

Creating a New Nationally-Representative Sample of Immigrants to the U.S.: Design and Results from the 2017 PSID New Immigrant Refresher

Narayan Sastry
Institute for Social Research
University of Michigan

Abstract (150 word limit)

This paper describes the design and results of the 2017 New Immigrant Refresher Sample to the Panel Study of Income Dynamics. PSID is a nationally-representative panel survey of U.S. families that began in 1968 and completed a prior new immigrant refresher in 1997. PSID launched a major new immigrant refresher in 2016, in conjunction with the Health and Retirement Study (HRS). In screened households that did not meet HRS recruitment criteria, respondents were screened for the PSID new immigrant sample. Eligible families were interviewed for the 2017 wave of PSID. We describe the results of the screening process, the baseline PSID interviews, and a multiplicity sampling procedure designed to fill a hole in the PSID new immigrant sample caused by HRS recruitment of all new immigrants born from 1960 to 1971. Finally, we assess the representativeness of the PSID New Immigrant sample using data from the American Community Survey.

Extended Abstract (2-4 pages, including tables)

Introduction

The Panel Study of Income Dynamics (PSID) is a longitudinal survey of a nationally-representative sample of U.S. families that began in 1968 (McGonagle et al., 2012). A total of 40 waves of data have been collected as part of the main PSID study, with annual data collection from 1968 to 1997 and biennial interviews from 1999 to 2017. PSID began with a sample of 5,000 families, but children who split-off to form their own households are recruited to the sample as primary respondents. This design feature, combined with high response rates, leads to growth in the PSID sample over time, and the sample currently comprises approximately 10,000 households. In addition to the main PSID interview, PSID conducts major supplemental studies such as the PSID Child Development Supplement and the PSID Transition into Adulthood Supplement.

In the 50 years since PSID was launched, immigration to the U.S. has transformed the nation's population, affecting both its size and its composition based on age, race and ethnicity, reproductive behavior, and geographic distribution (Kandel 2011). In 1968, when PSID was launched, the U.S. foreign-born population had reached its lowest point in the 20th century, comprising 4.7% of the total population (Greico et al. 2012). During the next three decades, the nation's foreign-born population grew substantially, increasing in size over three-fold by 2000 when it comprised 11.1% of the total population.

A design goal of PSID since its inception has been to represent the national population of U.S. families with sampling weights. In 1997, in response to the rapid growth of the immigrant population since the study began, PSID introduced a refresher sample of post-1968 immigrant families and their adult children (see PSID 2000). The size of this new sample was designed to be proportional to its population representation, which was estimated to be 7.5% based on data from the 1997 Current Population Survey. A total of 511 immigrant families were added in 1997,

representing 7.6% of the total PSID sample in 1997 of 6,679 families. The new immigrant sample was 52% Hispanic, 21% Asian, 12% white, 8% black, and 7% other.

Since 1997, the U.S. immigrant population has continued to expand. Estimates and projections based on microdata from the American Community Survey for 2005–2017 show that post-1997 immigrant households have increased from 2.4% of all households in 2005 to 5.2% in 2013, and are expected to have reached 6.5% by 2017. New immigrant families tend to be younger and to have higher fertility (Martinez et al. 2012; Parrado 2011), and, as a consequence, the share of children from these households has increased at an even faster rate. We estimate that 10.6% of children in the U.S. in 2017 are from post-1997 immigrant households, up from 3.5% in 2005. For younger children, the fraction is larger yet: by 2017, nearly one in seven children under age 5 years will be from a post-1997 immigrant household.

In order to support research on this important and growing population segment and for PSID's coverage of the U.S. population to remain high, PSID in 2016 launched a major project to add a refresher sample of post-1997 immigrants. The goal was to achieve nearly complete coverage by PSID of the national population of U.S. families in 2017. The new immigrant refresher sample will also help PSID to capture the increasing diversity of the U.S. population, especially by race and ethnicity and by immigration status.

