HHS Public Access

Author manuscript

Ann Am Acad Pol Soc Sci. Author manuscript; available in PMC 2017 April 24.

Published in final edited form as:

Ann Am Acad Pol Soc Sci. 2015 January ; 657(1): 208–246. doi:10.1177/0002716214552773.

Potential Data Sources for a New Study of Social Mobility in the United States[1]

John Robert Warren

Minnesota Population Center, University of Minnesota

In this article I evaluate potential sources of data for a new study of social and economic mobility in the United States. After reviewing the major costs and benefits to be considered in such an evaluation, I discuss the strengths and limitations of four sets of options: (1) ongoing smaller-*n* surveys that might be supplemented to include a social mobility component; (2) ongoing larger-*n* surveys—the Current Population Survey and the American Community Survey—that might serve as a basis for a new mobility study; (3) linked administrative records; and (4) a new stand-alone survey. I conclude by summarizing the most viable and advantageous options.

Parameters: Defining Costs and Benefits

In this section I define the criteria I use to evaluate the costs and benefits of the several sources of data that might be used in a new study of social mobility in the United States. For each criterion, I describe the ideal-typical situation—the quality of a data resource that would be observed in a perfect world. Later I compare real data resources to these standards.

Population definition and coverage

The ideal data resource would allow analysts to generalize their conclusions about social and economic mobility to a well-defined population. On a basic level, observations would (at least after weighting) reflect the characteristics of that entire population and of sub-groups that have traditionally been difficult to interview or enumerate (Groves and Couper 1998; Olson and Witt 2011). This may include people in institutions (e.g., prisons, dormitories, homeless shelters); people who are exceedingly busy or mobile (e.g., single mothers, over-the-road truck drivers); people who may have fears about the consequences of cooperation (e.g., unauthorized immigrants, elderly people); and people for whom responding is simply more difficult (e.g., non-native English speakers, people with severe disabilities).

Beyond these basic coverage issues, however, the specification of a “well-defined population” depends on the way in which social and economic mobility will be studied. If the fundamental question is, “How do social origins affect people's adult outcomes?” then it is sufficient to begin with a cross-section of American adults at one point in time (which I

[1]This report was prepared for the National Research Council's Committee on Population's June 10, 2013 “Meeting on Developing a New National Survey on Social Mobility.” Elements of this report were informed by conversations with J. Trent Alexander and Todd Gardner at the U.S. Census Bureau; Steve Ruggles, Gina Allen, and Carolyn Liebler at the Minnesota Population Center; Andrew Halpern-Manners at Indiana University; and Eric Grodsky at the University of Wisconsin-Madison. However, errors and omissions are my responsibility. For questions, please contact me at 612.624.2310 or warre046@umn.edu.

will call Generation 0, or G0); to observe the attributes of their parents (who are in Generation -1, or G-1); and to model their outcomes as a function of their parents' attributes. Under this specification, people included in G-1 are *not* a representative cross-section of the population of Americans. Some members of G-1 are *over*-represented under this specification because they had more than one child who survived to join G0; others are *under*-represented because they produced no such children (e.g., because they left the U.S. before having children, because they produced no offspring, or because their children did not survive). That is, the "well-defined population" is a population of offspring, not of parents.

Conversely, if the basic question amounts to, "How do people transmit their social and economic advantages and disadvantages to their offspring?" then it is sufficient to begin with a cross-section of American adults at one point in time (G0); to observe the attributes of their adult children (who are in Generation +1, or G+1); and to model their children's outcomes as a function of their own characteristics. Here, people included in G+1 are not a representative cross-section of all Americans. In particular, people who entered G+1 via immigration are not represented under this design. The "well-defined population" here is thus a population of parents, not of offspring.

If the core question about social and economic mobility is posed a different way—"How does the distribution across social and economic positions in one generation evolve into the distribution of those positions in the next generation?"—then the "well-defined population" might pertain to both parents *and* offspring. Starting with data that represent a cross-section of Americans at one point in time *and* data that represent a cross-section of Americans at a subsequent point in time (where the two times are separated by a few decades), it would be possible to study both intergenerational mobility *and* the several demographic processes that reshape distributions of advantage and disadvantage over time and that allocate people to places in those distributions (Mare 2011; Mare and Song 2011). This specification would permit both of the analyses described above. However, it would also facilitate analyses of both (1) intergenerational social and economic mobility and (2) the roles of differential fertility, mortality, and migration in transforming distributions of outcomes across generations.

Each of the specifications above could be modified to pertain to populations of *families* as opposed to population of *individuals*. For example, one could begin with a representative cross-section of pairs of siblings in G0 and then observe either those siblings' parents (in G-1) or their children (in G+1). Likewise, one could begin with a representative cross-section of people in G0 and then observe the attributes of people's parents (G-1) and grandparents (G-2) or else of their children (G+1) and grandchildren (G+2); these could be expanded to include more complete ancestry or descendent trees. For the reasons described above, however, each of these designs would only produce representative samples of people or sibling pairs in G0.

Sample size

The ideal data resource for a new study of social and economic mobility would facilitate separate analyses of important social and demographic subgroups. At minimum this would require a sufficiently large sample (in G0) to consider groups defined by the cross-

classification of gender, race/ethnicity, and nativity (e.g., U.S.-born white men, foreign-born Hispanic women). Ideally this would allow for separate analyses by geography (e.g., by state, by urban/rural/suburban status). Sufficiently large samples of smaller population subgroups could be obtained by strategic oversampling or by use of large-scale administrative data that include measures of these social, demographic, and geographic attributes.

Topical coverage

A new study of social and economic mobility would require measures of education, occupation, and income for parents and children. These measures form the cornerstone of research on intergenerational mobility. Most prior research has measured these for children and their *fathers*. In light of evidence of the importance of *mothers* (e.g., Beller 2009; Kalmijn 1994), a new study would require measures of children's and both parents' educations, occupations, and incomes.

Beyond these core measures, the ideal data resource might also include a number of other measures of children's and both parents' social and economic conditions. These might include (in no particular order): cognitive and non-cognitive skills; job quality and working conditions; poverty status; wealth and debt; subjective social class; voting behaviors; family structure; incarceration experiences; health and disability; and neighborhood characteristics. With respect to each measure, there are at least nascent research literatures on intergenerational mobility; most of these literatures express the need for larger and richer data resources.

Aside from measures of social and economic circumstances, a new study of mobility would also need to include measures of children's and parents' race/ethnicity, gender, nativity, and geographic location. Depending on the design of the study, it may also be necessary to obtain all of these social, economic, demographic, and geographic measures for siblings, spouses, and/or ancestors or descendants who are more distant than one generation.

The above measures would be sufficient to describe patterns of intergenerational mobility in social and economic circumstances, separately for population sub-groups. However, additional measures would be required to fully consider the roles of various factors that mediate those patterns—that is, to describe the mechanisms through which mobility and immobility occur. For example, the ideal data resource might include measures of childhood health, birth weight, and attributes of intrauterine conditions; school characteristics, opportunity to learn, and college selectivity; significant others' influence and social and cultural capital; genetic markers for things like depression or mental illness; and personality and taste.

Temporal issues

A key motivation for the present effort is that it has been four decades since the last major, dedicated study of social and economic mobility in the United States. Implicit in this motivation is the notion that the U.S. has changed in important ways—demographically, economically, politically, etc.—such that we need to update our assessment of mobility

patterns and processes. As a result, any new study of social and economic mobility needs to produce results that pertain to the contemporary U.S. population, society, and economy.

However, the ideal data resource would also allow for valid comparisons over time periods *and* across birth cohorts. Sociologists and others have learned a great deal, for example, by comparing the 1962 and 1973 Occupational Changes in a Generation samples (OCG; e.g., Featherman and Hauser 1976a; Featherman and Hauser 1976b; Hauser and Featherman 1976) and by modeling time trends using multiple years of the General Social Survey (GSS; e.g., Hout 1988; Wolfinger 1999). Both have facilitated analyses of change over periods and cohorts. A new data resource for studying social and economic mobility should allow researchers to update these temporal comparisons while also describing social mobility among contemporary Americans (and separately by birth cohorts).

Spatial issues

The ideal data resource for a new study of social and economic mobility would permit two sorts of spatial comparisons. First, it would allow for comparisons of the U.S. to other countries. Cross-national comparative research design are very useful for understanding the roles that macro-level political, economic, social, and demographic factors play in shaping mobility patterns (e.g., Breen and Jonsson 2005; Smeeding, Jäntti and Erikson 2011).

Second, it would allow for careful consideration of the role of spatial factors *within* the U.S. in shaping mobility patterns and processes. Previous work has considered the ways in rural or farm residence, life in segregated neighborhoods in central cities, and residence in particular geographic regions of the country have shaped opportunity structures within and across generations. The design of a new study of mobility should permit such geographic comparisons.

Sustainability

Outside of the GSS, the United States has no data resource for routinely monitoring patterns of intergeneration social and economic mobility. Although a new study would provide a needed update to our understanding of patterns and processes of intergenerational mobility, it too will soon become dated if it only involves collecting or assembling data at one point in time. The ideal data resource for a new study of social and economic mobility would thus represent just the first iteration of an ongoing effort to monitor this important aspect of American society.

Financial expense

A new data resource with broad population coverage; with a sample size suitable for sub-group analysis; and with broad topical coverage will be expensive—especially if it is to become a sustainable and replicable effort going forward.

Perhaps the most financially costly way to accomplish these goals would be to conduct a new stand-alone data collection effort. From a *financial* point of view, it would probably be more efficient to instead supplement ongoing data collection efforts (e.g., GSS) or to utilize administrative record data (e.g., Internal Revenue Service records). However, these monetary

efficiencies would likely be offset by reduced control over the specification of the study population; the design of the sample; the execution of the fieldwork or other data collection; and the focus and breadth of the measures that can be included. In this respect, there is no “ideal” new data resource. Every option will involve some trade-off between the ideals described above and the need to keep the costs of the new data resource reasonable (and sustainable).

Privacy and data access

The mainstream social science model has been to collect survey data; to produce a public-use version of the data that prevents direct or deductive disclosure of identifying or otherwise confidential information; and to disseminate the data and associated metadata. This model becomes more problematic when data resources include (or entirely consist of) administrative record data that either cannot easily be de-identified or that cannot, by law, be as freely disseminated as most survey data. Major U.S. longitudinal surveys—such as High School & Beyond, the National Longitudinal Study of Adolescent Health, the Health and Retirement Survey, and the National Longitudinal Study of Youth—have been linked to administrative record data (most commonly school transcripts, Social Security Administration records, and mortality records). Typically, researchers must apply for access to these augmented data, and frequently they must access the data in a secure data enclave or using some other restrictive protocols that enhance data security.

