Actes de conférences
Enquêtes longitudinales sociales et de santé
dans une perspective internationale

Conference proceedings
Longitudinal Social and Health Surveys in
an International Perspective

The Panel Study of Income Dynamics:
Overview & Summary of Scientific Contributions After Nearly 40 Years

Katherine A. McGonagle & Robert F. Schoeni

The Panel Study of Income Dynamics was born as a result of the desire to assess the success of President Lyndon Johnson's War on Poverty on the economic well-being of U.S. families. In 1966 and 1967, the Office of Economic Opportunity (OEO) directed the U.S. Bureau of the Census to conduct a national assessment of the success of the War on Poverty by conducting a study called the Survey of Economic Opportunity (SEO), which completed interviews with about 30,000 households. Interest in continuing this national study led OEO to approach the Survey Research Center (SRC) at the University of Michigan about interviewing a sub-sample of approximately 2,000 low-income SEO households. Professor James N. Morgan, who became the new study's director at SRC, argued successfully for adding a fresh cross-section of households from the SRC national sampling frame so that the new study would be representative of the entire population of the United States, including non-poor as well as poor households. In addition, it was fortuitously decided to follow members of the families who moved away from their original households, such as children who came of age during the study. In this way, the sample could remain representative of the nation's families and individuals over time. This study became what is now called the Panel Study of Income Dynamics (Hill, 1992).

Originally conceived as a five-year project, the PSID will soon mark its 40th anniversary. Nearly 70,000 people have participated in the PSID, 4,000 sample members have died, and as many four generations within given families are represented. The mission of the PSID has broadened well beyond its original more narrow focus on income and poverty dynamics. As a result, the PSID has been used in over 2,000 peer-reviewed publications, and the user base has grown increasingly diverse, drawing in psychologists, medical researchers, public health scholars, geographers, and others. In this study we provide an overview of the PSID.

We were asked to describe the most significant scientific findings based on the PSID, and to discuss the strengths and weaknesses of the survey. Before we address these questions, we begin by describing the key features of the design and content as well as data collection and distribution technologies. Section II summarizes the key scientific advances made possible by the PSID. Sections III and IV explicitly address the strengths and weaknesses of the PSID, respectively.

1. Overview of the Panel Study of Income Dynamics

1.1. Survey Design

The PSID was originally designed to study the dynamics of income and poverty. Thus, the original 1968 PSID sample was drawn from two independent samples: an over-sample of roughly 2,000 low income families from the Survey of Economic Opportunity (the "SEO sample") and a nationally representative sample of roughly 3,000 households designed by the Survey Research Center at the University of Michigan (the "SRC sample"). The oversampling of families who were poor in the late 1960s resulted in a sizable subsample of African Americans. Probability–of-selection weights enable analysts to make estimates from the sample that are representative of the U.S. population. From 1968 to 1996, the PSID interviewed and re-interviewed individuals from families in these two samples every year, whether or not they were living in the same dwelling or with the same people. Adults have been followed through the full life course and there is now a confidentiality-protected mortality data file that includes over 4,000 PSID sample members who have died since 1968. Children have been interviewed as their own family unit after they have left their parents' household. This unique self-replacing design implies that for many PSID families, the data include self-reported information on three (and occasionally now, four) generations within the same family at various points in their life course. Through multiple waves collected over long time periods on a national population, the PSID is the only data set ever collected on life course and multigenerational economic conditions, well-being, and health in a long-term panel representative of the full U.S. population. Comparisons of PSID data with simple cross-sectional benchmark studies – the March Current Population Survey for income (Gouskova and Schoeni, 2002a), Survey of Consumer Finances for wealth, National Health Interview Survey for health status and health behaviors (Gouskova and Schoeni, 2002b), and the Consumer Expenditures Survey for expenditures (Charles et al., 2004) -- support the claim that the sample, with weights, remains representative of the U.S. population for a given temporal snapshot.