In this paper, we describe the design and results of the 2017 New Immigrant Refresher Sample to the Panel Study of Income Dynamics. PSID launched a major new immigrant refresher in 2016, in conjunction with the Health and Retirement Study (HRS). In screened households that did not meet HRS recruitment criteria, respondents were screened for the PSID new immigrant sample. Eligible families were interviewed for the 2017 wave of PSID. We describe the results of the screening process, the baseline PSID interviews, and a multiplicity sampling procedure designed to fill a hole in the PSID new immigrant sample caused by HRS recruitment of all new immigrants born from 1960 to 1971. Finally, we assess the representativeness of the PSID New Immigrant sample using data from the American Community Survey.

Design of the PSID New Immigrant Sample

An independent, stand-alone screening operation to identify new immigrants for PSID was determined to be prohibitively expensive. We instead launched a collaboration with the Health and Retirement Study (HRS) to obtain the PSID new immigrant refresher sample, which led to considerable cost savings. In 2016 HRS launched a major nation-wide screening operation to identify and enroll its next two new cohorts, comprised of Late Baby Boomers, who were born from 1960 to 1965 (i.e., aged 51–56 years in 2016), and GenX1, born from 1966 to 1971 (i.e., aged 45–50 years in 2016). HRS agreed to add two questions to their screener to ascertain immigration status in households that did not screen into their sample (i.e., heads and partners who were younger than 45 years and older than 56 years), and to provide PSID with the cases that met the PSID new immigrant initial screening criteria. The two questions that were added to the HRS screener asked whether anyone in the household was foreign-born and, if so, whether the foreign-born members arrived in the past 20 years. Cases that screened positively to these two questions were transferred to PSID for a second-stage screening in which a short (10 minute) questionnaire was completed to establish more precise eligibility and collect detailed contact information.

HRS kept for its sample all households with a head aged 45–56 years, some of whom are post-1997 immigrants and hence were eligible to be selected for the PSID new immigrant refresher sample. To fill this missing segment of the PSID new immigrant sample, we designed and

implemented a “multiplicity sampling” approach (Sirken 1970). This approach allowed us to leverage the sample we do obtain from HRS to identify cases in the missing age-segment, and has been used previously in similar applications (Wagner & Lee 2014; Rothbart et al. 1982; Kalton & Anderson 1986). Because it is a probability sampling technique, we can construct weights that will enable proper inferences to the target population.

The multiplicity sample method uses the 2017 PSID New Immigrant sample as an initial probability sample of families and households. Respondents in this sample were asked to identify post-1997 immigrant family units within their family network that have a head or spouse in the target age range of 45–56 years. To ensure accuracy in reporting the status of family network members, the network was limited to parents, children, and siblings residing in the U.S. Respondents were asked to provide a complete roster of these individuals, which allows us to compute inclusion probabilities for each family identified for the new immigrant sample through this multiplicity process. Note that multiplicity sampling is distinct from “snowball sampling” and “peer nomination” methods that often recruit respondents from undefined or undefinable networks and hence provide no basis for determining inclusion probabilities. Multiplicity sampling provides a cost-effective way to identify a rare population segment, and the intergenerational and intragenerational ties among family members it provides is a distinct strength given PSID genealogical sample design.

Fieldwork Results

HRS began screening operations in late February 2016, and released a total of 47,591 cases to the field. In February 2017, HRS stopped PSID immigrant screening (but continued screening for HRS). This decision reflected long-standing concern about the effects of asking immigrant questions on respondent cooperation at both the household and neighborhood levels. HRS had previously conducted an experiment to assess whether there were any negative effects of asking the PSID immigrant screener questions, but did not find any evidence of such effects. However, the 2016 presidential election and the national climate concerning immigration policy and immigrants raised significant concerns. Although there were no specific negative events in the field, there was a general sentiment that seeking to interview immigrants was challenging and potentially problematic, particularly in the context of HRS already experiencing a separate set of fieldwork challenges associated with declining rates of respondent cooperation.