The security and confidentiality of study participants' data is thus a major consideration in designing a new study of social and economic mobility. To the extent that the new data resource contains administrative records or includes private information—about participants' identities, to be sure, but also about sensitive subjects such as income, assets, or health—researchers will have to accept a trade-off between data access and data security. The ideal data resource would certainly include measures of sensitive issues (e.g., income, health) and may well include administrative record data. Although such a data resource might be “ideal” with respect to the considerations described above, researchers may have to access the data via secure data enclaves or via facilities like the U.S. Census Bureau's Research Data Centers (RDCs). Such restrictions on data access may be necessary to maintain the confidentiality of data; however, their likely down side would be that fewer people would be willing or able to access the data.

Here again, in this respect there is no “ideal” new data resource. Every option will involve some trade-off between data security and data access. A new and highly-secure data resource that can only be analyzed in an RDC will probably be used by substantially fewer analysts; not everyone has easy access to RDC's. In contrast, a new data resource with no links to administrative data and without questions on sensitive topics may be easy to disseminate and easy for researchers to access. However, such a survey would be relatively less rich with respect to the scope and depth of data and measures.

Evaluating Existing Smaller-Scale Surveys

In the appendix I describe the design and content of nine (relatively) smaller-*n* sample surveys of the U.S. population that might be expanded or otherwise supplemented to serve

as a data resource for a new study of intergenerational social and economic mobility. These nine include GSS, HRS, HSB, Add Health, NLSY-79, NLSY-97, PSID, PT, and SIPP (see the appendix for an explanation of the acronyms). In this discussion I do not include smaller-*n* panel surveys for which there are no plans for future re-interviews (e.g., the older National Longitudinal Study cohorts, the National Education Longitudinal Study of 1988) or smaller-*n* panel surveys of adolescents who will not soon reach full adulthood (e.g., the National Children's Study, the Education Longitudinal Study of 2002, the High School Longitudinal Study of 2009). In this section, I summarize the costs and benefits of using these nine (relatively) smaller-*n* sample surveys in light of the parameters outlined above. I then conclude with a "bottom line" statement about the prospects for basing a new study on one or more of these nine resources.

Population definition and coverage

Several of the (relatively) smaller-*n* sample surveys entirely exclude institutionalized individuals. GSS, SIPP, and HRS are samples of household-based adults, and thus excluded *adults* who were institutionalized. PT and HSB began as samples of in-school adolescents, and thus excluded *children* who were institutionalized.

Each of the nine smaller-*n* surveys allows researchers to generalize their results to the population of people included in their main samples; following the convention above, call these people Generation 0 (or G0). Six of the nine surveys—NLSY-79, NLSY-97, Add Health, HRS, PT, and HSB—only include people in particular birth cohorts. Eight of these nine surveys—all but NLSY-97—are designed such (1) the social and economic circumstances of people in G0 are observed directly and (2) members of G0 are asked to report the social and economic circumstances of their parents (who constitute G-1). Three of them—NLSY-97, Add Health, and HSB—also include supplemental surveys with people who are themselves in G-1 by virtue of being a parent of someone in G0. At least in some years, three of the surveys—GSS, PSID, and NLSY-79—also include information about the social and economic circumstances of the *children* of people in G0 (who constitute G+1). As described in the appendix, NLSY-97 also includes information about people in G-2 (grandparents of people in G0) and PSID includes information about people in G+2 (the grandchildren of G0). Thus four surveys—GSS, PSID, NLSY-79, and NLSY-97—include at least some information about *three* generations of people. Finally, seven of the nine surveys —all but SIPP and HRS—observe siblings' social and economic characteristics in G0 or G +1.

For reasons described above, however, these nine surveys are only representative of a cross-section of Americans in G0. People in G-1 (or G-2) are not observed at all if they had no surviving children, and are over-represented to the extent that they had multiple surviving children. Because GSS (in 1994), PSID, and NLSY-79 include information about *all* of the children of G0, it is tempting to conclude that these samples are representative cross-sections of people in both G0 *and* G+1. However, people in G+1 who entered the U.S. via immigration (instead of being born to someone in G0) cannot be included in G+1 in those surveys (unless they are part of a subsequent supplementary oversample of immigrants, as in the PSID). As described above, the fact that these nine data resources are only representative

of G0 broadly constrains the sorts of questions that can be asked using these data. It would difficult, for example, to use any of them to fully consider the roles of fertility and migration in transforming across generations the distribution of social and economic resources in the United States.

Sample size

Each of these nine sample surveys includes enough observations of people in G0 to permit analyses of social and economic mobility between G-1 and G0 and/or between G0 and G+1. However, some—notably the GSS—are not sufficiently large to permit detailed sub-group analyses; a few—NLSY-79, NLSY-97, and Add Health—only allow for sub-group analyses by virtue of having over-sampled some population sub-groups. As described below, however, the capacity to conduct sub-group analyses using these nine data resources is sometimes constrained by the availability of information about detailed geography.

Topical coverage

Each of these nine smaller-*n* surveys includes observations of the educations, occupations, and incomes of people in G0. However, only one—PSID—observes all three for two generations (G0 and G+1). In most cases, income is only observed for G0 and not for adjacent generations; NLSY-97 and PT also include more subjective "relative family income" measures (i.e., average, above average, etc.) for G-1.

To varying degrees, these nine surveys include information about other social and economic circumstances for people in G0 such as job quality, poverty status, wealth, subjective social class, voting behaviors, incarceration, cognitive and non-cognitive skills, health, disability status, and neighborhood conditions. Only in exceptional cases—mainly in the PSID—are any of these things observed for either G-1 or G+1, thereby generally precluding analyses of social mobility in these other dimensions of social and economic conditions. A few exceptions: military service is observed for people in both G0 and G+1 in NLSY-79, and summary measures of health are available for people in G0 and G-1 in NLSY-97 and Add Health.

Temporal issues

If a primary motivation for a new social mobility study is to understand the mobility experiences of contemporary Americans, then these nine studies vary greatly with respect to their utility. Figure 1 depicts temporal dimensions of these nine surveys and also includes the older NLS cohorts and the Occupational Changes in a Generation (OCG) surveys for comparison purposes. The horizontal axis shows the range of years in which people in G0 in each survey were born. The vertical axis shows the year in which members of G0 were first interviewed (which would be the *only* year people were interviewed for cross-sectional surveys). Whereas the OCG samples were first contacted long ago (and are thus low on the vertical axis) and included few people born after the middle of the 20th century (and are thus low on the horizontal axis), a new social mobility study would ideally occur in the near future and would include people from a wide range of birth cohorts including recent ones (thus putting it high on both axes).

By this standard, the most promising of these nine surveys is the GSS and the new panel—presuming the latter includes items akin to those included in the 1986-1988 SIPP topical modules. Both will include people from a wide range of birth cohorts including very recent ones, and both will describe representative cross-sections of contemporary Americans in G0. HRS respondents and the PSID children are also first observed recently, but are from a constrained range of birth cohorts. People in G0 in the PSID in 1968, in the NLS surveys, or in Add Health, PT, HSB, or OCG were first observed long ago and/or are also from a truncated range of birth cohorts. Of course, surveys that are low on the vertical axis in Figure 1 are useful for making inferences about trends over time and/or across birth cohorts. However, they do not reflect the experiences of contemporary Americans from a range of cohorts.

Spatial issues

These nine surveys collect detailed geographic information, but generally make that information available to researchers only through restricted data use agreements. In most cases, only state or region and information about rural/suburban/urban residence is available publicly. To varying degrees, each of the nine surveys is designed facilitate cross-national comparisons. GSS is part of the International Social Survey Program, which makes some effort to harmonize measures across surveys in various countries. HRS is designed in coordination with parallel surveys in the United Kingdom and several other countries, thereby making cross-national comparison easier.

Sustainability

If a goal is to design a data system that would regularly provide updated information about social and economic mobility in the United States, then of these nine surveys only two—the GSS and SIPP—are suitable targets of opportunity (again presuming the new SIPP surveys might include necessary survey questions). It is difficult to imagine routinely fielding new cohort-based studies like NLSY-79, NLSY-97, PSID, Add Health, PT, or HSB. It seems likely that HRS will continue to include new birth cohorts over time, but by design the HRS will always include people born in earlier birth cohorts (and who thus entered the labor force, for example, some time ago). GSS and SIPP, on the other hand, are likely to continue to include new cross-sections of Americans born across multiple birth cohorts and could—perhaps with some urging—continue to include relevant measures.

Financial expense

It is difficult to gauge the precise financial cost of basing a new social mobility study on supplements to or extensions of these nine surveys. It seems safe to assume, however, that including necessary topical modules on existing, ongoing surveys—such as SIPP and GSS—would be far less expensive than fielding large new data collection efforts.

Privacy and data access

Data from each of these nine surveys is generally freely available to researchers. As noted above, however, in some cases content like detailed geography is only available via restricted use data agreements. Because the producers of these data resources withhold

identifying information and take steps to prevent deductive disclosure, the use of these surveys presents relatively little risk to subjects' privacy.

Bottom line

If a new study of social and economic mobility is to be based on (or represent an extension of) one of these nine surveys, then GSS and SIPP are the best candidates. As GSS is now configured, however, detailed sub-group analyses are difficult and the scope of information available about G-1 is limited to education and occupation; information about G+1 or about the siblings or other family members of G0 has not been collected since 1994. A new GSS topical module might include a more detailed income measure for people in G0 and G-1 and measures of education, occupation, and income for all of the children born to people in G0 (aka, G+1). Given its larger sample size, SIPP might be a more attractive option—but only if the new panel (and future panels) includes detailed information about the social and economic circumstances of people in G-1 and/or G+1. Another possibility is to include new modules of questions in the HRS that more fully capture the social and economic conditions of people in G+1—the children of HRS panel members. Unfortunately, none of these three surveys includes institutionalized individuals in G0; this is a potentially serious omission, especially for some population subgroups (Wakefield and Uggen 2010).

Beyond their exclusion of institutionalized people from G0, basing a new mobility study on GSS or SIPP would also be limiting in two other important respects. First, even if rich topical modules of questions on the social and economic circumstances of people in G-1 or G+1 can be added to the GSS or SIPP, our information about G-1 or G+1 would be limited to just a few measures (e.g., of education and occupation). Second, because the GSS and SIPP samples are representative cross-sections of people in G0 (and not of people in G-1 or G+1), the nature of the research questions we can ask is constrained. With information about G-1 and G0, it would be possible to ask, “How do people in G0's social and economic outcomes depend on their families of origin?” With information about G0 and G+1, it would be possible to ask, “How do people in G0 pass along social and economic advantage and disadvantage to their children in G+1?” But it would not be possible to *fully* consider the roles of demographic factors—differential fertility, child mortality gradients, international migration, and assortative mating—in reproducing and transforming distributions of social and economic resources over time.

Evaluating Existing Larger-Scale Surveys I: The Current Population Survey

The Current Population Survey (CPS)—the design and content of which is summarized in the appendix—has twice before served as a vehicle for major national studies of social and economic mobility (Blau and Duncan 1967; Featherman and Hauser 1978). In this section, and in light of the parameters outlined at the outset, I describe the costs and benefits associated with returning to the CPS for a new social mobility study. Note that I say nothing here about the political or practical obstacles that one might encounter in trying to implement such a study using the CPS. In fact, the political, administrative, and perhaps legal hurdles may be high.

