Future Research in the Social, Behavioral, and Economic Sciences with the Panel Study of Income Dynamics

15 October 2010

Charles Brown, University of Michigan
Dan Brown, University of Michigan
Dalton Conley, New York University
Vicki Freedman, University of Michigan
Kate McGonagle, University of Michigan
Fabian Pfeffer, University of Michigan
Narayan Sastry, University of Michigan
Robert Schoeni, University of Michigan
Frank Stafford, University of Michigan

Abstract: There are extraordinary opportunities to address the next generation of research challenges in the social, behavioral, and economic sciences that build on the Panel Study of Income Dynamics (PSID). First, PSID offers untapped opportunities to examine questions of relevance to our understanding of environmental sustainability. Second, cross-national harmonization of PSID with other national panel surveys will be instrumental for developing and facilitating new research on the effects of policies and institutions. Third, measuring genetic information in PSID will open a wide range of new studies on social and economic behavior and outcomes. Advances in these areas will provide a foundation for future research and for new interdisciplinary collaborations.

Abstract word count (200 word limit): 108 words

Main Word Count (2000 word limit): 1950 words

Introduction

There are extraordinary opportunities to address the next generation of research challenges in the social, behavioral, and economic sciences. We describe three such opportunities relating to the Panel Study of Income Dynamics (PSID) that focus on human-environment interactions, cross-national research, and genetics. PSID is the longest-running nationally representative panel survey in the world and is an important component of the National Science Foundation's investment in research infrastructure for the social, behavioral, and economic sciences.

PSID and Human-Environment Interactions

First, PSID offers untapped opportunities to examine questions of relevance to our understanding of environmental sustainability. Sustainability science has emerged as an important paradigm for investigating the bi-directional linkages between human actions and natural-environmental processes with the goal of helping to solve a variety of environmental problems. The existing NSF Program on the Dynamics of Coupled Natural and Human Systems aims to investigate these problems, with a particular emphasis on modeling approaches. As this program and others like it matured, research results have highlighted the needs for data on longitudinal, economic, geospatial, cultural, and behavioral dimensions of human activity that can complement process-level understanding and data in the natural sciences. Although some new social science data collection programs are emerging as part of existing or planned environmental observatories (like NEON and WATERS; Braden et al. 2009), social science data collection efforts are still not implemented at the scale envisioned for these systems. PSID offers a unique opportunity to demonstrate the value of individual-level longitudinal information in the investigation of the human role and response in environmental systems.

We identify two major areas of environmental research that might benefit from engagement with and expansion of PSID. We believe that pursuing this research will generate important new understandings of the role and response of humans in environmental processes, but also demonstrate the value to the scientific community and to society of further investment in significant empirical social scientific research.

One important area that can benefit from existing and expanded PSID data is in our understanding of the social and economic determinants of energy, water, and materials consumption. Although not measured directly in PSID, consumption in these areas is strongly related to existing PSID data (including information on travel, housing, utilities, and location). Some additional information, through new questions or a new module, would allow investigation of the direct consumption variables and their association with these more indirect measures. Furthermore, existing information on various expenses, contributions to charities, income levels, education, and so forth can be used to investigate the social, economic, and cultural determinants of consumption behaviors. Influences of intergenerational processes on consumption behaviors might also be investigated. An important outcome of these investigations might be a better understanding of the degree to which changes in the economic, information, or natural environments might be most likely to yield changes in behavior.

Another important area for investigation and application of the PSID is the locational characteristics of participants' places of residence and places of work and travel. These studies would likely be most profitable in combination with environmental data on land use and cover, linked through the geocode. These studies could be investigated at multiple scales, focusing on demographic processes of migration and urban-rural movements, on neighborhood social and physical characteristics to understand residential preferences across a number of land markets, and at lot scales to investigate characteristics of residential land consumption and management. Detailed information on the housing and land related expenditures, house and land value, and movement can be used to better understand the residential choices of participants, how those choices are influenced by contextual factors at multiple scales, and, ultimately, how those choices influence the development of urban forms and structures we observe.

PSID and Cross-National Research

Second, in today's increasingly interconnected world, many industrialized countries face common challenges, such as population aging, the integration of immigrants, and social and economic inequalities. Although many of these problems affect modern nations similarly, their severity and impact on individual lives can differ markedly across countries. High quality, nationally representative data that cover a wide array of topics from different spheres of life have already served to illuminate some of the common challenges as well as their differential consequences. Many of the most pressing concerns, however, are often only captured in a dynamic perspective. For instance, the complexity of educational careers and life-long learning in the knowledge society and transitions into and out of unemployment and poverty in times of economic downturn require this dynamic perspective and consequently rely on longitudinal data.

Beyond description, future research must strive to identify policy solutions that have proven successful in other nations. There is ample supply of novel policy approaches and alternative institutional arrangements around the globe. In other words, in a globalized world *nations* may serve as the new laboratories of social and economic policy. However, studying the causal role of specific aspects of different policies and institutions based on cross-national comparative research suffers an important inherent problem: there are many more explanations for cross-national differences than there are countries to compare.

There are several strategies for investigating the role of institutional characteristics and policies in explaining observed cross-national differences. One promising strategy begins by reliably establishing the individual-level mechanisms that account for the observed phenomena in each nation. Understanding *why* given social or economic phenomena occur increases our chances for understanding *how* a given policy or institutional arrangement may affect these phenomena. For example, detecting barriers to educational access among disadvantaged children yields important information for inferring why different forms of educational financing do or do not impact educational opportunities. In short, pinning down the causal mechanisms that are at work at the individual-level is integral to our effort to make meaningful cross-national comparisons that have the potential to identify a "best practice" policy or institutional arrangement that may be transferable to a different nation.

