the effects of these factors. We use regression analysis and inverse probability of treatment weights to control for these factors in our analysis, and also exploit a discontinuity in selection based on children's ages.

Once we incorporate an appropriate set of controls for selection into the supplemental study, we hypothesise that the invitation to participate in the supplemental study will lead to higher attrition and, among those who do not attrit but instead continue in the study, with greater fieldwork effort needed to contact, track, and persuade respondents to participate in panel interviews in subsequent waves. We conceptualise these worse subsequent fieldwork outcomes to be the result of the higher burden associated with completing interviews for the supplemental study, especially when we refine our comparison group to observationally similar panel respondents who just missed being eligible for participation in the supplemental

study. The magnitude of the negative effect of participation is difficult to predict because respondents often enjoy the interviews and receive a significant financial incentive for participating. We also conceptualise a negative effect of the invitation that operates regardless of the burden of the supplemental study, which arises because the supplemental study focuses on a different topic – children's development – than the focus on family economics of the main study in which respondents originally agreed to participate. Such a switch in topic could lead to respondents reassessing their decision about participating in the original study and deciding to end their participation.

4.3 Previous Research

Although a number of studies have examined the effects of survey experiences on subsequent participation in the context of panel studies, there are only two studies of which we are aware (aside from Pashazadeh et al. in Chapter 5 of this book) that have considered whether participation in between-wave supplemental surveys affects participation in an ongoing panel study. First, Ofstedal and Couper (2008) examined the impact of supplemental requests on panel non-response in the US Health and Retirement Study (HRS). The overall findings were that most supplemental requests had no significant effects on subsequent panel attrition in HRS. Supplements with high topic relevance had positive effects on subsequent panel retention (an internet survey and diabetes mail survey).

Second, Deeg et al. (2002) assessed whether differential inclusion in a variety of supplemental studies affected participation in subsequent waves of the Longitudinal Aging Study Amsterdam (LASA) among a sample of adults aged 55–85 years at baseline in 1992 ($N = 3805$). Supplemental study topics included health, social networks, widowhood, and depression. The authors' main conclusion is that the risk of attrition from the main panel study is increased by approaching respondents to participate in a supplemental study. Further, the authors note that respondent burden of the supplement (as measured by questionnaire length, effort, and subject matter) was unrelated to subsequent attrition.

In a related study, Kantorowitz (1998) reported that in the Israeli Labour Force Survey conducting either 'easy' or 'heavy' supplements, in contrast to having no supplement, was not associated with higher levels of non-response in the next cycle. Phillips et al. (2005) tested two conditions to assess whether a request to complete a supplemental questionnaire influenced the likelihood that a respondent would complete the primary questionnaire. In the first condition, respondents were asked to complete the secondary questionnaire depending on their responses to items in the primary questionnaire. In the second condition, which was intended to impose respondent burden, respondents were asked to complete

the supplemental questionnaire unconditionally. Those in the latter group were less likely to complete either questionnaire, but among those who did, response rates on the supplemental questionnaire were higher compared to the screener condition. McCarthy et al. (2006) considered whether frequency of contact and cumulative interview length affected the likelihood of participating in agricultural surveys sponsored by the US Department of Agriculture's National Agricultural Statistics Service. In this case, the agency fields frequent surveys on a variety of topics in a relatively small population. As a result, individual agricultural operations have a relatively high probability of being selected into multiple, independent surveys over time. Frequency of contact and cumulative interview length over a three-year period were not routinely associated with refusal to participate in subsequent interviews.

There is related research that considers whether interview length and other measures of respondent burden affect subsequent panel attrition (e.g. Lynn 2014; Hart et al. 2005; Phillips et al. 2005; Porter et al. 2004; Rolstad et al. 2011). For instance, Sinibaldi and Karlsson (2016) identified individuals selected into more than one sample for all Statistics Iceland general population household surveys over a 12-year period to examine whether the decision to participate in a second survey is influenced by the amount of time since the first survey. The results show a weak linear positive effect of length of time since the first survey on participation in the second survey, but this is explained by both demographic characteristics of respondents and survey indicators. Overall, these studies present a mixed set of findings suggesting that the survey burden generally does not affect subsequent attrition but may increase it under some limited circumstances.

In summary, previous research suggests that an invitation to participate in a supplemental study may cause respondents to recalibrate the perceived costs to participating in a panel study. Those who determine that the cost of participating in the supplement is too great, or who were already considering ending their participation in the panel study, will withdraw at the point of being invited to the supplementary study, while those who do participate will remain committed to future cycles of the supplementary study as well as the panel study in which it is embedded. Although not reviewed here, salience of the study topic and appeals to respondents' unique value also may help to retain respondents selected for supplemental studies.

4.4 Data and Methods

The PSID is the world's longest-running household panel survey. It began in 1968 and collects nationally representative data for the United States through interviews conducted annually through 1997 and biennially thereafter

(McGonagle et al. 2012). One adult is interviewed in each household; respondents report information about themselves, their spouse/partner, and all other family members. PSID has achieved response rates of 95–98% for the continuing panel in most years. PSID has a number of supplemental studies, which began in 1997 with the original CDS. CDS collected information on up to two randomly selected children ages 0–12 years and their caregivers in 2380 PSID families, including detailed information on health, skills, behaviour, time use, parenting and the home environment, and many related topics (McGonagle and Sastry 2015). Interviews and assessments with children were conducted during in-home visits, and with caregivers using in-person visits and the telephone. Two additional rounds of CDS interviewers were conducted in 2002 and 2007.

Our analysis focuses on the effects of selection for Wave I of the original CDS in 1997 on PSID outcomes in subsequent years. We analyse outcomes for a focal PSID adult sample member in each PSID family who is either the household head or the head's spouse. Most PSID households have just a single such person; however, in the small number of households in which there are two such individuals, we select one of these individuals at random as our focal PSID adult sample member.