There are two basic designs by which the CPS might be leveraged for this purpose. The first design, which I will call “OCG III,” involves fielding a follow-on survey *after* the CPS, administering to CPS respondents an additional battery of questions about their own and their family members' social and economic circumstances. The second design, which I will call the “February Mobility Supplement,” involves fielding a new topical supplement (in February) *as part of* the CPS's rotation of topical supplements. Both designs would involve asking additional questions of CPS participants who were eligible to respond to the March Annual Social and Economic Supplement (ASEC); this is because the ASEC collects detailed income and other economic information about each household member that would be costly to collect anew.

OCG III

The 1962 and 1973 Occupational Changes in a Generation (OCG) surveys involved leave-behind and mail back surveys administered to adult men in the March 1962 and March 1973 CPS's, respectively (Blau and Duncan 1967; Featherman and Hauser 1978). OCG III would involve follow-up surveys with people who (1) have recently completed their eighth month in the CPS 4-8-4 rotation and (2) were also interviewed in the month of March. By design, this strategy would mean focusing on the ~65,000 individuals each year who complete their eighth month in the CPS rotation in the months of March, April, May, or June (Drew, Flood and Warren Forthcoming). After weighting to account for sampling errors and non-response, these ~65,000 individuals would represent a cross-section of non-institutionalized Americans.

These ~65,000 people could then be surveyed—by phone, web, mail, or a mixture of modes—to ascertain the social and economic circumstances of their family members, including parents, children, spouses, in-laws, and siblings. Many of the measures of key concepts—education, occupation, income, veteran status, disability, and so forth—would need to parallel CPS measures in order to allow for comparisons across and within generations.

Although the design of OCG III may be inspired by the original OCG surveys, practically speaking the modern analogue is the American Time Use Survey (ATUS; U.S. Census Bureau 2013). ATUS is administered by the Census Bureau and collects information about how Americans use their time. ATUS draws it sample—generally about 13,000 people annually in recent years—from the larger CPS sample. Two months after completing their eighth month in the CPS rotation, people are eligible to be included in the ATUS.

February Mobility Supplement

Because of the CPS's 4-8-4 design, three quarters of respondents to the March ASEC are also interviewed February. (Although three quarters of March ASEC respondents are also interviewed in *April*, the CPS already typically features an April supplement on child support, at least in even-numbered years.) Thus a second CPS-based design would be to field a February Mobility Supplement. As part of the main CPS data collection operation, the survey instrument would ascertain the social and economic circumstances of CPS respondents' family members. Another virtue of a February supplement is that three quarters of the respondents would also be eligible to participate in the CPS in January, which

frequently (at least in even-numbered years) features supplemental modules on job tenure, occupational mobility within careers, and displaced workers. Indeed half of the February respondents would, by design, be eligible for *both* the January and March supplements.

Most CPS topical supplements are funded by federal agencies with policy interest in their content. For example, the October School Enrollment supplement is partially funded by the National Center for Education Statistics and the periodic Tobacco Use supplements are sponsored by the National Cancer Institute. One possibility would be to partner with the Census Bureau to field the supplement and to develop a series of reports on this important aspect of American social and economic life.

Population definition and coverage

The OCG III and February Mobility Supplement designs would each permit generalizations to G0—the cross-sections of Americans included in the CPS sample. Although it would be possible to study social and economic mobility from G-1 to G0 and/or from G0 to G+1, the resulting data would not be representative of the populations of people in either G-1 or G+1. What is more, because of the sampling design of the CPS, people living in institutions—most importantly, prisons—would not be included in G0.

Sample size

As described above, OCG III—which would involve interviews with people who have completed their eighth month in the CPS rotation and who were interviewed in March—would include about 65,000 individuals. The February Mobility Supplement design would include about 140,000 individuals, although three in four (or about 105,000) would also be eligible for the March ASEC. These sample sizes would generally permit analyses of social and economic mobility among important population sub-groups.

Topical coverage

The CPS basic monthly survey collects information about educational attainment, labor force status, occupation, military service, veteran status, and demographic characteristics. The March ASEC further includes information about wages and income, as well as (in recent years) a basic measure of self-reported health. Either OCG III or the February Mobility Supplement could expand the scope of information gathered about the social and economic circumstances of members of G0 and could collect parallel information about their family members. As described above, the February Supplement would be folded into CPS fieldwork, and so the length of that supplement may be more constrained than a separate standalone survey like OCG III.

Temporal issues

Either OCG III or a February Mobility Supplement would include very recent cross-sections of Americans from multiple birth cohorts. As such, they would reflect the mobility experiences of contemporary Americans. Furthermore, either would produce results that could be comparable to those produced by the earlier OCG surveys (as long as the new measures were collected in a way that permitted valid temporal comparisons).

Spatial issues

Publicly-available CPS data include information about respondents' state, metropolitan area, and urban vs. rural vs. suburban residence. Restricted-use versions of the CPS—available in RDCs—include more detailed geographic information. Thus either the OCG III or February Mobility Supplement designs could facilitate analyses that consider the role of “place” in the process of social and economic mobility.

Sustainability

The CPS is an important and ongoing federal survey, and has implemented the 4-8-4 rotation design for more than half a century. Although either the OCG III or February Mobility Supplement designs could be implemented once, it is also possible to imagine repeating them multiple times going forward. This would provide a continually updated portrait of the mobility circumstances of the contemporary American population, and would also facilitate analyses of change over time in mobility rates and processes.

Financial expense

Because these two CPS-based designs would build on the Bureau of Labor Statistics' ongoing sample design, respondent tracking, fieldwork, and data processing and management operations, it seems likely that either the OCG III or February Mobility Supplement designs would be far less costly than fielding an entirely new stand-alone survey, especially if that survey included anything like 65,000 to 140,000 respondents.

Privacy and data access

The CPS data are publicly available from the Bureau of Labor Statistics, either directly (via Data Ferret) or from organizations like the National Bureau of Economic Research and the Minnesota Population Center. The data are released after recoding key variables in such a way as to minimize the possibility of deductive disclosure. It is thus conceivable that OCG III or February Mobility Supplement data could also be subject to procedures designed to prevent deductive disclosure and be made available to the public without undue risk to subjects.

Bottom line

Supplementing the CPS—using either of the two designs described here—would produce new data on contemporary patterns of social and economic mobility that would include large samples and high quality measures of the social and economic circumstances of sample members and their family members. On a per-respondent basis, this option would likely be less expensive than a new stand-alone survey. Finally, because of its affiliation with a federal government agency, either CPS option is likely to generate higher response rates than could be obtained in stand-alone surveys conducted by universities or private organizations.

The main limitation of any CPS-based approach—beyond the political, administrative, and legal hurdles one might face in implementing it—is that institutionalized individuals are excluded by design from the sample. Given high and changing rates of incarceration for subgroups of the U.S. population, this may be an important limitation. The February

Mobility Supplement might also be limited with respect to the range of concepts that could be measured for G-1, G+1, or both; in general, CPS topical supplements are not exceptionally long or complicated. However, for reasons described above, the February Mobility Supplement design would yield a considerably larger sample of people in G0 than the OCG III option. Finally, as with the nine smaller-*n* sample surveys described above, it would not be possible to use CPS-based data to fully consider the roles of demographic factors—fertility, child mortality, international migration, and assortative mating—in transforming and reproducing distributions of social and economic resources across generations. Again, this is because the CPS sample is representative only of people in G0, not of people in their parents' or children's generation.

Evaluating Existing Larger-Scale Surveys II: The American Community Survey

The basic design and content of the American Community Survey (ACS) is described in the appendix. There are two basic ways in which ACS data might be leveraged for the purpose of a new social mobility study. First, ACS records could be linked to existing survey and administrative data to provide information about multiple generations of people and family members. This possibility is described in the *next* section. Second, it is possible to conduct a follow-up survey of ACS respondents. I explore the latter possibility in *this* section. Again, I do not dwell on the possible political, administrative, and legal hurdles that might be encountered in trying to implement either design.

A follow-up survey of ACS respondents would likely be superior to a completely independent, stand-alone social mobility survey in three important respects. First, it would take advantage of the ACS's efforts to draw a representative sample of a cross-section of Americans, including those in institutions and other group quarters. Second, the ACS data already include detailed measures of the education, income, occupation, labor force status, veteran status, family circumstances, health insurance coverage, and disability status of everyone in the cross-sectional sample of people in G0. This means that an ACS-based social mobility study would need only to relocate and interview people who previously participated in the ACS; collect additional information about them; and collect parallel information about their family members. Even a subset of ACS sample members—especially if strategically selected based on attributes observed in the main ACS (e.g., race/ethnicity, educational attainment, country of origin, institutional residence)—would likely yield a sufficiently large sample to sustain analyses of detailed population subgroups. At the same time, the expense per respondent would likely be less than for a stand-alone social mobility study. Third, by virtue of being a federal survey, an ACS follow-up is likely to enjoy higher response rates than stand-alone surveys conducted by universities or private organizations.

Such a design would not be unprecedented. A good model might be the National Survey of College Graduates (NSCG; National Research Council 2008: Ch. 3). Since 1962 the NSCG has re-interviewed a sample of highly educated people as identified in the most recent decennial census or ACS. Most recently, members of the 1993, 2003, and 2010 NSCG

longitudinal panels were selected from among those who received the long forms of the 1990 or 2000 decennial enumerations or the post-2000 ACS. The NSCG then interviews people (about their work, income, and other social and economic outcomes) and links responses to census or ACS records.

Population definition and coverage

Unlike the GSS, SIPP, or CPS options discussed above, the ACS follow-up design would generate a representative cross-section of Americans (in G0)—including those in institutions. The ACS is certainly not perfect with respect to population coverage or response rates, and it has historically faced challenges enumerating people in group quarters (Beaghen and Stern 2009). Nonetheless, its coverage of the institutionalized population is a major strength. As with most of the data resources discussed to this point, this ACS-based design would yield information about people in G0 and about their parents (G-1) and/or children (G+1) and potentially other family members. The resulting data would describe the mobility experiences of a single cross-section of Americans (G0) from multiple birth cohorts.

Sample size

In recent years the ACS has sampled more than two million households and collected information about their occupants. Subject to financial constraints, an ACS follow-on study could thus strategically sample ACS respondents in such a way as to maintain adequately high numbers of sample members within strategically selected population subgroups.

Topical coverage

As noted above, the ACS already collects information about each household member's demographic characteristics, education, income, occupation, labor force status, veteran status, family circumstances, health insurance coverage, and disability status. The follow-on survey could expand the range of social and economic variables observed for members of G0, and could add parallel information for people in G-1, G+1, or other family members.