In 1990, the PSID added 2,000 Latino households, including families originally from Mexico, Puerto Rico, and Cuba. But while this sample did represent three major groups of immigrants, it missed out on the full range of post-1968 immigrants, Asians in particular. Because of this crucial shortcoming, and a lack of sufficient funding, the Latino sample was dropped after 1995. In 1997, in order

to keep the study representative of the U.S. population, two major changes were made to the PSID sample: 1) a reduction of the core sample and 2) the introduction of a refresher sample of post 1968 immigrant families and their adult children. The original core sample was reduced from nearly 8,500 families in 1996 to approximately 6,168 in 1997. The majority of the cuts were taken from the SEO sample. However, 43% of the SEO sample, or 1,714 families, remained in the active sample. Through natural sample growth generated by split-offs, the SEO sample now includes 2,279 families as of 2005. Because the SEO sample consists largely of low-income African-American families, it supports research on economic transitions, poverty, and disparities in health and other resources. A sample of 511 immigrant families was added in 1997/1999 to update the PSID sample of U.S. households by adding a representative sample of recent immigrants to the United States. Immigrants eligible for the supplemental sample were those who have immigrated to the United States after 1968 and who were not married to persons who were living in the United States at the time of the original PSID sample selection in 1968. These recent immigrants were not represented in the PSID sample as of 1997. By 2005, the immigrant sample re-interview response rate had increased to more than 94%, approaching the 98% re-interview response rate of the original sample. Moreover, because some of the 511 immigrant families created split-off families, the number of post-1968 immigrant families has increased to 572. In total, as of 2005 there are 8,041 PSID families.

The steady-state panel design, which is ideal for studying the dynamics of human and social behavior, is depicted in Figure 1. Flows of people into the panel come from three sources: the original 1968 sample (N); the 1997 refresher sample of post-1968 immigrants (n); and births and marriages in existing families (d). The intergenerational element is represented by children who split off as adults to form their own households (s). Because of the unique follow-status rules (f_1, f_2), and because of success in bringing in new families (i.e., boosting f_1, f_2) and minimizing attrition (a_1, a_2), unlike studies worldwide that have adopted the PSID steady-state design in Figure 1, the PSID sample has grown despite losses due to mortality ($m1$, $m2$, $m3$, $m4$) and attrition from the active panel (P). Also, by attempting to re-contact (r_1) families refusing to be interviewed or not located ($A1$ and $A2$) in previous waves, the active panel can be boosted. Decisions about the efforts used to shape the various flows (f_1, f_2 a_1, a_2, and n) allow the long-run sample composition to be 'steered' to some extent.

1.2. Consistently High Response Rates
The PSID has achieved high response rates. Table 1 displays the 2003-2005 response rates for different segments of the PSID sample. A response rate of 97% has been consistently obtained for the "core re-interview sample." The "core sample" consists of all PSID families except the 1997 immigrant refresher

families. The "re-interview sample" includes families interviewed in the previous wave. The "core re-interview sample" accounts for roughly 85% of the entire PSID sample in a given year. The "re-contact sample" consists of families who did not respond in the previous wave, but were respondents in the wave prior to that; this sample is labeled the "attrition stock" in Figure 1. PSID attempts to re-contact these non-response families as a way to minimize long-run attrition and maintain the representativeness of the sample.

A high annual continuation rate is critical to the success of a long panel study. While most cross-sectional surveys have experienced rapid declines in response rates over the past several years, the PSID has experienced no such decline. In fact, the response rates in the most recent waves have been as high as any in the entire 37-year history of the PSID. These high rates are most likely due to a number of factors:

- a substantial incentive payment of $60 in 2005 that is mailed within a few days of the interview;
- paying respondents $10 during years PSID is not collecting data if respondents provide PSID with their current contact information;
- maintaining a set of experienced interviewers from wave-to-wave, many of whom have interviewed the exact same families for several years;
- a staff person experienced in refusal conversion who writes persuasion letters tailored to reluctant respondents;
- keeping respondents informed about the content and timing of upcoming data collections;
- sending families an annual respondent newsletter describing the scientific value of the study and thanking them for their continued participation;
- having a specialized corps of interviewers specifically trained in refusal conversion to encourage participation of reluctant respondents.