The treatment of interest is whether the PSID adult sample member lived in a family unit in which one or more children were selected to participate in the first wave of CDS in 1997. The main treatment indicator does not distinguish between whether or not the family participated in CDS, and hence represents an intent-to-treat (ITT) indicator. An ITT analysis is appropriate for addressing our first research question about whether the *invitation* to participate in the supplemental study is an important determinant of panel attrition. Because there is higher panel attrition among non-compliers (i.e. CDS non-respondents), the ITT approach provides an upper bound on the effects of treatment on the treated. In order to examine the effects of actual participation, our second research question, we need to contrast subsequent panel outcomes of these two groups (CDS respondents and non-respondents) with each other and with the control group – while acknowledging that the decision to respond in CDS is endogenous. We cannot directly assess the effects of CDS participation on subsequent attrition without considering the non-compliers, which do not belong with the control group (because they received the invitation to participate) and, at the same time, should not be omitted from the analysis (because they were present at baseline and hence were eligible for the treatment). We use an ITT approach that distinguishes between compliers and non-compliers as a way to gain insights into the effects of actual participation in the absence of an appropriate causal analysis approach to adequately control for the participation decision. Note that we do not separately examine the effects of subsequent CDS-related participation, which includes either one or two additional waves of CDS in 2002 and 2007 and

participation in the biennial PSID Transition into Adulthood Supplement (TAS) from 2005 onwards.

The main outcome of interest for this analysis is the post-1997 attrition of the focal PSID adult sample member. PSID classifies a household as a permanent panel refusal if the household is non-response for two consecutive waves. We treat such cases as having attrited at the first wave in which they do not respond. In addition, a single clear and explicit request by a respondent to be removed from the ongoing sample can also lead to a permanent panel refusal. A PSID adult sample member can leave the study through death and can miss a wave through being declared ineligible following institutionalization or moving into another panel household. Observations are censored at the time of death for deceased sample members and are omitted for waves in which sample members were classified as ineligible.

A second set of outcomes we examine are five indicators of fieldwork difficulty associated with completing interviews in each wave for continuing sample members. These indicators capture the effort required at various fieldwork stages to contact and initiate or to complete an interview (the number of telephone calls to complete an interview and whether a face-to-face visit was necessary), respondent cooperation with the interview request (any resistance to completing an interview and any interview suspension that requires the interview to be completed in two or more calls for any reason), and respondent residential mobility (whether any tracking was required in order to find a sample member and conduct an interview).

Tables 4.1 and 4.2 present the outcome data for our analysis. We begin with 6308 observations in 1997, as shown in Table 4.1, which comprise all households that completed a PSID interview in that year. The attrition rate averaged 3.5% per wave over the subsequent eight biennial waves while ineligibility rates averaged 0.5% per wave. Table 4.1 shows that we have a total of almost 44 000 person-wave observations for our attrition analysis. Table 4.2 shows that the average number of telephone calls needed to complete an interview increased dramatically from 1997 to 2015, almost tripling between 1997 (when an average of 6.0 calls were needed to complete an interview) and 2015 (when 16.2 calls were needed). The fraction of the sample receiving face-to-face visits or exhibiting any resistance both followed a U-shaped pattern of decline and then increase, ending the period at a similar level as at the beginning. Interview suspensions followed an inverse U-shaped pattern, as did tracking rates.

Summary statistics for the covariates used in our analysis are presented in Table 4.3. The main independent/treatment variable is the family's CDS status in 1997. Approximately 40% of PSID families had a sample member eligible for CDS, and 87% of these families participated in the CDS survey.

The next set of covariates in Table 4.3 describes various sample characteristics. The PSID sample has several sources: the original 1968 PSID sample came from

Table 4.1 PSID sample observations by biennial wave for 1997–2013 among the 1997 baseline sample.

Year	Interviewed	Attrited	Ineligible	Died	Total
1997	6308	–	–		6308
	100.0%	–	–		100.0%
1999	5985	249	29	45	6308
	94.9%	3.9%	0.5%	0.7%	100.0%
2001	5770	174	20	50	6014
	95.9%	2.9%	0.3%	0.8%	100.0%
2003	5560	161	36	33	5790
	96.0%	2.8%	0.6%	0.6%	100.0%
2005	5382	156	30	28	5596
	96.2%	2.8%	0.5%	0.5%	100.0%
2007	5145	190	31	46	5412
	95.1%	3.5%	0.6%	0.8%	100.0%
2009	4940	174	28	34	5176
	95.4%	3.4%	0.5%	0.7%	100.0%
2011	4696	204	29	39	4968
	94.5%	4.1%	0.6%	0.8%	100.0%
2013	4463	210	24	28	4725
	94.5%	4.4%	0.5%	0.6%	100.0%
Total	41941	1518	227	303	43989
	95.3%	3.5%	0.5%	0.7%	100.0%

Note: The first column includes all interviewed cases. Cases were coded as attrited in the first wave they were non-response, with attrition occurring after two successive waves as non-response. Ineligible cases comprise respondents who were institutionalised or joined another sample family unit as a non-head/non-spouse. The last column is the row sum and is also equal to the sum of cases from the previous wave that were either interviewed or ineligible.

the Survey of Economic Opportunity (SEO sample) or a nationally representative sample frame maintained by the University of Michigan's Survey Research Center (SRC sample). The CDS sample also included new immigrants that were added to PSID in 1997. Approximately one-quarter of the 1997 PSID sample were from the SEO sample, two-thirds were from the SRC sample, and the remaining 7% were from the new immigrant refresher. The final sample-related variables describe whether the focal respondent did not respond to any prior wave of PSID, whether the family unit had recently split from another PSID family to form a new family

Table 4.2 Summary statistics for PSID fieldwork outcomes.

Year	Continuing sample	Number of calls	Any face-to-face	Any resistance	Interview suspension	Any tracking
1997	6308	6.0	–	–	–	–
1999	5985	7.0	–	–	–	–
2001	5770	7.3	–	–	–	–
2003	5560	9.2	11.6%	4.0%	–	10.6%
2005	5382	9.4	11.6%	3.3%	–	12.7%
2007	5145	9.5	8.4%	3.0%	–	16.3%
2009	4940	11.0	8.9%	3.0%	22.9%	16.8%
2011	4696	11.3	3.4%	3.1%	39.4%	15.7%
2013	4463	11.9	5.3%	3.4%	29.9%	18.3%
2015	4145	16.2	9.2%	5.0%	29.5%	8.7%

Note: Number of calls is the count of telephone calls made to respondents. No data were collected prior to 2003 for face-to-face visits, resistance, or tracking and prior to 2009 for interview suspensions

unit, the type of respondent who was interviewed in 1997 (most often the head or the spouse, but occasionally the head's partner or another individual), and the family's PSID sample weight.