Temporal issues

An ACS follow-on survey would provide information about a representative cross-section of contemporary Americans. It would be possible to restrict the sample to non-institutionalized people or to people in particular age ranges to compared results to those generated using GSS or OCG data.

Spatial issues

A key strength of the ACS is that it can be used to make local-area estimates. Although an ACS follow-up study on social mobility would not likely seek to interview most ACS respondents (for financial reasons), the availability of detailed geographic information—accessible in RDCs—means that researchers could study the impact of neighborhoods and other geographic factors on mobility rates and processes.

Sustainability

Current plans call for continuing the ACS into the foreseeable future. Thus it is conceivable that an ACS follow-up study for the purposes of studying social and economic mobility could be replicated periodically. Indeed it may be best to design the follow-up study as a multi-year effort to follow-up a percentage of ACS respondents each year for multiple years.

Financial expense

The cost per respondent of fielding an ACS follow-up study would be lower than fielding a new stand-alone survey. This is because the follow-up study would piggy-back on the ACS's sampling and fieldwork operations. Because of the likely impossibility of re-interviewing anything close to all ACS respondents, and thus the need to sample them, the sampling rate and strategy could be designed so as to make efficient use of resources.

Privacy and data access

Just as the Census Bureau now provides a public use ACS file—after recording variables and otherwise perturbing the data to protect respondents' privacy—it is conceivable that a public use file of an ACS follow-up survey could be prepared and disseminated in such a way that minimizes risks to respondents and does not require researchers to utilize RDC's or other restricted use data access protocols.

Bottom line

An ACS follow-on survey—modeled, for example, on the NSCG—would produce new data on contemporary patterns of social and economic mobility that would include very large samples and a wide range of measures of the social and economic circumstances of sample members and their family members. As with the CPS options outlined above, on a per-respondent basis an ACS follow-on survey would likely be less expensive than a new standalone survey.

Compared to either of the CPS-based approaches described above, an ACS follow-on study would have two key advantages. First, unlike the CPS, the ACS would include people living in institutions. Second, the size of the ACS sample is such that a follow-on study of social mobility could be very strategic in selecting sample members. The larger of the two CPS-based options would include a maximum of 140,000 individuals; this would simply involve surveying everyone, with no strategic oversampling. On the other hand, an ACS-based study could strategically over-sample large numbers of particular geographic or demographic groups (e.g., foreign-born Hispanics, people living in central cities, college-educated African Americans) and conduct the same overall number of interviews.

However, as with the nine smaller-*n* sample surveys and either of the CPS options, it would not be possible to use ACS-based data to fully consider the roles of demographic factors—fertility, child mortality, international migration, and assortative mating—in transforming and reproducing distributions of social and economic resources across generations. On its own, the ACS sample is representative only of people in G0, not of people in their parents' or children's generation.

Evaluating Administrative Record Data

Administrative records—from the Social Security Administration (SSA), the Internal Revenue Service (IRS), or elsewhere—could serve as useful *supplements* to any or all of the options described above. A number of the smaller-*n* surveys describe earlier have already been linked to such records, and the availability of Personal Identification Keys (PIKs) on the Census Bureau's internal versions of CPS and ACS data make possible linkages of those data to records from SSA, IRS, the Veterans Administration, the National Death Index, and others. These are all useful *supplements* to these several data resources. In general, however, these administrative record linkages primarily expand the set of measures available within a data resource—for example, adding additional detail about income or benefit receipt. In most cases, existing plans or strategies for administrative record linkage do not constitute a stand-alone data resource for studying social and economic mobility.[2] In this section I describe one model of administrative record linkage that would constitute a major new social mobility study; later, I reflect on the proposed Intergenerational Master File which is similar in some respects.

PIKs are internal Census Bureau identifiers that uniquely represent each individual and that remain constant for individuals across surveys and over time; they are not equivalent to Social Security Numbers, but instead are random numbers that are related to them. The Census Bureau has assigned PIKs to internal versions of many of its data products, including ACS records and the 2000 decennial census. It is possible, with some effort and expense (described below), for the Census Bureau to PIK records from the 1990 and previous decennial censuses.

A new study of mobility could be conducted by linking 1990 U.S. Census records with ACS records from recent years. The strategy would be to identify people (1) who were children living with at least one parent in the 1990 Census and (2) who were also ACS respondents as young adults between (for example) 2006 and 2014. This would allow us to observe parents' attributes in 1990 and their children's attributes 16 to 24 years later.

How many people would appear in both the 1990 Census (as 9 to 16 year old children of household heads) and in the ACS (as adults) in some year between 2006 and 2014? To find out, I first selected the 1,090,488 people in the 2006-09 public-use ACS files who were born between 1974 and 1981 (and who thus would have been between the ages of 9 and 16 in 1990). Second, I excluded the 157,380 individuals who were born abroad and who moved to the U.S in 1990 or later (and who thus could not have appeared in the 1990Census). These two steps left 933,108 people who would have been living in the U.S. and who would have been between the ages of 9 and 16 in 1990. Third, because we would have five additional ACS years (2010 through 2014), I inflated this figure by 9/4, so that (I estimate that) we would observe 933,108×(9/4)=2,099,493 people in the combined 2006-2014 ACS files who were born between 1974 and 1981 and who lived in the U.S. in 1990. Fourth, I multiplied this figure by 1/6 to account for the fact that only 1 in 6 of these 2,099,493 individuals

[2]One exception is ongoing work at the Stanford Center on Poverty and Inequality to link IRS records across generations. This effort is novel and will produce valuable new insights, but the size of the sample and the narrow scope of the available measures will limit its broader utility.

would have lived in a household that was selected to complete the long form of the 1990 Census. This means that we would expect to observe 2,099,493×(1/6)=349,916 age-appropriate individuals whose household completed the 1990 Census long form and who also appeared in the ACS between 2006 and 2014.[3]

Population definition and coverage

Both the ACS samples and the 1990 Census long form samples represent cross sections of all Americans. By design, people in institutions and other group quarters are sampled. Neither the 1990 Census nor the ACS is perfect with respect to population coverage and response rates, but neither systematically excludes entire sub-groups.

The linked 1990 Census-ACS sample of (approximately) 350,000 individuals—who comprise G0—would be representative of the population of people who were born between 1974 and 1981; who were alive and still lived in the U.S. in 1990; and who survived long enough to be observed in the ACS in some more recent year. As described to this point, this sample would *not* constitute a representative cross-section of people in this birth cohort in 2006-2014, since people who entered the U.S. from another country after 1990 would be excluded. It would also *not* include a representative cross-section of children in 1990, since those who did not survive to 2006-2014 would be excluded. Finally, the sample of 350,000 parents (or pairs of parents) in 1990 (G-1) would *not* constitute a representative sample of people in G-1 since people without children living in their household would be excluded and people with more than one child living in their household would be over-represented. Unlike the GSS, SIPP, CPS, or other ACS designs described above, this design would describe just one birth cohort (albeit a recent one).

A unique virtue of this design, however, is that *unlinked* individuals are also observed. That is, this design begins with representative cross-sections of individuals in 1990 and 2006-2014. The 350,000 linked records are just a subset of the full samples. This makes it possible to study the ways in which immigration of young people after 1990 shapes the distribution of social and economic circumstances. It also makes it possible to study the role of selective mortality among young people after 1990 in shaping these distributions. In the parents' generation, the availability of the entire 1990 Census file makes it possible to study the impact of differential fertility. Altogether, the availability of the full 1990 Census file, the full 2006-2014 ACS file, and the linked subset of ACS-Census records facilitates research on both (1) rates of intergenerational mobility and (2) the role of demographic processes in reproducing and transforming distributions of social and economic outcomes across generations.

[3]Going forward, in the longer term it would be possible to conduct annual mobility studies by linking widely spaced ACS records. The strategy would be to identify people (1) who were children living with at least one parent in the ACS in Year X and (2) who were also ACS respondents as young adults 25 to 30 years later. This would allow us to observe parents' attributes in Year X and their children's attributes as young adults. In recent years, the ACS has included about 650,000 people under age 18. Assuming that the ACS continues to sample about one percent of the U.S. population going forward, this means that we would expect about 650,000×0.01=6,500 linked records each year. Although linkage rates are not perfect and some individuals may have died in the interim, these data would be inexpensive to obtain since the ACS data are already routinely PIK'd. Of course, data from such a design could not be obtained or analyzed for another two or three decades.

Sample size

The tabulations above suggest that about 350,000 people in this birth cohort would appear in both the 1990 Census long form sample and the ACS sample in some year between 2006 and 2014. Of course, not *all* of these 350,000 individuals would have been living with both their mother and father at the time of the 1990 enumeration. To produce a *conservative* estimate of the number of observations for which we would expect to observe parents' attributes in 1990, I deflated the expected number of observations by the sex- and race/ethnic-group specific percentage of 9 to 16 year olds who lived with their *fathers* at the time of the 1990 enumeration. Analyses that utilized *mothers'* (or *either* parents') characteristics would always have *higher* numbers of observations than those reported below.

For example, we would expect to observe 38,540 age-appropriate non-Hispanic U.S.-born blacks in both the 1990 Census and the 2006-14 ACS. However, only 43.5% of people in this group lived with their fathers in 1990. Consequently, we would expect to observe 38,540×0.435=16,765 age-appropriate non-Hispanic U.S.-born blacks—and 271,461 age-appropriate individuals overall—who were enumerated as living with their fathers on the 1990 Census long form and who also appear in the ACS between 2006 and 2014. Consequently, this design would permit separate analyses for the full sample (n=271,461), men (n=133,062), women (n=138,399), US-born non-Latino Whites (n=205,634), US-born non-Latino Blacks (n=16,773), US-born Latinos (n=22,703), foreign-born Latinos (n=1,938), US-born American Indians and Alaskan Natives (n=3,371), US-born Asians and Pacific Islanders (n=4,689), and foreign-born Asians and Pacific Islanders (n=4,044). This procedure would also likely produce sufficient numbers of observations in urban/rural/suburban areas.

Topical coverage

Both the 1990 Census and the 2006-2014 ACS will have gathered information about each household member's educational attainment; labor force status; industry and occupation; place of birth and citizenship; language use; military service; and personal income (including wage and salary income, self-employment income, interest income, SSI receipt, retirement income, and other income). It is thus possible to consider mobility from G-1 to G0 with respect to any of these measures of social and economic circumstances.

Temporal issues

This design would provide information about both rates of intergenerational mobility and the demographic factors that shape the distribution of social and economic resources among a birth cohort that has recently reached young adulthood.

However, it would be difficult to compare rates of intergenerational mobility observed using this design to those observed in any other U.S. mobility study. Unlike the GSS, CPS, or SIPP designs described above, G0 in the linked 1990 Census-ACS design is not a representative cross-section of people in any generation. One possibility, of course, would be to use a parallel design to link (publicly available) records from the 1850 through 1940 U.S. Censuses.