The pool of HRS released cases eligible for PSID screening questions was estimated at approximately 35,000 households. The HRS screener response rate was 54 percent and a total of 1,378 eligible new immigrant households were identified, from which a total of 1,490 new immigrant families were released to the field for PSID screening (there were a higher number of families due to split-offs and the presences of multiple eligible family units in some dwellings). In the second-stage PSID screening, 63 percent of families completed a screener interview and a total of 614 families were determined to be eligible for the PSID new immigrant refresher sample. As part of the 2017 wave of PSID, 455 interviews were completed among the screened new immigrant households, representing a 76 percent response rate. These 455 PSID families in turn identified a total of approximately 50 relatives who were eligible for the PSID multiplicity sample. Screening was recently completed with these multiplicity sample members and PSID interviews with these families are planned as part of the 2019 wave of PSID. In 2019, the 455 new immigrant families will be eligible for panel reinterviews as well as all other study components—such as the 2019 PSID Transition into Adulthood Supplement and the 2019 PSID Child Development Supplement. In addition, we will attempt to complete interviews with approximately 144 eligible new immigrant families that did not complete a PSID interview in 2017 (a small number of the 614 released new immigrant households were determined to be

ineligible). Overall, we expect by 2019 to have added a total of approximately 500 new immigrant families to PSID, representing (on an unweighted basis) approximately 5 percent of the total PSID sample.

Plan for Completing Paper for PAA 2019

We have a full set of fieldwork results from the 2016–2017 HRS screening effort, as well as the secondary PSID screening effort conducted during this same period. We also have final results from the 2017 PSID interviews with new immigrant families identified through the screening effort. Finally, we have complete results from the process of identifying the multiplicity sample members, which was conducted as part of the PSID interviews, and screening these cases.

We are currently in the process of finalizing all response rates and constructing the sampling weights, which will provide important insights into the fieldwork outcomes for the 2017 PSID new immigrant refresher effort.

Lastly, we will undertake an analysis as part of the post-stratification weighting process to assess the coverage and representation of the PSID new immigrant sample compared to a benchmark nationally-representative sample of new immigrant individuals and families identified in the 2017 American Community Survey.

We will report final results on all of these study components as part of the paper we will prepare for the 2019 PAA Annual Meeting.

References

- Grieco, E.M., Trevelyan, E., Larsen, L., Acosta, Y.D., Gambino, C., De La Cruz, P., Gryn, T., & Walters N. (2012). The Size, Place of Birth, and Geographic Distribution of the Foreign-Born Population in the United States: 1960 to 2010. Population Division Working Paper No. 96, U.S. Census Bureau, Washington, D.C.
- Kalton, G., & Anderson, D.W. (1986). Sampling Rare Populations. *Journal of the Royal Statistical Society. Series A*, 65-82.
- Kandel, W.A. (2011). U.S. Foreign-Born Population: Trends and Selected Characteristics. DIANE Publishing.
- Martinez, G., Daniels, K., & Chandra, A. (2012). Fertility of Men and Women Aged 15-44 years in the United States: National Survey of Family Growth, 2006-2010. *National health statistics reports*, (51), 1-28.
- McGonagle, K.A., Schoeni, R.F., Sastry, N., & Freedman, V.A. (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.
- Parrado, E.A. (2011). How high is Hispanic/Mexican fertility in the United States? Immigration and Tempo Considerations. *Demography*, 48(3), 1059-1080.
- PSID. (2000). Information on the PSID Immigrant Sample Addition of 1997/1999, Ann Arbor, MI: Panel Study of Income Dynamics Technical Paper 00-04. Retrieved from: http://psidonline.isr.umich.edu/Publications/Papers/tsp/2000-04_Imm_Sample_Addition.pdf
- Rothbart, G.S., Fine, M., & Sudman, S. (1982). On Finding and Interviewing the Needles in the Haystack: The Use of Multiplicity Sampling. *Public Opinion Quarterly*, 46(3), 408-421.
- Sirken, M.G. (1970). Household surveys with multiplicity. *Journal of the American statistical Association*, 65(329), 257-266.
- Wagner, J., & Lee, S. (2014). Sampling Rare Populations. *Health Survey Methods*, 77-104.