The next set of variables describes demographic and socioeconomic characteristics of the focal respondents and their families in 1997, including the person's age, sex, race or ethnicity, marital status, education, income, rural–urban place of residence, and whether the family had moved since the last interview. The average age in 1997 was 44 years, and well over half of focal respondents were female. Due to the presence of the SEO sample, PSID includes an oversample of poor families and African Americans – with the latter group accounting for one-in-three sample members. Nearly two-fifths of families were headed by an unpartnered individual, with the remainder either married (55%) or cohabiting (6%). The average education level was 12.8 years of schooling, and one-quarter of families resided in rural areas. One-in-five families had moved since the previous interview. Not shown in the table (or included in our analysis) is the fact that just under half of PSID families in 1997 had children aged 0–17 years.

To describe and analyse the effects of CDS on sample attrition in PSID, we use survival curves and univariate and multivariate discrete time hazard models. Survival analysis is the unifying approach, which allows us to examine if a respondent attrited ($y = 1$) or did not attrit ($y = 0$) from PSID in each wave from 1999 to 2013,

Table 4.3 PSID 1997 baseline sample characteristics.

Variable	Percent or mean (std. dev.)
1997 CDS-I status	
Not selected	60.8%
Response	34.1%
Non-response	5.2%
PSID sample source	
SEO	24.8%
SRC	68.2%
Immigrant	7.0%
Nonresponse prior to 1997	
No	92.2%
Yes	7.8%
Family unit split off in 1997	
No	97.1%
Yes	2.9%
Respondent in 1997	
Head	68.3%
Spouse	28.4%
Partner	1.8%
Other family member	0.7%
Proxy, not family member	0.8%
Family sample weight	24.91 (17.39)
Age (years)	44.02 (16.29)
Sex of Head	
Female	58.1%
Male	41.9%
Race/Ethnicity of Head	
White	61.0%
Black	29.4%
American Indian	0.5%
Asian	1.7%
Hispanic	4.5%
Other	2.9%

(Continued)

Table 4.3 (Continued)

Variable	Percent or mean (std. dev.)
Marital status of head	
Unpartnered	38.5%
Married	55.4%
Cohabiting	6.1%
Education of head (years)	12.76 (2.67)
Income-to-needs ratio 1996	3.65 (4.10)
Place of residence in 1997	
Rural	26.0%
Urban	74.0%
Family moved since last interview	
No	80.6%
Yes	19.4%
Observations	6308

while appropriately accounting for observations that were censored due to reaching the end of the observation period or missing a wave due to being ineligible.

To describe and analyse the fieldwork outcomes, we use linear regression panel models for the number of telephone calls to complete an interview in each wave and logistic regression panel models for the remaining binary dependent variables ($y = 1$ if a respondent had a face-to-face visit, any resistance, an interview suspension, or any tracking in each wave; $y = 0$ otherwise).

The key methodological challenge is to control for differences between families that were and were not selected for CDS, which is complicated by the fact that all families with a child in the target age range of 0–12 years were selected for CDS. The controls amount to achieving balance between families with and without children in a particular age range, and hence represent a family-level model for childbearing and, in particular, the timing of childbearing. We use three complementary approaches to address this issue. First, we use regression analysis to control for the observed covariates listed in Table 4.3. Second, we use inverse probability of treatment weights, based on a propensity score for treatment – i.e. being selected for CDS – that is estimated using logistic regression to model whether a household was eligible ($y = 1$) or was not eligible ($y = 0$) for participation in CDS. We estimate a propensity model for the full sample, as well as separate models for comparisons based on sub-samples described below. Third, we stratify the

CDS and non-CDS samples and compare outcomes among sub-samples that are substantially more similar to each other based on the ages of their children, essentially providing a form of discontinuity analysis. In particular, we narrow the comparisons to focus on observationally-similar families who just missed eligibility for CDS participation due to: (i) their oldest children being ages 0–2 years in 1999 and hence born just after the youngest children selected for CDS and (ii) their youngest children being ages 13–17 years in 1997 and thus born just before the oldest children selected for CDS. We compare these two non-CDS sub-samples with CDS families having the youngest and oldest CDS children, respectively, as well as with all CDS families. We apply each of these three methods on their own and in combination.

Four logistic regression models of selection for CDS were estimated in order to construct the inverse probability of treatment weights for our main analysis. The first is estimated on the full PSID sample. The second model is restricted to families with children (all CDS families as well as non-CDS families whose youngest child was aged 13–17 years in 1997 and non-CDS families whose oldest child was aged 0–2 years in 1999, with a further restriction for all groups to those who responded in 1999). The third model is estimated on families with younger children (CDS families with a child aged 0–2 years in 1997 and who responded in 1999 and non-CDS families whose oldest child was aged 0–2 years in 1999), and the fourth on families with older children (CDS families with a child aged 10–12 years in 1997 and non-CDS families whose youngest child was aged 13–17 years in 1997). In all cases, the models fit the data well and reveal several covariates that consistently predict selection for CDS, including being a split-off family, having a younger household head, Hispanic ethnicity, being married, having higher family income, and living in an urban area. The inverse probability of treatment weights was constructed from the propensity score, which is the predicted probability from these models. Full results of the models are presented in Part A of the online supplementary material. The predicted probabilities of selection for CDS of those sample members who were selected and those who were not selected for CDS in each of the four sub-samples corresponding to the four different logistic regression models are compared in Figure A.1. Overall, the results show that there are indeed PSID sample members who, based on observed covariates, appear to have had a reasonably high likelihood of being selected for CDS, even though they were, in fact, not selected. The least overlap between the CDS and non-CDS samples is observed in the first model, based on the full PSID sample, which reflects the inclusion of a large number of families that had no children eligible for CDS and little likelihood of having such children due to their age, marital status, and other basic demographic factors. However, these families would receive low weights, and hence have only small influence on the results. The other three panels in Figure A.1 show considerable overlap between

families that were and were not selected for CDS, and, in particular, a significant proportion of cases that were not selected for CDS but that had a likelihood of greater than 50% of being selected. Overall, the results suggest that the inverse probability of treatment weights should perform reasonably well in improving the comparability of those who were and were not invited to participate in CDS.