Spatial issues

As with an ACS follow-on design, a key strength of the ACS for this purpose is that it can be used to make local-area estimates. Again, the availability of detailed geographic information in the ACS—accessible in RDCs—means that researchers could study the impact of neighborhoods and other geographic factors on mobility rates and processes.

Sustainability

Given the termination of the decennial census long form after 2000, it will not be possible to precisely replicate this design in the future. It would be possible going forward to link future ACS records to 1990 decennial census records for people born in this cohort; we would expect one percent of people observed as 9 to 16 year olds in the 1990 Census to appear in each new ACS. It would also be possible going forward to link people across ACS years, although the sample would be considerably smaller (and although there are design elements in the ACS that prevent people from being interviewed too frequently). Finally, this entire design could be replicated by linking the 2000 decennial census to 2016-2024 ACS records.

Financial expense

If records from the 1990 Census were PIK'd, then the per observation monetary expense of linking the records and executing the analyses would be *trivial* relative to the cost of new data collection. However, the 1990 U.S. Census has *not* been PIK'd—because the names have not been digitized. In my communications with Census Bureau staff, a ballpark estimate of the expense of digitizing the names for everyone in the long form sample is $5 million. It is conceivable that this cost could be reduced substantially by digitizing only the names of people living in households that contain a member of the focal birth cohort. However, it is less clear that the Census Bureau would be willing to do such a partial PIK'ing of the 1990 long form sample (because it would be less useful for their purposes).

Privacy and data access

Personally identifiable ACS records are tightly controlled by the Census Bureau. Only Census Bureau staff would be authorized to digitize or PIK 1990 U.S. Census records, and linked ACS-1990 Census records would only be available to researchers in an RDC. Indeed a major hurdle to executing this design would be securing necessary permissions from the Census Bureau to digitize names, PIK the 1990 long form sample, and link records from two surveys to which people are legally obligated to respond. Even should all of these permissions be obtained, researchers would need to obtain pre-approval for all analyses, and their results would need to pass through disclosure review before they could leave the RDC.

Bottom line

Linking ACS records to the 1990 U.S. Census for one birth cohort of Americans—i.e., those born between 1974 and 1981—presents unique opportunities and also brings challenges that most of the other designs do not face.

On the positive side, this strategy would yield a sample of some 350,000 young adults and would include detailed measures of their own and their parents' social and economic

characteristics—indeed, more of those characteristics for both G0 and G-1 than for any other data resource considered to this point. This strategy would include institutionalized individuals in both G0 and G-1, and would likely be far cheaper from a financial point of view because no new fieldwork would be required. Finally, the combination of representative cross-sectional data on people in this cohort in the 2006-2014 ACS; of linked records for members of this birth cohort; and of representative cross-sectional data on their parents' generation in 1990 would allow us to consider the roles of migration, fertility, and mortality—as well as intergenerational mobility—in reshaping distributions of social and economic resources over time.

On the other hand, this design would yield data on just one birth cohort of Americans, albeit a recent one. What is more, the challenges associated with obtaining necessary permission to PIK the 1990 records and link them to ACS records are considerable, especially given that the resulting data would only be available in the highly restrictive RDC setting.

Evaluating a New Stand Alone Survey

One could field a new data collection operation that allowed for an updated assessment of rates and patterns of social mobility in the United States. These new data could potentially be linked to SSA, IRS, school transcript, Census Bureau, and other administrative data. This endeavor would likely entail multiple modes of interview (i.e., by phone, via the web); would be constrained with respect to sample size and design complexity; but could enjoy wider coverage of topical domains. In this section I review the costs and benefits of this choice relative to the criteria outlined above.

Population definition and coverage

It would be straightforward to follow the example of many of the surveys described above and generate a sample that reflects the characteristics of the population of individuals in G0 (including those in institutions). One could then ask questions about sample members' parents and/or children and model mobility from G-1 to G0 and/or from G0 to G+1. If questions were asked about extended family members, one could implement a sibling design and one could model mobility across more than two generations.

With more creativity (and resources), however, one could design a study that facilitate research on both (1) rates and patterns of intergenerational mobility and (2) the ways in which fertility, mortality, and migration re-shape or reproduce distributions of social and economic resources over time. The “standard” sampling design is to draw a representative sample of people in G0 and then ask people in G0 about themselves, their parents and parents-in-law (G-1), and/or their children (G+1). This design could be augmented in such a way that would generate representative data on all three generations, G-1, G0, and G+1. For example, after drawing a sample that represented the cross-section of people in G0 (including those in institutions), one could then add a supplementary sample of people in G-1—specifically, those who had no children or whose children did not survive to be at risk of observation in G0; in some cases it may be necessary to collect information from widows or other proxies for deceased people. These additional members of G-1 would have no corresponding children in the G0 sample. One could likewise add a supplementary sample of

people in G+1—specifically, those whose parents were not eligible to be included in G0 because they lived outside of the United States. Again, these new members of G+1 would have no corresponding parents in G0. This design—a representative cross-section of G0 with supplementary samples of people in G-1 and G+1—would allow for analyses of social and economic mobility across three generations and would also allow for analyses of the ways in which demographic impact the distribution of social and economic resources over time.

Sample size

A new data collection effort would likely be constrained with respect to total sample size (at least as compared to alternative designs based on the CPS or ACS). However, as with some of the smaller-*n* sample surveys described above, strategic oversampling might allow for sub-group analyses.

Topical coverage

A key virtue of a new data collection effort is that it would allow for a richer set of measures —obtained for people in each focal generation—than is likely possible in any of the options discussed to this point. It would be possible to obtain measures of routinely collected attributes like education, income, and occupation, but it would also be possible to observe things like (in no particular order): cognitive and non-cognitive skills; job quality and working conditions; wealth and debt; subjective social class; voting behaviors; incarceration experiences; health and disability; and neighborhood characteristics.

What is more, it would be possible to obtain measures of things hypothesized to serve as mechanisms through which advantage and disadvantage are transmitted across generations.

Temporal issues

A study of social and economic mobility using newly collected data would include observations of a contemporary cross-section of Americans. Although the sampling design may be more complicated than previous studies, it would seem feasible to collect new data in such a way that facilitates comparisons over time (e.g., by allowing for comparisons to NLSY-79 or the 1986-1988 SIPP panels).

Spatial issues

Another virtue of new data collection is that one could collect detailed geographic information and other spatial data (e.g., detailed neighborhood conditions). It would also seem feasible to design a new data collection effort that allowed for comparison to parallel samples in other countries.

Sustainability

It would be important to view a new data collection effort as a long-term effort to establish a data system for continually monitoring social and economic mobility in the United States—not as a one-time effort. That is, it would be beneficial in the long-term to design from the outset a data collection strategy that would be sustainable.

Financial expense

A key downside of collecting new data for this purpose is the likely financial expense. Designing and drawing samples, developing new multi-method fieldwork protocols, and implementing new data cleaning and data access systems are all expensive. The GSS, SIPP, CPS, and ACS based alternative described above all economize by building on the broader efforts of those existing data collection operations. The consequence of the additional expense of new data collection would probably be trade-offs in terms of sample size and complexity and in the likely size or scope of the survey instrument itself.

Privacy and data access

New data collection for this purpose would not necessarily entail any novel challenges with respect to privacy or data access. It would seem quite feasible to follow standard protocols for restricting access to sensitive information or to information that might be used to deduce the identities of respondents. There would certainly be fewer issues with respect to data access as compared to the CPS or ACS designs described above.

Bottom line

Fielding a new survey for the purposes of studying contemporary patterns of social and economic mobility is an attractive option for a number of reasons. New data collection would allow for a wider range of measures of social and economic circumstances than are typically available in extant surveys; the new data could also be linked to administrative record data. A new data collection effort could employ strategic sampling designs to expand the sort of analyses that could be performed and the range of questions that could be addressed. Finally, a new data collection effort could be designed from the outset to observe sibling pairs, to observe extended family networks, and to be comparable both over time and across countries.

The key downside of new data collection is its considerable financial expense. The high cost of new data collection would likely mean tradeoffs in terms of the scope and content of surveys and in terms of the complexity and size of the sample. It is not clear whether the main virtues of new data collection—a strategically drawn sample, rich and sophisticated measures of social and economic circumstances—could be fully realized given their likely price tag. Another downside is that social surveys administered by universities or private research organizations are likely to realize lower response rates that surveys like the CPS, SIPP, or the ACS.

Evaluating the Proposed Intergenerational Master File

Finally, how does the planned Intergenerational Master File (IMF) described in this volume (Grusky, Smeeding and Snipp Forthcoming) compare to the ideal-typical scenario outlined above and to the other options evaluated in this article? To begin, let us assume that (a) PIK'ing the 1990 U.S. Census can be accomplished for the vast majority of individuals who responded to the long form and (b) it is possible to achieve very high linkage rates across censuses, ACS's, and administrative records. Let us also assume that the political, administrative, and legal hurdles to linking these records can be overcome (Johnson, Massey,

and O'Hara Forthcoming). These are very strong assumptions, and it is not entirely clear that they can ever be met, but even if they can: How would the resulting IMF compare to the ideals outlined above?

Population definition and coverage

None of the data resources that form the core of the IMF systematically exclude entire sub-groups of the population; all are designed to cover institutionalized and non-institutionalized individuals alike.

Whereas the design described above (which involved linking children in the 1990 Census to their responses to subsequent ACS surveys) would generate a sample that generalized only to a particular birth cohort, the IMF would include individuals from a wide range of cohorts. However, because of the complex design of the IMF, researchers will have to think carefully about the population to which their results can be generalized. For example, imagine using the kidlink file to identify parent-child matches, and then imagine linking IRS records about individuals' incomes to those of their children. One could then presumably study intergenerational income mobility. However, neither the parents nor the children would be representative of any particular population. The older generation would exclude people who had no children who survived to adulthood (and would over-represent those who had multiple children); the younger generation would exclude those with parents who lived abroad. The fact that the data resources that comprise the IMF also observe people who cannot be linked is a virtue in this regard; it is possible to consider the exclusivity or representativeness of various lined samples. However, unlike any of the designs described above, researchers will have to work much harder to establish the external validity of their results.

Sample size

The basic design of the IMF suggests that it will include a large number of observations. However, the more data elements that are added, the lower the overall linkage rates and the greater the number of sample exclusions. For example: It is not clear how many people responded to the 1990 Census long form, have appeared in the ACS, and also have IRS records available for themselves and their children or parents. Especially for smaller population sub-groups, it is not yet clear how large the IMF will be.

Topical coverage

To the extent that the IMF is based on the 1990 Census and the ACS, it will observe core outcomes for multiple generations of people, including: educational attainment; labor force status; industry and occupation; place of birth and citizenship; language use; military service; and personal income (including wage and salary income, self-employment income, interest income, SSI receipt, retirement income, and other income). Plans to supplement the IMF with records from the SSA, the IRS, and elsewhere would seem most likely to improve the quality of income and program participation measures; only occasionally would they add new topical content. However, the capacity to link the IMF to records from stand-alone surveys would seem to have the potential to enhance content in useful ways. Of course, the

latter option is premised on the capacity to obtain survey respondents' permission to make such links.