Large-scale longitudinal surveys provide a strong foundation for the study of the causal mechanisms underlying a wide range of social and economic dynamics. The most important data requirement for future research, however, is that of cross-national comparability, both in terms of sample construction as well as measurement. The social sciences have profited immensely from existing large-scale projects that provide comparable data for a number of nations. Different organizational models have been successful. Some international collaborations are dedicated to the ex-post harmonization of existing surveys (such as the *Luxembourg Income Study*) while others have accomplished ex-ante standardization of a set of core questions that are asked in a large number of countries (for instance, the *World Values Survey*). Another successful model, and one that may be predicted to gain in importance in the future, is that of harmonization by imitation, in which important data collection efforts in one country inspire and guide similar projects in other countries. For instance, PSID has served as model for the *German Socio-Economic Panel*, the *British Household Panel Study*, and many other national panel studies.

Because these data sources provide the most potent basis for cross-national research, the main future challenge will be to further increase the harmonization of measures between these datasets. So far, ex-ante harmonization has mostly involved informal cooperation among the founding survey administrators while ex-post harmonization is beginning to take place in more formal initiatives, such as the *Cross-national Comparative Equivalent File* project. Although the latter efforts should be expanded, continued opportunities for ex-ante harmonization should be pursued wherever possible, for instance in the case of new panel surveys or new topical modules in existing surveys. NSF support will be instrumental for developing and facilitating cross-national data harmonization. This will improve the foundation for fruitful cross-national comparative research by providing high-quality, partially harmonized, nationally representative, longitudinal data that not only allow a dynamic view on important social and economic phenomena but that also facilitate the search for best practices that have proven successful in other countries and that hold promise in being applied to the U.S.

PSID and Genetics Research

Third, from the days of Francis Galton's eugenic theories of the heritability of intelligence and criminality through the controversial, bestselling book *The Bell Curve* (1994), introducing genetics to discussions of social behavior in humans has been morally suspect. This has led to an intellectual firewall between mainstream social science and biological data. Although recently there has been increased interest in collecting biomarkers, in general, and genetic data, in particular, in social science surveys, no existing social science study collecting genetic information is intergenerational in nature, nor are these nationally representative samples of the entire adult population across the age spectrum. Further, the focus to date of such studies has generally been on health dynamics. Meanwhile, those U.S. studies that focus primarily on collecting social, demographic, and economic data have not yet embraced the integration of genetic information.

The time is right for a nationally representative socioeconomic study to collect genetic markers. And there is no study better positioned to maximize the intellectual return on investment in this area than the PSID. Only the PSID provides what would be a "full service" socioeconomic dataset with gene markers. For example, no present study, aside from the PSID, provides an

opportunity to obtain genetic information across three generations. If the PSID were to collect genetic markers, a number of important research questions will be able to be answered, such as:

- How does the distribution of haplotypes (unique sets of polymorphic markers in an individual) vary by race, class, and region in the U.S.? Has accelerated immigration since the 1960s affected this distribution? How much genetic in-breeding occurs in the U.S.?

- How do the phenotypes and genotypes of our family and household environments affect individual outcomes? Does, for example, the expression of genetically-based propensities toward depression depend on growing up with a depressed parent? Does the effect of an individual's genetic background on social and economic outcomes depend not only on the observed behavior of family members but also on their (unexpressed) genetic makeup? For example, it could be adaptive to have a putatively more emotionally "reactive" allele when one is the only offspring to be homozygous for this allele, thereby garnering more parental attention. If the behavioral phenotype of an individual is not just contingent on her/his own genotype but that of her/his siblings, then it suggests non-independence of the units of analysis for classic heritability analysis.

- How do genes interact with exogenous economic shocks? The basic logic until now has been the following: a certain proportion of a population sample is found to have a variant of a particular allele. If this allele is shown to be randomly distributed across demographic subgroups (or, for example, within a particular subgroup such as ethnic group), and, likewise, it is found to be associated with a specific social outcome or tendency (such as addictiveness, shyness, or schizophrenia) *within* that same population (or subgroup), then researchers often look for specific outcomes which covary with the presence or absence of that particular allele. This has been the approach of most work to date in both the social and biological sciences that have used observational data. However, a problem is that alleles are not necessarily distributed randomly across sub-populations thus potentially biasing the observed phenotypic associations with those alleles. PSID would allow for within-family (cross-sibling or cross-cousin) and across-time (within-person) analysis what would alleviate some of these population stratification concerns.

In closing, the three research opportunities described above will provide a foundation for the next generation of social, behavioral, and economic research on human-environment interactions, cross-national research, and genetics. They will be a national resource for conducting transformative research that will also strengthen links between the social, behavioral, and economic sciences, on one hand, and the environmental, development and learning, and genetics sciences on the other hand.

References

Braden, J.B., Brown, D.G., Dozier, J., Gober, P., Hughes, S.M., Maidment, D.R., Schneider, S.L., Schultz, P.W., Shortle, J.S., Swallow, S.K., and Werner, C.M. 2009. Social science in a water observing system. *Water Resources Research*, 45: W11301.

Herrnstein, R., and Murray, C., 1994. *The Bell Curve*. New York: Free Press.