4.5 Results

Our findings are presented in Tables 4.4 and 4.5 and Figures 4.1 and 4.2. Figure 4.1 shows observed group differences in attrition over the 16-year observation period. The panel in the top left compares all CDS families with all non-CDS families, and shows that families selected for CDS are substantially less likely to attrit from PSID over the observation period. This result is affected by compositional differences between the two groups, which in the top-right panel we partially control for by restricting attention to non-CDS families with either a slightly younger or slightly older child. The top-right panel reveals that respondents in non-CDS families with a child were slightly less likely to attrit from PSID than CDS families. However, the lines are close together, and the 95% confidence intervals indicate that the difference in attrition is not statistically significant. The bottom two panels of Figure 4.1 present comparisons focusing separately on younger children (bottom left) and older children (bottom right). In both cases, individuals from families not selected for CDS were less likely to attrit from PSID. The differences are not statistically significant for either group, although the difference appears larger among families with the youngest children.

The results from the figure are replicated in the first row of Table 4.4, which shows the estimated effect of CDS selection on subsequent attrition from PSID without controlling for any observed covariates. Column 1 compares all 6308 individuals based on CDS vs. non-CDS status, and indicates that likelihood of attriting from PSID is $(1 - 0.502 = 0.498)$, or approximately 50% lower for individuals whose families were selected for CDS. This result is statistically significant at the 0.001 level. Column 2 restricts the comparison to all CDS families and non-CDS families with a slightly younger or older child, and reveals that CDS selection was associated with marginally higher attrition – although this effect is not statistically significant. The same finding emerges for the remaining two comparisons in Row 1: in Column 3 between CDS families with children aged 0–2 years in 1997 and non-CDS families with children aged 0–2 years in 1999 and in Column 4 between CDS families in 1997 with children aged 10–12 years and non-CDS families with children aged 13–17 years. Sample sizes are much smaller for the more focused comparisons in Columns 3 and 4, but the point estimates for the CDS effect are close to unity.

Table 4.4 Discrete time hazard model regression results for the effects of CDS selection on subsequent PSID attrition, 1999–2013.

	Comparison			
	(1)	**(2)**	**(3)**	**(4)**
Model 1 (no covariates)				
CDS selection	0.502*** (0.030)	1.042 (0.120)	1.120 (0.282)	1.022 (0.202)
Model χ^2 (7 df)	187.48***	29.89***	16.01*	9.64
Model 2 (all covariates)				
CDS selection	0.949 (0.116)	1.043 (0.158)	2.148# (0.938)	0.824 (0.199)
Model χ^2 (37 df)	1218.91***	149.94***	60.23**	77.90***
Model 3 (inverse probability of treatment weights)				
CDS selection	0.863 (0.242)	0.956 (0.169)	1.077 (0.308)	0.984 (0.222)
Model χ^2 (7 df)	125.01***	20.68**	22.43**	6.098
Model 4 (inverse probability of treatment weights and all covariates)				
CDS selection	0. 919 (0.186)	0.989 (0.171)	1.954 (0.903)	0.921 (0.251)
Model χ^2 (37 df)	535.87***	148.77***	100.37***	81.86***

(*continued*)

Table 4.4 (Continued)

	Comparison			
	(1)	**(2)**	**(3)**	**(4)**
Model 5 (inverse probability of treatment weights and all covariates)				
CDS response	0.628# (0.156)	0.822 (0.145)	1.861 (0.857)	0.698 (0.204)
CDS non-response	4.005*** (1.438)	2.235*** (0.442)	2.420 (1.374)	3.756** (1.907)
Model χ^2 (38 df)	528.68***	223.90***	104.40***	99.56***
Comparison groups				
Non-CDS	All non-CDS FUs	Non-CDS FUs with children just younger and older	Non-CDS FUs with children just younger (age 0–2 years)	Non-CDS FUs with children just older (age 13–17 years)
CDS	All CDS FUs	All CDS FUs	CDS FUs with youngest children (age 0–2 years)	CDS FUs with oldest children (age 10–12 years)
Person-periods of observation				
Non-CDS	25 392	4410	969	3441
CDS	17 921	17 921	2711	1893
Total	43 313	22 331	3680	5334

Note: Parameters are relative risks; standard errors in parentheses; *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; # $p < 0.10$; FU: 'family unit.' Model 1 covariates are survey year and CDS selection; Models 2–5 add the covariates listed in Table 4.2 plus squared terms for age, sample weight, years of education, and income-to-needs ratio. The inverse probability of treatment weights are derived from the predicted probabilities of logistic regression models of participation in CDS-I with the same set of covariates as used in Models 2–5; separate logistic regression models were estimated and separate inverse probability of treatment weights were constructed for the different comparison groups (see the online supplementary material).

Table 4.5 Panel data regression model results for the effects of CDS selection on PSID fieldwork outcomes.

	Comparison			
	(1)	**(2)**	**(3)**	**(4)**
A. Number of Telephone Calls to Complete a PSID Interview, 1999–2015				
Model 1 (no covariates)				
CDS selection	2.968*** (0.128)	1.878*** (0.215)	−0.041 (0.576)	1.201*** (0.343)
Model 2 (inverse probability of treatment weights and all covariates)				
CDS selection	0.041 (0.330)	0.351 (0.376)	1.099 (0.869)	0.212 (0.399)
B. Any Face-to-Face Visits to Complete PSID Interview, 2003–2015				
Model 1 (no covariates)				
CDS selection	1.562*** (0.080)	1.349*** (0.118)	0.943 (0.153)	1.113 (0.181)
Model 2 (inverse probability of treatment weights and all covariates)				
CDS selection	0.870 (0.121)	1.189 (0.157)	1.223 (0.381)	0.947 (0.199)
C. Any Resistance to Completing a PSID Interview, 2003–2015				
Model 1 (no covariates)				
CDS selection	1.364*** (0.080)	1.311** (0.134)	0.919 (0.173)	0.971 (0.191)
Model 2 (inverse probability of treatment weights and all covariates)				
CDS selection	1.319# (0.209)	1.344# (0.216)	1.052 (0.368)	0.753 (0.163)

(*continued*)

Table 4.5 (Continued)