Will the IMF involve higher-quality measures of core topical areas than might be achieved in stand-alone surveys? A stated virtue of the IMF is “the capacity to exploit high-quality administrative data and high-quality Census products rather than field new and almost inevitably lower-quality surveys” (Grusky, Smeeding and Snipp Forthcoming). In my view, this remains to be seen. It seems quite plausible that IRS income records are of higher quality than can be obtained using self-reported income from surveys. Perhaps the same is true for program participation. However, it is not clear that administrative records of people's occupations are of very high quality (either in absolute terms or compared to surveys). And, none of the components of the IMF rely on reports of educational attainment beyond respondents' self-reports.

Temporal issues

It is not clear how the IMF could be used to study trends over time in mobility, at least looking backward. Once established, and after some time has passed, the IMF could certainly be used to study trends in mobility from perhaps the 1990s onward. However, it would not seem possible to easily compare results generated using IMF data to any pre-IMF results (because the IMF sample design is so different from anything that precedes it). Nonetheless, because the design of the IMF includes individuals from multiple birth cohorts, it would be possible to study trends across birth cohorts in social and economic mobility.

The design of the IMF is such that researchers would also have to pay attention to a few additional temporal issues. In particular, the design allows for observing parents and children at different ages. Should the attributes of G0 only be measured when people in G+1 were children? Or at any point in the life course of the G+1 generation? Should the attributes of G+1 be observed in young adulthood, later adulthood, or averaged across ages? In research on intergeneration mobility, should we ever consider observations when the attributes of children (e.g. income) are observed temporally prior to the parallel observation for parents? For example, the design of the IMF is such that we may observe income for people in G+1 in 1990 and the income of their (presumably elderly) parents in G0 in 2000. How should we treat such observations?

Spatial issues

The IMF would not likely permit cross-national comparison. Even in countries with administrative record data that permit analyses of intergenerational mobility, the design of those data files would seem to be sufficiently different from the IMF to make comparative analyses very difficult.

However, the design of the IMF is such that spatial analyses within the U.S. (e.g., of rural/urban differences or by region) would be possible. Indeed observing the same individuals at multiple points in time (or, perhaps, continuously in administrative record data) makes possible analyses of the role of geographic mobility in processes of socioeconomic mobility.

Sustainability

A stated virtue of the IMF is its sustainability: “[T]he development of a monitoring infrastructure that, by virtue of being automatically ‘refreshing,’ sidesteps the problems with unrepresentativeness that plague other long-running panels (e.g., the PSID)” (Grusky, Smeeding and Snipp Forthcoming). While it is true that the components of the IMF—the ACS, IRS data, etc.—will continue to be collected for the foreseeable future, this does not guarantee sustainability. First, the capacity to link newly collected data (e.g., a new ACS, or another year of IRS data) depends on various sorts administrative and political good-will. It is not clear that the ability of construct the IMF using currently existing data will necessarily mean that new data can be added as they are collected. Also, like any of the data options that rely on restricted access to administrative data via RDCs or similar arrangements, access to the data is tightly restricted. The sustainability of any mobility study based on such data is thus tenuous going forward (as compared to options that rely only on publicly-available data).

Financial expense

Beyond the expense of PIK'ing the 1990 Census, the IMF is among the least expensive options. Putting aside efforts to supplement the IMF with survey data, the primary expense of the IMF going forward would be to link new records as they become available.

Privacy and data access

As I noted at the outset, every data option for a new social mobility project involves a trade-off between data security and data access. The IMF, by virtue of relying on some of the most restricted data about individuals, would thus require considerable efforts to maintain its security. Even the *least* restricted IMF data would be available only in an RDC. A serious downside to this fact is reduced data access. The investment of time, money, and administrative capacity to gain access to the IMF will be more than many researchers can afford.

Bottom line

The IMF represents a bold new design (at least for the United States) for studying social and economic mobility. It makes creative and ambitious use of currently existing administrative data, and takes advantage of recent advances in record linkage technology, the growth of Census RDCs, and the willingness of the Census Bureau and other agencies to entertain such possibilities. On the other hand, the analytic virtues of the IMF remain unclear. The sample may be large, but to what population can results be generalized? How high are linkage rates likely to be, especially when multiple sources of data are linked, and what are the implications for sub-group analysis? How can we use results from the IMF to compare social and economic mobility over time or across countries? How do we maintain the strict data security provisions that the IMF will require while at the same time encouraging innovation and discovery by making the data accessible as widely as possible?

Conclusions

None of the options I have described is ideal with respect to the parameters I outlined at the outset. Each brings its own unique strengths and limitations and costs and benefits. I have outlined the pros and cons of investing in the IMF. But what if the IMF proves to be impractical, for financial, administrative, or legal reasons? Below I review (in no particular order) five promising opportunities for generating data that would facilitate new research on social and economic mobility in the contemporary United States. My view is *not* that we should select just one of these, *or* that we should try to pursue all five. Some combination of them, however, would facilitate a new generation of research on intergenerational mobility—regardless of whether the IMF goes forward.

1. Continue to Invest in the GSS

Despite its smaller sample size and limited topical content, the GSS has long served as a key resource for studying social and economic mobility in the United States. In particular, it will almost certainly continue to be the way we analyze longer-term trends in mobility patterns. Regardless of what investments we make in new data resources, we should continue to invest in the GSS.

Beyond its current sample design and topical content, we should continue and increase efforts to make the GSS useful for mobility research. For example, recent initiatives to recode all GSS occupations to consistent and modern classification schemes will expand the sorts of analyses that we can do. We might also continue to press for richer measures of income, or to expand the range of questions that we ask about GSS respondents' parents. Finally, we should invest in a new topical module on social mobility that would allow for temporal comparisons to the 1994 GSS family mobility module; in particular, this module should collect information about respondents' spouses, spouses' parents, and children. All of these efforts are inexpensive (compared to the options below) and would produce important new information.

2. Add Relevant Content to the New SIPP Panel

The planned 2014 SIPP panel provides a great opportunity for research on social and economic mobility. At least compared to other smaller-*n* surveys, SIPP has traditionally had a larger sample size. As depicted in Figure 1, the new SIPP panel would reflect the contemporary characteristics of a cross-section of Americans from multiple birth cohorts. Finally, the core content of the SIPP has traditionally included key social and economic information about people in G0, including education, labor force activities, occupation, and rich measures of income and program participation. Some of that information could be validated by linkages to administrative record data.

If the new SIPP panel also obtained parallel information about SIPP respondents' parents and/or children, it would immediately serve as an important new data resource for studying social and economic mobility. We should press to have SIPP include—perhaps as part of a topical module or supplement—questions about the social and economic characteristics of panelists' family members. To facilitate comparisons over time, some of these questions

might parallel those asked in the 1986-1988 SIPP supplements; however, it would also be necessary to extend and improve upon those items.

3. Supplement the CPS

Supplementing the CPS—using either the "OCG III" or "February Mobility Supplement" designs described above—would produce new data on contemporary patterns of social and economic mobility. These data would include large samples and high quality measures of the social and economic circumstances of sample members and their family members. On a per-respondent basis, this option would likely be less expensive than a new stand-alone survey.

Which of the two designs we should implement depends on their relative practicalities and expense. Whereas there are multiple models in place for funding and fielding new monthly supplements, there are fewer such models for fielding follow-on surveys like ATUS. For reasons outlined earlier, the February Mobility Supplement would permit a larger sample size but would likely allow for a smaller number and narrower range of new survey questions. A key virtue of either CPS-based design—at least as compared to new data collection—-is that it would take advantage of the fact that we already know a great deal about respondents from the CPS basic monthly and supplemental surveys, and in particular the March ASEC.

4. Supplement or Link the ACS

The ACS's enormous sample size, its broader population coverage, and its broad range of measures of individuals' social and economic circumstances make it an especially attractive resource for gathering new data on social and economic mobility.

An ACS follow-on study would have the key virtues of either of the CPS-based options described above, but it would also have several comparative advantages. Unlike the CPS, the ACS includes people living in institutions (and particularly prisons). Furthermore, the size of the ACS sample is such that a follow-on study could be very strategic in selecting sample members so as to efficiently facilitate sub-group analyses. Although an ACS follow-on study might include the same total number of respondents as a CPS follow-on study, the former would include a more complete cross-section of Americans and would more efficiently allow for an assessment of how mobility processes vary across social, demographic, and geographic groups.

Linking ACS and 1990 Census records as described above would yield a sample of about 350,000 current young adults and would include detailed measures of an especially rich array their own and their parents' social and economic characteristics. This strategy would likely be far less expensive than collecting new data or than fielding a CPS or ACS follow-on survey because no new fieldwork would be required. Finally, by combining (1) representative cross-sectional data on people in one birth cohort in the 2006-2014 ACS, (2) linked ACS-Census records for members of same birth cohort and (3) representative cross-sectional data on their parents' generation in 1990 we would be able to consider the roles of migration, fertility, and mortality in reshaping distributions of social and economic resources over time. Despite its virtues, this design would be limited to studying just one birth cohort. Finally, the challenges of obtaining necessary permissions to implement this design are

considerable, and the resulting data would not be as easily available to researchers as the data produced using other designs.

5. Collect New Data

Collecting new data would allow great flexibility with respect to survey content and sampling strategy. We could observe aspects of people's and their family members' social and economic conditions that are not typically observed elsewhere, and we could measure things hypothesized to mediate the effects of social origins on life outcomes. At the same time, we could draw our sample in such a way as to maximize the sorts of questions we can pose using the data. However, new data collection is very expensive. Instead of building on existing survey operations' sampling, fieldwork, instrumentation, and data processing capacities, we would build much of this from scratch. These additional expenses would have the consequence of constraining sample size and survey length, thereby eroding the benefits of new data collection.

Appendix: Descriptions of Existing Surveys

American Community Survey (ACS)

The ACS—first fully implemented in 2005 by the Census Bureau as a replacement for the decennial census's long form (U.S. Census Bureau 2006b; U.S. Census Bureau 2009)—samples Americans living in households and group quarters (which are most commonly prisons). Since 2005, the ACS has sampled 1 in 40 U.S. addresses each year (a sample density sufficient to produce small-area estimates by pooling data across three to five years). The ACS collects demographic, social, economic, and housing characteristic data every year. Although the survey observes each household member's education, income, occupation, labor force status, veteran status, family circumstances, health insurance coverage, and disability status, it includes no information about respondents' parents or children unless they also live in the sampled household.

Like many other federal data, ACS records have been assigned Personal Identification Keys (PIKs) by the Census Bureau. Consequently, ACS records can technically be linked to any other federal data that have also been PIK'ed—including other Census Bureau data (e.g., the decennial census), the Current Population Survey, Internal Revenue Service data, Veterans' Administration data, and Social Security Administration records. Such linkages can only be made with permission from and in cooperation with the Census Bureau within an RDC.