1.3. The Child Development Supplement
Children have always played an important role in the PSID. However, only limited data were collected about children until the first wave of the PSID-Child Development Supplement (CDS-I) was fielded in 1997. The CDS-I collected information on up to two randomly selected 0-12-year-old children residing in PSID families, resulting in a total sample of 3,563 children in roughly 2,400 families. The objective was to provide researchers with a comprehensive, nationally representative (with child-based weights), longitudinal database of young children and their families to study the dynamic process of early human capital formation. The age group was selected to fill an important gap in our knowledge about the development of young children.
In 1997, the following measures were collected: (1) cognitive, behavioral, and health status; (2) assessment of basic skills in reading, mathematics, and written language; (3) a comprehensive accounting of parental and caregiver parenting

styles and time inputs; (4) two 24 hour time diaries for each child – one for a random weekday and another for a random weekend day; (5) other-than-time-use measures of other resources, for example, the learning environment in the home; teacher and administrator reports of school resources and decennial-census-based measurement of neighborhood resources; and (6) child-reported indicators of socio-emotional characteristics.

The second wave of the CDS was successfully fielded in 2002/2003 with a response rate among primary caregivers of 91 percent. The same children originally interviewed in PSID-CDS-I, now aged 5-18, were re-interviewed. Much of the same information was collected from parents, teachers, and the children, as well as new data addressing issues relevant to adolescent transitions. In sum, the PSID-CDS provide unique nationally representative panel data in several domains: (1) high-quality annual measurements of economic and demographic conditions at both the family and neighborhood level, dating from birth, (2) teacher and school characteristics at the time of the interview; (3) information from multiple caregivers, including fathers, whether present in the household or in another household; (4) high-quality measures of child cognitive, social, and behavioral development; (5) rich information on family processes, (6) unique comprehensive national time diary measures on children of all ages in several settings which has detailed coding supporting the comparisons with the single day diary cross-sectional data from the 2003 BLS American Time Use Survey (ATUS) – which collected diaries down to age 15; and (7) large and representative samples of siblings to examine within-family disparities.

1.4. Survey Content: Current Domains
Since its inception in 1968, the PSID has collected extensive information on income and family demographics. With input from the PSID Board of Overseers and the broader scientific community, new material has been continually added to the study's content domains to allow the study of emerging scientific and policy interests, particularly during the current NSF funding cycle (2002-2006). As can be seen in Table 2, PSID collects data on a wide array of social, demographic, economic, geospatial and psychological topics, supporting unique multidisciplinary research. In 2005, the 74 minute interview collected data on: employment, earnings, income from all sources, expenditures covering 72% of total household spending (Charles et al., 2004), housing, geospatial data, health status, health behaviors, and health insurance, marriage and fertility, participation in government programs, vehicle ownership, wealth and pensions, and philanthropy. Many of these areas have been included in the instrument since 1968. Hundreds of additional variables that fall into other domains have been collected in various waves throughout the history of the PSID. These variables can be easily identified in our on-line Data Center or website using various tools and added to a user's customized data extract.

For many domains, the interests of mission agencies, including the National Institute on Aging (NIA), the Assistant Secretary of Planning and Evaluation (ASPE) at HHS, the Economic Research Service at the Department of Agriculture, (USDA/ERS), and the Department of Housing and Urban Development (HUD), coincided closely with the interests of the larger social science research community. For instance, an extensive battery of questions on health, wealth, and pensions has been supported through investigator-initiated awards made by NIA, and this information increases substantially the value of the core data supported by NSF. Questions about food security and housing have been funded by USDA and HUD, respectively. And in 2001, 2003, and 2005 a philanthropy module was included with support from the Center on Philanthropic Study at Indiana University Purdue University at Indianapolis. New scientific domains will constantly be assessed to keep the PSID up-to-date with the needs of the research community. Potential supplemental data collection in four new domains are currently being considered: time use, biomarker assessment, education and information technology, and transition to adulthood.

1.5. Data Collection Processes

Since 1993, the PSID has been collected via a computer-assisted telephone interview (CATI), and in 2003 Blaise software was used to program the questions, along with SurveyTrak, an internally developed software, to manage sample and administrative information about the family. The PSID instrument was converted to Windows-based Blaise (Statistics Netherlands) in 2003 after a comprehensive evaluation of software technologies concluded it was the most powerful computer-assisted interviewing tool and survey processing software that could reliably handle the complexity of surveys such as PSID, which are lengthy and depend on many routing conditions, calculations, and edits.