	Comparison			
	(1)	**(2)**	**(3)**	**(4)**
D. Any Interview Suspension when Completing a PSID Interview, 2009–2015				
Model 1 (no covariates)				
CDS selection	1.384*** (0.051)	1.241*** (0.076)	0.811# (0.097)	1.214 (0.132)
Model 2 (inverse probability of treatment weights and all covariates)				
CDS selection	1.465*** (0.159)	1.382*** (0.138)	0.901 (0.189)	1.054 (0.145)
E. Any Tracking to Complete a PSID Interview, 2003–2015				
Model 1 (no covariates)				
CDS selection	1.552*** (0.048)	1.586*** (0.089)	1.424** (0.156)	1.355** (0.129)
Model 2 (inverse probability of treatment weights and all covariates)				
CDS selection	1.034 (0.095)	0.963 (0.094)	1.003 (0.186)	1.007 (0.124)
Comparison groups				
Non-CDS	All non-CDS FUs	Non-CDS FUs with children just younger and older	Non-CDS FUs with children just younger (age 0–2 years)	Non-CDS FUs with children just older (age 13–17 years)
CDS	All CDS FUs	All CDS FUs	CDS FUs with youngest children (age 0–2 years)	CDS FUs with oldest children (age 10–12 years)

Note: Linear regression coefficients reported in Panel A and odds ratios reported in Panels B–E; standard errors in parentheses. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; # $p < 0.10$. Covariates for Model 1 are survey year and CDS selection; for Model 2 are survey year, CDS selection, and the variables listed in Table 4.2 plus squared terms for age, sample weight, years of education, and income-to-needs ratio. See text for description of the inverse probability of treatment weights. Person-period observations for Comparisons 1–6 are, respectively, 52 355, 27 313, 4624, and 6506 for Panel A; 34 382, 18 402, 3107, and 4378 for Panels B, C and E; and 18 307, 10 039, 1681, and 2385 for Panel D.

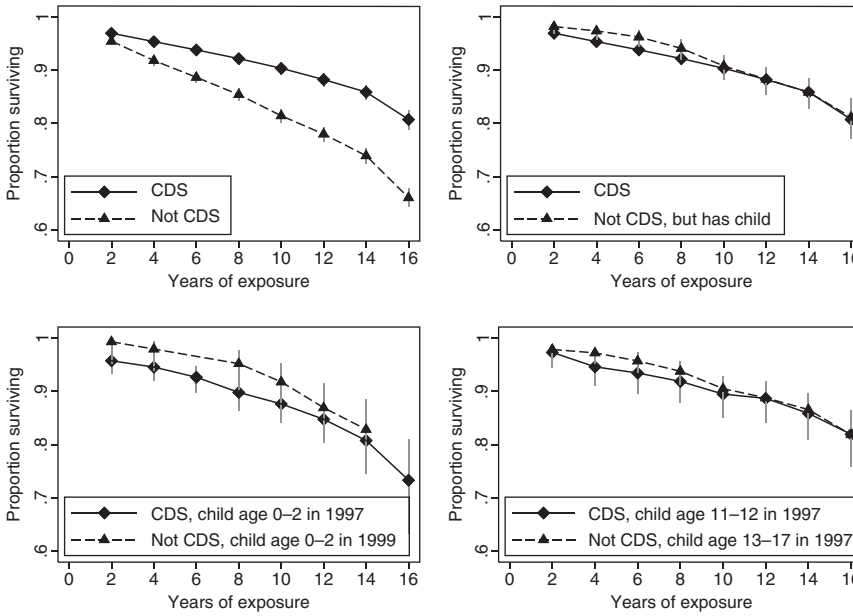

Figure 4.1 Observed trends in PSID response by CDS selection status for full sample and sub-samples with children, 1997–2013. Note: Clockwise from top left: full sample in 1997; family units with children aged 0–17 years in 1997 or 0–2 years in 1999; family units with children aged 11–17 in 1997; family units with children aged 0–2 years in 1997 or 1999. Observed trends are shown, without any covariate controls or weights.

The results for Model 2 in Table 4.4 include covariate controls for observed respondent characteristics. The results in the first column change substantially, with the apparent lower attrition for individuals from CDS families being accounted for entirely by observed characteristics and rendering statistically insignificant the difference in their likelihood of attrition. The results in Columns 2–4 continue to show no statistically significant differences in the effects of CDS participation in subsequent attrition – with the exception of the results in Column 3 which suggest that, compared to individuals in families with children just younger than the youngest children in CDS, the likelihood of attriting from PSID is 2.15 times higher, an effect that is statistically significant at the 0.10 level.

The results for Model 3 in Table 4.4 use the alternative approach of inverse probability of treatment weighting to control for the likelihood of CDS selection. This approach provides a near-uniform set of findings suggesting that CDS selection leads to no change in attrition from PSID in subsequent years. Model 4 in Table 4.4 extends these results to include covariate controls in addition to inverse probability of treatment weights, providing us with 'doubly-robust' estimates of the effects