Current Population Survey (CPS)

The CPS is conducted jointly by the U.S. Census Bureau and the Bureau of Labor Statistics (U.S. Census Bureau 2006a). The CPS sample is representative of the civilian, household-based population of the United States. In recent years, each monthly CPS has included about 140,000 individuals living in about 60,000 households. Occupants of sampled households are interviewed in four consecutive months, are rotated out of the panel for eight months, and are then re-interviewed in four consecutive months. Because the CPS samples housing

units, not people, individuals who move out of CPS households are not followed and new people who subsequently move into CPS households are included in the panel.

The basic monthly CPS survey gathers demographic data and extensive information about labor force activities, occupation and industry, and unemployment. In most months the CPS also includes a topical supplement; the number, topics, and timing of these supplements has changed over time. Since 1968, the October supplement has gathered additional information about school enrollment and educational experiences. For many years the March supplement (now called the Annual Social and Economic Supplement) has gathered detailed data on wages and other income. In March of 1962, Blau and Duncan (1967) administered a “leave behind” paper survey to 20,700 20 to 64 year old men in the CPS; their Occupational Changes in a Generation (OCG) survey collected additional information about respondents' parents and families of origin. In March of 1973, Featherman and Hauser (1978) largely replicated the OCG with a mailout/mailback survey of 20 to 65 year old male CPS respondents.

As with the ACS, PIKs can be used to link CPS records to other federal data that have also been PIK'ed. For example, Davern et al. (2009) recently used linked CPS-Medicaid data to consider rates of underreporting of Medicaid use in the CPS.

General Social Survey (GSS)

The GSS—administered by NORC (Smith et al. 2013)—observes samples of non-institutionalized adults. In most years between 1972 and 2004, the GSS interviewed fresh cross-sections of adults. In 2006, it switched to a rotating panel design. Under the new design, about 2,000 respondents are interviewed (usually in person) and then re-interviewed two and four years later. From 2010 onward, the GSS will include interviews with a fresh cross-section of about 2,000 respondents; interviews with about 2,000 respondents first surveyed two years earlier; and interviews with about 2,000 respondents first surveyed four years earlier. The GSS collects measures of a core set of social, economic, demographic, and attitudinal attributes of respondents; in most years it has also included topical supplements. Since its inception the GSS has included survey questions about respondents' parents' educations and occupations (e.g., Diprete and Grusky 1990). A 1994 topical module also gathered extensive socioeconomic information about respondents' parents, spouses, siblings, and children (e.g., Goldstein and Warren 2000).

Health and Retirement Survey (HRS)

The HRS—also conducted at the Institute for Social Research at the University of Michigan—is a longitudinal study that now surveys more than 26,000 Americans over the age of 50 every other year. Members of the original HRS cohort were born between 1931 and 1941, were first interviewed in 1992, and have been re-interviewed every two years through 2012. Over time, earlier and later cohorts have been added; the earliest cohort (born 1890 to 1923) was first interviewed in 1993, while the most recent cohort (born 1954 to 1959) was first interviewed in 2010.

The HRS routinely collects information about income, assets, and debt; pension and insurance coverage; labor market activities; family structure and dynamics; care-giving and receiving and financial inter-transfers; disability, health and physical functioning; cognitive functioning and psychological well-being; and health care and other expenditures. The HRS also includes experimental topical modules typically focused on issues related to retirement, health and wellness, and financial issues. HRS records have been to data from the Social Security Administration, the National Death Index, Medicare data, and pension records. Since its inception in 1992, the HRS core survey has asked about respondents' parents' educations, father's joblessness, and relative family income. The survey also includes information about panel members' children's education and labor force activities.

High School & Beyond (HSB)

HSB—funded by the National Center for Educational Statistics (NCES) and fielded by NORC—began in 1980 with a sample of 30,030 sophomores and 28,240 seniors in 1,015 public and private high schools. Unlike PT, HSB sampled relatively small numbers of students within schools; however, the design included oversamples of some groups (e.g., Catholic schools, Hispanic students). The 1980 survey gathered information about educational plans and experiences; cognitive and non-cognitive skills; peers groups; employment activities; and occupational plans and aspirations. Subsets of respondents were then re-interviewed---with relatively high rates of participation—in 1982, 1984, 1986, and 1992. The latter surveys measured educational, employment, and family circumstances and transitions; included additional cognitive tests; and gathered secondary and post-secondary transcripts. A new follow-up with the HSB sophomore cohort was conducted in 2013-2014, and a follow-up with the HSB senior cohort is planned for 2015.

The 1980 HSB instrument collected information about parents' educations and occupation and family composition, and follow-up surveys gathered information about labor force and occupational outcomes; the 2013-2015 surveys include batteries of questions about labor force circumstances. The sampling design of the 1980 survey included all twins related to sampled students. As a result, the sample includes 524 pairs of twins and 810 other sibling pairs (U.S Department of Education 2001).

National Longitudinal Study of Adolescent Health (Add Health)

Add Health—conducted at the University of North Carolina at Chapel Hill (Harris 2011)—is following a sample of more than 20,000 adolescents who were in grades 7 through 12 during the 1994-1995 school year. In-school and in-home surveys of children and interviews with parents were conducted in 1994-1995; these were followed by in-home interviews, biomarker collection, and other assessments in 1996, 2001-2002, and 2008-2009 (Harris et al. 2009). Add Health was designed to include oversamples of several population groups including 784 pairs of twins and 1,251 other full sibling pairs (Harris et al. 2006).

Add Health focuses on respondents' social, economic, psychological, and health characteristics; the design is such that is also provides contextual data on respondents' families; neighborhoods; schools; and peer and social networks. The adolescent and parent

questionnaires in the earlier survey rounds collected information about sample members' parents' educations, employment circumstances, demographic characteristics, and health.

National Longitudinal Study of Youth, 1979 (NLSY-79)

NLSY-79—fielded by NORC on behalf of the Bureau of Labor Statistics—is following a sample of 12,686 Americans who were born between 1957 and 1964. The study includes both a representative cross-sectional sample and an oversample of Hispanics and African Americans. Panel members were between the ages of 14 and 22 when they were first interviewed in 1979; they were then re-interviewed annually through 1994 and every other year since then. Because each household member born between 1957 and 1964 was included in the sample, the study began with nearly 2,900 sibling pairs.

NLSY-79 is designed to understand employment and labor force dynamics, but also includes detailed information about job training; military service; income, assets, and program participation; schooling; cognitive and non-cognitive skills; family formation and processes; health; and crime and substance use. The 1979 baseline survey ascertained sample members' parents' educations, occupations, and country of birth. Since 1986 the NLSY-79 has also included interviews with a sample of children of NLSY-79 women; 5,255 of those children were interviewed in 1986. These surveys have focused on home environments, health, and cognitive and psychological developments, educational experiences, and (for young adults) early career and military service experiences.

National Longitudinal Study of Youth, 1997 (NLSY-97)

NLSY-97—also fielded by NORC on behalf of the Bureau of Labor Statistics—is following a sample of 8,984 Americans who were born between 1980 and 1984. The study includes both a representative cross-sectional sample and an oversample of Hispanics and African Americans. Panel members were between the ages of 12 and 17 when they were first interviewed in 1997, and they have been interviewed annually since then. Because each household member born between 1980 and 1984 was included in the sample, the study began with nearly 2,000 sibling pairs.

Like NLSY-79, NLSY-97 is primarily designed to understand employment and labor force dynamics, but also includes detailed information about income, assets, and program participation; schooling; cognitive and non-cognitive skills; family formation and processes; health; and crime and substance use. In addition to collecting information from panel members, in its base year NLSY-97 fielded a parent questionnaire that gathered information about sample members' parents' educations, employment circumstances, income and assets, demographic characteristics, and health. The 1997 parent questionnaire also collected information about the educational attainments of parents' parents—that is, about the educations of panel members' grandparents.

Panel Study of Income Dynamics (PSID)

The PSID—conducted at the Institute for Social Research at the University of Michigan (Duncan, Hofferth and Stafford 2004)—began in 1968 with a household sample of more

than 18,000 Americans living in 5,000 families (including an oversample of low-income families). Original panel members have been followed prospectively each year through 1997 and then biennially since then. In addition, the children of original panel members have been followed prospectively as they have moved away to form their own households; as a result, PSID includes a sibling design. The PSID sample has been augmented over time to add post-1968 immigrants, and since 1997 a Child Development Supplement questionnaire has periodically been administered to children (Andreski et al. 2012).

Although the PSID was originally designed to focus on income and poverty dynamics, its content and focus has broadened over time; it is now a study of family and labor force dynamics, education, religion, health and healthcare, and aging and the life course. Original panel members were asked questions about the social and economic circumstances of their families of origin, and as those original panel members' children grew older the PSID came to include information about the social and economic circumstances of multiple generations within families. PSID has been most widely used to study intergeneration income mobility (e.g., Corcoran et al. 1992; Solon 1992), but it has also been used to study other outcomes (e.g., Smeeding, Jäntti and Erikson 2011; Torche 2011).

Project Talent (PT)

PT—housed in the American Institutes for Research (AIR) and first designed and funded by the U.S. Office of Education in the late 1950s—began as a survey of about 440,000 high school students in 1960 (Flanagan et al. 1960). All 9^{th} through 12^{th} graders in more than 1,300 schools were selected to participate. PT attempted re-interviews with selected panelists one, five, and eleven years later, but generally experienced low participation rates to those surveys (Carrel, Potts and Campbell 1975). The 1960 survey included an extensive battery of tests of cognitive skills, non-cognitive skills, and academic achievements and aptitudes, as well as questions about educational and occupational experiences and aspirations. AIR is now planning and testing a new PT follow-up.

The 1960 PT survey included questions about family income parents' occupations, in addition to other aspects of home environments. Because effectively all students in each school were selected into the sample, PT includes about 2,300 pairs of twins and more than 35,000 other pairs of siblings (Prescott et al. 2013).

Survey of Income and Program Participation (SIPP)

SIPP—conducted by the Census Bureau—is a periodic longitudinal survey of adults in U.S. households (U.S. Department of Commerce 2009). New SIPP panels—each with between 12,400 and 41,000 households (National Research Council 2009)—were begun each year between 1984 and 1993, and then every few years thereafter through 2008. Through the 2008 panel, respondents were interviewed every four months over the course of three or four years. Beginning with the planned 2014 panel, respondents will be interviewed annually (and will include a modestly larger number of households). SIPP monitors respondents' income, assets, liabilities, taxes, insurance, and public program participation; it is a major resource for modeling the effectiveness of state and federal economic policy and transfer

programs. Like the GSS, SIPP has frequently included topical supplements—including a topical supplement on family background in the 1986 through 1988 panels (Hauser et al. 2000; Warren, Sheridan and Hauser 1998). At various times in its history, SIPP files have been matched to administrative records from the Social Security Administration (e.g., Mazumder 2005), the Internal Revenue Service, and other agencies (National Research Council 2009; U.S. Department of Commerce 2009).