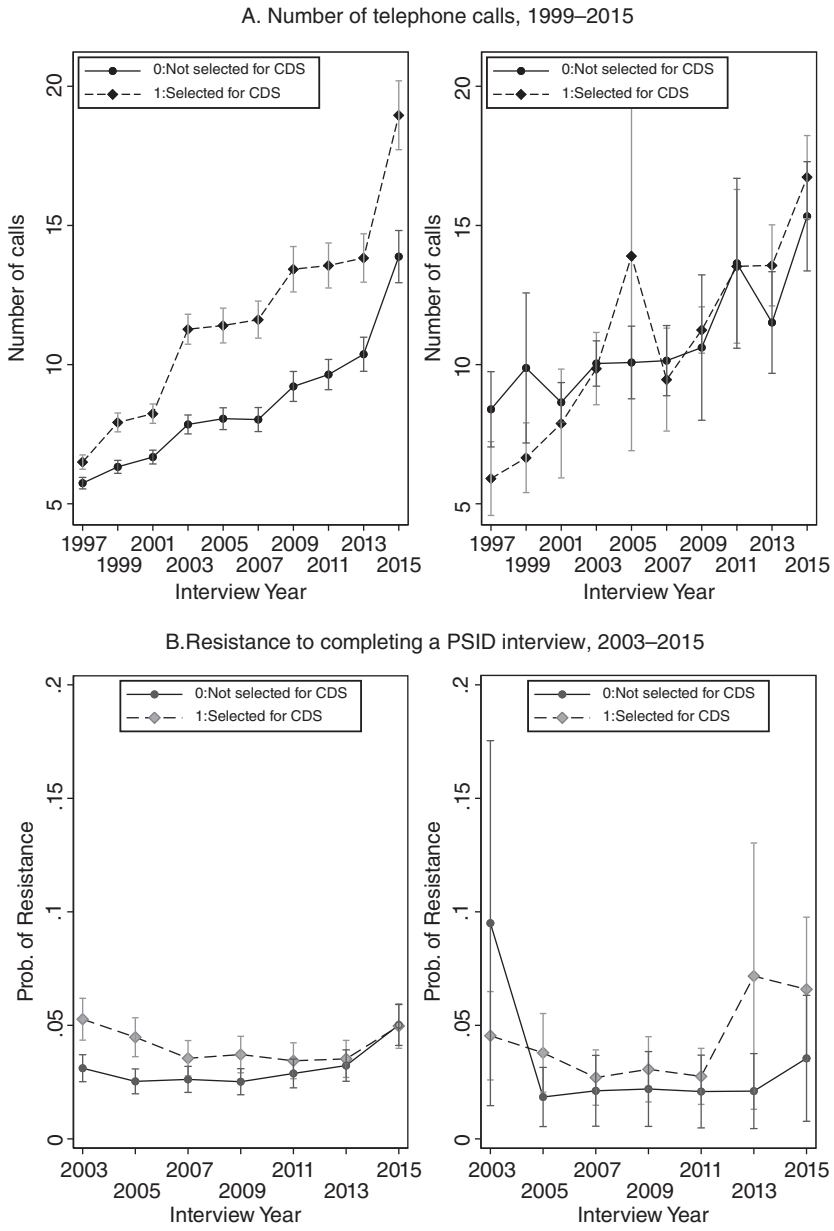

Figure 4.2 Observed and adjusted trends in PSID fieldwork outcomes by CDS selection. A. Number of telephone calls, 1999–2015. B. Resistance to completing a PSID interview, 2003–2015. C. Tracking to complete a PSID interview for full PSID sample, 2003–2015.

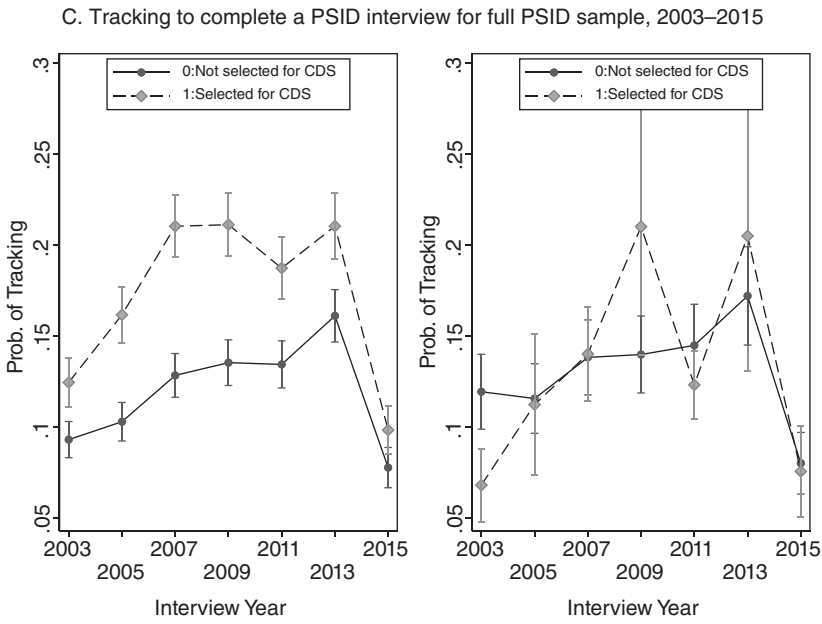

Figure 4.2 (*Continued*)

of CDS participation on subsequent attrition (Robins et al. 1995; Lunceford and Davidian 2004). These are our preferred results because of the comprehensive and flexible way in which they control for differences between observationally similar sample members who were and were not selected to participate in CDS. The results from Model 4 indicate that CDS selection is not associated with a change in attrition. However, there is a suggestion that CDS selection may be associated with higher attrition for families with the youngest children, although the higher attrition estimate is not statistically significant.

Finally, the results for Model 5 in Table 4.4 compare CDS respondents and non-respondents with those not selected for CDS in terms of propensity for subsequent attrition from PSID. These results are based on the modelling approach used for Model 4, which combines the inverse probability of treatment weights with covariate adjustment. The findings from Model 5 suggest that those who participated in CDS had a similar risk of subsequent attrition from PSID to those who were not selected for CDS; however, CDS non-respondents had a higher subsequent attrition risk, except in the comparison based on families with the youngest children. The relative risks of subsequent attrition range between 2.2 and 4.0 times higher for CDS non-respondents compared to those who not were not selected for CDS.

For each attrition model, we also examined whether there were time-varying effects of CDS selection on subsequent attrition. However, we found no evidence based on non-significant results of a joint statistical test of interactions between wave of interview and CDS selection status.

We next present results, summarised in Table 4.5, of the effects of CDS selection on the five fieldwork outcomes other than attrition (number of telephone calls, any face-to-face visits, any resistance, any interview suspension, and any tracking). We examine the four sample comparisons from the preceding set of results, but focus on two model specifications. The first model specification includes no covariates, while the second is based on our preferred specification that incorporates inverse probability of treatment weights and all covariates. More extensive results are presented in Part B of the online supplementary material.

Focusing first on the observed relationships between CDS selection and other fieldwork outcomes, based on Model 1 in Panels A–E in Table 4.5 we see that for all outcomes, and for all comparisons, there is either a statistically significant deleterious (i.e. positive) effect of CDS selection on subsequent fieldwork outcomes or no statistically significant effect. For example, the effects of CDS selection on the number of telephone calls to complete a PSID interview based on the full sample (Comparison 1) and shown in Panel A indicates that families selected for CDS required three more calls to complete the PSID interview in each year. Results (not shown) suggest that the presence of time-varying effects of CDS selection on the number of calls to complete a PSID interview, with the higher number of calls needed to complete interviews with PSID respondents who were selected for CDS increasing modestly over time. Similarly, the results for any resistance to completing a PSID interview (shown in Panel C) for the full sample (Comparison 1) indicate a 36% increase in the likelihood of resistance as a result of CDS selection. A last example in Panel E shows the effects of selection for CDS on tracking in PSID; in this case, the likelihood of tracking for the full sample (Comparison 1) is raised by 55% as a result of CDS selection. Again, there is evidence in these descriptive findings of time-varying effects of CDS selection on each of these fieldwork outcomes, although for some outcomes the effects over time of CDS selection increase while for others they decrease; in all cases, the basic conclusions are not altered in terms of the direction of the effects.

When we adjust for covariates and incorporate inverse probability of treatment weighting, in all three of these illustrative cases there is no statistically significant increase in fieldwork difficulty associated with CDS selection. This overall set of results can be seen clearly in Figure 4.2, which shows the differences in probabilities for each of these three illustrative fieldwork outcomes (number of telephone calls, resistance to completing a PSID interview, and whether tracking was necessary) between those selected and not selected for CDS in the full sample based on the observed (left panel, Model 1) and adjusted (right panel, Model 2) results.

For all three of these fieldwork outcomes, CDS selection is associated with higher fieldwork effort in the unadjusted models but in the adjusted models there is no clear difference (which is confirmed by formal statistical tests).

For the outcome of any interview suspension (Panel D), we do find a statistically significant effect at the 0.001 level of CDS selection on increasing the likelihood of an interview suspension even after incorporating inverse probability of treatment weights and covariate adjustment for the first two comparisons that are based on all cases (an odds ratio of 1.47) and on families with children (an odds ratio of 1.38). However, the results are not statistically significant for the two more focused comparisons based on younger children (Comparison 3) and older children (Comparison 4); moreover, the point estimates for these latter two comparisons are close to unity, suggest the true absence of an effect for these two subgroups.

Looking across all of the adjusted results presented in Table 4.5, we find no compelling evidence that CDS selection is associated with more difficult fieldwork for PSID. The only statistically significant result is that CDS selection is associated with greater fieldwork effort for completing a PSID interview (Panel D: any interview suspension for Comparisons 1 and 2). For all other fieldwork outcomes and all other comparisons, there is no statistically significant adverse effect of CDS selection on PSID fieldwork outcomes.

Finally, we note that in Part B of the online supplementary material we present a more complete set of results of the effects of CDS selection on PSID fieldwork outcomes that includes examining differences in effects based on cases that were CDS response compared to those that were CDS non-response. There is some scattered evidence that CDS non-response cases had worse subsequent PSID fieldwork outcomes – such as greater resistance to completing a PSID interview for Comparisons (1) and (2) and a higher likelihood of interview suspension for Comparison (2). But these results are not consistent or compelling, and are balanced by an occasional finding that CDS non-response was associated with better fieldwork outcomes (e.g. less tracking, based on Comparison 3). In addition, we examined whether there was variation over time in the effects of CDS selection on PSID fieldwork outcomes that could have affected these results but uncovered no systematic pattern of time-varying effects.

4.6 Conclusions

In this analysis, we examined whether selecting PSID respondents to participate in CDS – a major inter-wave supplemental study that collected detailed information on children's development – resulted in negative effects on response rates and fieldwork outcome measures such as the number of telephone calls to complete an interview. Overall, our results suggest that asking PSID families with children to

participate in CDS resulted in a generally small, statistically insignificant increase in attrition in subsequent years but no consistent pattern of negative effects on fieldwork outcomes. Our preferred estimates are based on Model 4 in Table 4.4, which uses inverse probability of treatment weights and covariate adjustment to control for observed differences in factors associated with being selected for participation in CDS. The estimates from this model suggest a statistically significant negative effect on response rates only for families with the youngest CDS children. The reason for the higher attrition among families with the youngest children is not clear. This result represents our only statistically significant result and hence is not unexpected with a nominal Type-I error rate of 5%. Nevertheless, there are several reasons why families with very young children might view their CDS experience unfavourably and attrit at higher rates in subsequent PSID waves. For instance, the motivation for the study, which was to capture children's development in the school and neighbourhood contexts, may not be salient for children aged 0–2 years. The youngest CDS children also were asked to participate in an extra wave of data collection (CDS-III in 2007/08), which may have increased respondent burden (though the negative effect on attrition occurred before this date).

The conceptual framework we use suggests some reasons for these generally small and statistically non-significant results. We would expect to find higher attrition rates as a result of the higher costs for respondents from participating in additional study components. The fact that the attrition rates are only modestly higher, and are only statistically significant in one instance, is likely due to several of the factors identified by Groves and Couper – especially reciprocation and social exchange that is associated with respondent incentives and engagement in the on-going PSID and CDS studies. The estimated effects of CDS selection on PSID attrition were not changed much by controlling for a wide range of covariates describing sample status, prior history in the study, and demographic and socioeconomic characteristics. This finding suggests that the overall net effect of the other factors identified in the Groves and Couper framework is relatively unimportant for attrition over the subsequent eight waves of PSID. However, the detailed regression results (not shown) do reveal some statistically significant findings that support certain hypotheses from this framework.

Our results identify one particular group that is at high risk of attriting from the panel based on the request to participate in the supplement – namely, individuals from families who refused to participate in the supplement. This group was selected for CDS, but did not receive the treatment of actually participating in the supplement. Note, however, that our analysis was designed to analyse the effects of being *invited* to participate in CDS, rather than the effects of actual partition in the supplement (the former is the intent to treat, while the latter is the effects of the treatment on the treated). A likely reason for the higher attrition

among the CDS non-respondents is that this group was already predisposed to ending their participation in PSID, and hence the CDS refusal is probably a marker of this intention rather than a causal factor that explains their subsequent attrition. (This latter question could be examined by adopting the modelling approach used in this paper and, in particular, by viewing the CDS non-response decision as an endogenous 'treatment' and comparing the subsequent PSID attrition among these non-respondents to observationally similar CDS respondents, although it would be difficult to account for differences in unobservable factors.) Nevertheless, this finding suggests that, for panel studies, there would be value in identifying families who are predisposed to attriting and targeting them for special retention efforts. For instance, these families might be offered higher incentives or presented with study materials that highlight key facets of the survey that are positively associated with the decision to continue panel participation; alternatively, these families could be excluded from the supplement altogether.