References

Andreski, Patricia, Beaule, April, Dascola, Mary, Duffy, Denise, Leissou, Eva, McGonagle, Katherine, Schlegel, Jay, Schoeni, Robert. PSID Main Interview User Manual: Release 2012.1. Institute for Social Research: University of Michigan; 2012.

Beaghen, Michael, Stern, Sharon. Usability of the American Community Survey Estimates of the Group Quarters Population for Substate Geographies. 2009 Joint Statistical Meetings: Proceedings of the Survey Research Methods Section American Statistical Association. 2009

Beller E. Bringing Intergenerational Social Mobility Research into the Twenty-first Century: Why Mothers Matter. American Sociological Review. 2009; 74(4):507–28.

Blau, Peter Michael, Duncan, Otis Dudley. The American occupational structure. New York: Wiley; 1967.

Breen R, Jonsson JO. Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. ANnual Review of Sociology. 2005; 31:223–43.

Carrel, Kathleen S., Potts, Clifford A., Campbell, Emily A. Project TALENT's Nonrespondent Follow-Up Survey: The 10th Grade Special Sample. Palo Alto, CA: American Institutes for Research; 1975.

Corcoran M, Gordon R, Laren D, Solon G. The Association between Mens Economic-Status and Their Family and Community Origins. Journal of Human Resources. 1992; 27(4):575–601.

Davern M, Klerman JA, Baugh DK, Call KT, Greenberg GD. An Examination of the Medicaid Undercount in the Current Population Survey: Preliminary Results from Record Linking. Health Services Research. 2009; 44(3):965–87. [PubMed: 19187185]

Diprete TA, Grusky DB. Structure and Trend in the Process of Stratification for American Men and Women. American Journal of Sociology. 1990; 96(1):107–43.

Drew, Julia, Flood, Sarah, Warren, John Robert. Making Full Use of the Longitudinal Design of the Current Population Survey: Methods for Linking Records Across 16 Months. Journal of Social and Economic Measurement. Forthcoming.

Duncan, Greg J., Hofferth, Sandra, Stafford, Frank P. Evolution and Change in Family Income, Wealth, and Health: The Panel Study of Income Dynamics, 1968-2000 and Beyond. In: House, James S., Arbor, Ann, editors. A Telescope on Society: Survey Research and Social Science at the University of Michigan and Beyond. University of Michigan Press; 2004. p. 156-200.

Featherman DL, Hauser RM. Changes in Socioeconomic Stratification of Races, 1962-73. American Journal of Sociology. 1976a; 82(3):621–51.

Featherman DL, Hauser RM. Sexual Inequalities and Socioeconomic Achievement in United-States, 1962-1973. American Sociological Review. 1976b; 41(3):462–83.

Featherman, David L., Hauser, Robert Mason. Opportunity and change. New York: Academic Press; 1978.

Flanagan, John C., Daily, John T., Shaycroft, Marion F., Gorham, William A., Orr, David B., Goldberg, Isadore. Project Talent Monograph Series, Monograph No 1: Designing the Study. Washington, D.C: Cooperative Researcg Program, U.S Office of Education; 1960.

Goldstein JR, Warren JR. Socioeconomic reach and heterogeneity in the extended family: Contours and consequences. Social Science Research. 2000; 29(3):382–404.

Groves, Robert M., Couper, Mick. Nonresponse in household interview surveys. New York: Wiley; 1998.

Grusky, David B., Smeeding, Tim, Snipp, C Matthew. The Case for Assembling an Intergenerational Master File in the United States. Annals of the American Academy of Political and Social Science. Forthcoming.

Harris KM, Halpern CT, Smolen A, Haberstick BC. The National Longitudinal Study of Adolescent Health (Add health) twin data. Twin Research and Human Genetics. 2006; 9(6):988–97. [PubMed: 17254442]

Harris, KM., Halpern, CT., Whitsel, E., Hussey, J., Tabor, J., Entzel, P., Udry, JR. The National Longitudinal Study of Adolescent Health: Research Design. 2009. http://www.cpc.unc.edu/projects/addhealth/design

Harris, Kathleen Mullan. Design Features of Add Health. Chapel Hill, NC: Carolina Population Center; 2011.

Hauser RM, Featherman DL. Equality of Schooling - Trends and Prospects. Sociology of Education. 1976; 49(2):99–120.

Hauser, Robert M., Warren, John Robert, Huang, Min Hsiung, Carter, Wendy Y. Occupational Status, Education, and Social Mobility in the Meritocracy. In: Arrow, Kenneth JosephBowles, Samuel, Durlauf, Steven N., editors. Meritocracy and Economic Inequality. Princeton, NJ: Princeton University Press; 2000. p. 179-229.

Hout M. More Universalism, Less Structural Mobility - the American Occupational Structure in the 1980s. American Journal of Sociology. 1988; 93(6):1358–400.

Kalmijn M. Mothers Occupational-Status and Childrens Schooling. American Sociological Review. 1994; 59(2):257–75.

Mare, Robert D. A Multigenerational View of Inequality. Demography. 2011; 48(1):1–23. [PubMed: 21271318]

Mare, Robert D., Song, Xi. Social Mobility in Multiple Generations. Unpublished Manuscript, Department of Sociology, UCLA. 2011

Mazumder B. Fortunate sons: New estimates of intergenerational mobility in the United States using Social Security earnings data. Review of Economics and Statistics. 2005; 87(2):235–55.

National Research Council. Panel on Assessing the Benefits of the American Community Survey for the NSF Division of Science Resources Statistics, Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press; 2008. Using the American Community Survey for the National Science Foundation's Science and Engineering Workforce Statistics Programs.

National Research Council. Reengineering the survey of income and program participation Committee on National Statistics. Washington, D.C: National Academy Press; 2009.

Olson K, Witt L. Are we keeping the people who used to stay? Changes in correlates of panel survey attrition over time. Social Science Research. 2011; 40(4):1037–50.

Prescott CA, Achorn DL, Kaiser A, Mitchell L, McArdle JJ, Lapham SJ. The Project TALENT Twin and Sibling Study. Twin Research and Human Genetics. 2013; 16(1):437–48. [PubMed: 23101474]

Smeeding, Timothy M., Jäntti, Markus, Erikson, Robert. Persistence, privilege, and parenting : the comparative study of intergenerational mobility. New York: Russell Sage Foundation; 2011.

Smith, Tom W., Marsden, Peter V., Hout, Michael, Kim, Jibum. General Social Surveys, 1972-2012: Cumulative Codebook. Chicago: National Opinion Research Center; 2013.

Solon G. Intergenerational Income Mobility in the United-States. American Economic Review. 1992; 82(3):393–408.

Torche F. Is a College Degree Still the Great Equalizer? Intergenerational Mobility across Levels of Schooling in the United States. American Journal of Sociology. 2011; 117(3):763–807.

U.S Department of Education. High School and Beyond, 1980: A Longitudinal Survey of Students in the United States; Codebook Volume II: Parent, Language, Twin and Sibling, Teacher Comment, and Friends Files. Washington, D.C: U.S Department of Education, National Center for Educational Statistics; 2001.

U.S. Census Bureau. Technical Paper 66. Washington, D.C: U.S Government Printing Office; 2006a. Current Popultion Survey: Design and Methodology.

U.S. Census Bureau. Technical Paper 67. Washington, D.C: U.S Government Printing Office; 2006b. Design and Methodology: American Community Survey.

U.S. Census Bureau. A Compass for Understanding and Using American Community Survey Data: What PUMS Data Users Need to Know. Washington, D.C: U.S Government Printing Office; 2009.

U.S. Census Bureau. American Time Use Survey User's Guide: Understanding ATUS 2003 to 2011. Washington, D.C: U.S Government Printing Office; 2013.

U.S. Department of Commerce. Survey of Income and Program Participation Users' Guide: Supplement to the Technical Documentation. Third. Rockville, MD: Westat; 2009.

Wakefield S, Uggen C. Incarceration and Stratification. Annual Review of Sociology. 2010; 36 36:387–406.

Warren JR, Sheridan JT, Hauser RM. Choosing a measure of occupational standing - How useful are composite measures in analyses of gender inequality in occupational attainment? Sociological Methods & Research. 1998; 27(1):3–76.

Wolfinger NH. Trends in the intergenerational transmission of divorce. Demography. 1999; 36(3):415–20. [PubMed: 10472505]

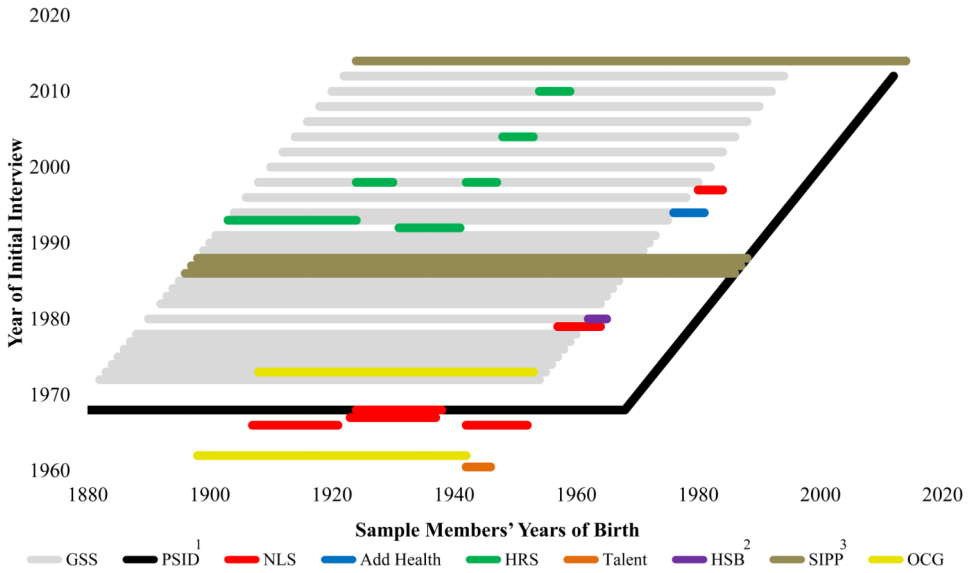

Figure 1. Major U.S. Surveys Useful for Intergenerational Mobility Research: Years of Sample Members' Birth & Years of Initial Interview

Note: To be included, the survey had to include information about the socioeconomic circumstances of at least two generations observed mid-career. For surveys with no maximum age at first interview, this figure assumes (for the sake of simplicity) a maximum age of 90.

[1]Children of PSID sample members are here counted as themselves being in the PSID upon their birth or adoption.

[2]Other NCES panels—e.g., NLS-72—are not included because the panels are inactive and/or respondents are not (yet) observed mid-career

[3]This assumes that the planned 2014 SIPP panel will include content on parents' and/or children's socioeconomic circumstances