We have used statistical methods for causal analysis to identify the effects of CDS participation on subsequent panel attrition and fieldwork outcomes – specifically, we used inverse probability of treatment weights calculated from the propensity score for CDS selection. This method provides a convincing approach to adjusting for the effects of observed covariates. But it is not an experimental approach and, in particular, it does not account for unobserved factors. However, the CDS design did not provide a mechanism for unobserved characteristics to directly affect the selection process; rather, unobserved factors, to the extent they played a role at all, were likely to have had subtler effects, such as through the timing of fertility and thus the direct comparability of families in the groups selected and not selected for CDS. In other words, the control group may not have observationally similar families to those selected for CDS because all CDS families chose to have their children in the same window based on factors related to the effects of age, period characteristics, and other variables.

Our analysis of CDS selection on other fieldwork outcomes, such as the number of telephone calls to complete an interview, also reveal that the invitation to participation in CDS did not have any lasting deleterious effects on fieldwork processes in PSID. The analytical design and statistical methods we used were important for revealing this result, because the observational patterns suggested that selection into CDS were associated with considerably worse outcomes for many different aspects of fieldwork.

Overall, our results suggest that a major supplemental study such as CDS has had, at most, relatively minor negative effects on attrition and fieldwork difficulties in PSID over a long follow-up period during which CDS respondents were asked to complete additional waves of data collection for CDS and its successor study that tracked these children into their young adult years. At the same time, the benefits of collecting these new data are significant, at least as measured by the

scientific contributions associated with over 600 publications based on these data according to the online PSID bibliography (available at www.psidonline.org). Our analysis provides support for continuing the CDS supplement to PSID, as well as an example of how the effects of supplemental studies can be evaluated for other panel studies.

Acknowledgements

Helpful comments were provided by participants at the 2016 Panel Survey Methods Workshop and the 2018 Methodology of Longitudinal Surveys II Conference, members of the PSID Board of Overseers, and Bob Schoeni. Programming assistance was provided by Meichu Chen. This project was supported by grant R01HD069609 from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD).

References

Barber, J., Kusunoki, Y., Gatny, H. et al. (2016). Participation in an intensive longitudinal study with weekly web surveys over 2.5 years. *Journal of Medical Internet Research* 18 (6): e105. PMCID: PMC4937177. DOI: https://doi.org/10.2196/jmir.5422.

Deeg, D.J.H., van Tilburg, T., Smit, J.H. et al. (2002). Attrition in the longitudinal aging study Amsterdam: the effect of differential inclusion in side studies. *Journal of Clinical Epidemiology* 55 (4): 319–328.

Fitzgerald, J., Gottschalk, P., and Moffitt, R. (1998). An analysis of sample attrition in panel data: the Michigan panel study of income dynamics. *Journal of Human Resources* 33 (2): 251–300.

Groves, R.M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly* 70 (5): 646–675.

Groves, R.M. and Couper, M.P. (1998). *Nonresponse in Household Interview Surveys*. New York: Wiley.

Groves, R.M., Singer, E., and Corning, A. (2000). Leverage-saliency theory of survey participation: description and an illustration. *Public Opinion Quarterly* 64 (3): 299–308.

Hart, T.C., Rennison, C.M., and Gibson, C. (2005). Revisiting respondent 'fatigue bias' in the National Crime Victimization Survey. *Journal of Quantitative Criminology* 21 (3): 345–363.

Kantorowitz, M. (1998). Is it true that nonresponse rates in a panel survey increase when supplement surveys are annexed? *ZUMA Spezial Nachrichten* 4: 121–135.

Lugtig, P. (2014). Panel attrition: separating stayers, fast attriters, gradual attriters, and lurkers. *Sociological Methods & Research* 43 (4): 699–723.

Lugtig, P., Das, M., and Scherpenzeel, A. (2014). Nonresponse and attrition in a probability-based online panel for the general population. In: *Online Panel Research: A Data Quality Perspective* (eds. M. Callegaro, R. Baker, J. Bethlehem, et al.), 135–153. New York: Wiley.

Lunceford, J.K. and Davidian, M. (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Statistics in Medicine* 23: 2937–2960.

Lynn, P. (2014). Longer interviews may not affect subsequent survey participation propensity. *Public Opinion Quarterly* 78 (2): 500–509.

McCarthy, J.S., Beckler, D.G., and Qualey, S.M. (2006). An analysis of the relationship between survey burden and nonresponse: if we bother them more, are they less cooperative? *Journal of Official Statistics* 22 (1): 97–112.

McGonagle, K.A. and Sastry, N. (2015). Cohort profile: the panel study of income dynamics' child development supplement and transition into adulthood study. *International Journal of Epidemiology* 44 (2): 415–422. PMCID: PMC4553706. DOI: https://doi.org/10.1093/ije/dyu076.

McGonagle, K.A., Schoeni, R.F., Sastry, N. et al. (2012). The panel study of income dynamics: overview, recent innovations, and potential for life course research. *Longitudinal and Life Course Studies* 3 (2): 268–284. PMCID: PMC3591471.

Ofstedal, M.B., Couper, M.P. (2008). Piling It On: The Effect of Increasing Burden on Participation in A Panel Study. Paper presented at the 2008 meeting of the Panel Survey Methods Workshop, University of Essex, U.K. (14 July 2008).

Phillips, C.B., Yates, R., Glasgow, N.J. et al. (2005). Improving response rates to primary and supplementary questionnaires by changing response and instruction burden: cluster randomised trial. *Australian and New Zealand Journal of Public Health* 29 (5): 457–460.

Porter, S.R., Whitcomb, M.E., and Weitzer, W.H. (2004). Multiple surveys of students and survey fatigue. *New Directions for Institutional Research* 2004 (121): 63–73.

Robins, J.M., Rotnitzky, A., and Zhao, L.P. (1995). Analysis of semiparametric regression-models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association* 90: 106–121.

Rolstad, S., Adler, J., and Rydén, A. (2011). Response burden and questionnaire length: is shorter better? A review and meta-analysis. *Value in Health* 14 (8): 1101–1108.

Sinibaldi, J. and Karlsson, A.Ö. (2016). The effect of rest period on response likelihood. *Journal of Survey Statistics and Methodology* 5 (1): 70–83.