Other data collection systems that have been developed internally for the PSID include web-based systems (i.e., 'Webtrak' and 'Casetrak') used to communicate between interviewers and project staff on issues such as refusal conversion, requests for persuasion letters, and tracking of respondents. An Oracle-based respondent address payment system (RAPS) was developed by PSID staff for the 2005 field period that will be used for many waves of data collection into the future. RAPS has several functionalities, including: the archiving of multiple respondent addresses, including the address at which they wish to receive their incentive, a secondary contact address, and the address at which the interview occurred; the automated generation of tailored letters for respondents, including different refusal aversion/conversion templates based on the type of reluctance the respondent is exhibiting (e.g., Groves & McGonagle, 2001), as well as sympathy letters, and advance contact letters; and the automated generation of respondent incentives immediately after the interview is completed, with most respondents receiving their incentive payment within 48 hours.

A major innovation introduced in the 2003 wave was the Event History Calendar (EHC), which provides 2-year long timelines of employment, residence, and features of employment across job transitions. The layout of the EHC from the 2003 CATI application is presented in Figure 2. Having 2-year data in these content areas has helped fill the gap of data caused by moving the study to a biennial data collection. The fine-grained EHC timeline data can be used to support the construction of traditional measures – such as weeks of employment, unemployment, and time out of the labor force. Methodological research has shown that the EHC interviewing methodology leads to consistently higher quality retrospective reports in comparison to traditional standardized question-asking methods (Belli et al.,, 2001; Belli et al, 2004). In addition, these timeline data can be used to analyze interrelated events such as the timing of auto purchases, residential moves, and employment transitions.

1.6. Data Distribution
The long period over which data have been collected, the extensive range of measures captured in the instrument, the unique sample following rules, and the fact that both individuals and families are followed make the PSID a valuable data source. At the same time, these qualities combine to make the PSID increasingly complex for users. To overcome what might otherwise be a major barrier, limiting the use and scientific advances from the archive, the PSID has developed tools to allow users to take full advantage of the many aspects of the unique data. In this section we discuss four specific ways in which PSID has made progress on this front: Internet-based data distribution; an on-line cross-year variable index, a family identification map application, and sensitive data available under contract.

1.6.1. Internet-based Data Center.
The PSID began distributing data through an on-line data center in 1996. An important milestone was the successful implementation in the fall of 2003 of a major transition to an Oracle-based Internet Data Center. Over the past three decades, thirty-four waves of data comprised of over 60,000 variables based on 8,000 families and nearly 70,000 individuals have been processed and disseminated via the Internet. We have used the best available technologies and implemented continual software improvements. The PSID's software infrastructure, although customized to meet the study's specific needs, is designed with an architecture that is highly generalizable and shares essential features with many other long-lived scientific data collections.

The recent redesign of the Data Center using Oracle allows users to more easily create customized longitudinal datasets from all waves of the PSID by choosing various options, creating customized codebooks specific to the data that has been downloaded, searching and browsing for variables, and archiving data downloads for shared and future use. Users can now "load up" data carts with

selected variables from selected observational years. When users are ready to download their data, or "check out," they first view the contents of their data cart. Users also see an "open-book" icon next to each variable, and pressing on this icon takes them to a window that contains the full codebook documentation for that variable. Users can then choose to edit their cart by removing any unwanted variables from their cart simply by checking the box in the column second from the right, or they can add variables by returning to the "data aisle" for more items. They may also choose to completely empty their data cart, or to proceed to download their data, i.e., "check-out."

Data carts may also be saved and named, allowing users to easily track specific data downloads. Users can choose from a range of output types including SAS, SPSS, STATA, dBase, and ASCII. Moreover, users can specify the data subset in a selection phase and work with smaller exploratory files or a more extensive dataset, as they so choose.

The greatly improved functionality of the Data Center supports dataset creation that is customized by the user in terms of both variables and sample connections. This approach is a streamlined, efficient way to create customized data subsets from all waves of the PSID, including selected supplemental data, CDS-I, and CDS-II. Customized subsetting allows users to create basic intergenerational analytical files by automatically linking data about PSID families, PSID individuals including children, caregivers, and teachers from CDS-I and CDS-II. The Data Center also allows the capability to generate customized merges of the family and individual data across multiple waves.

1.6.2. Online Cross-year Variable Index.
With thirty-four waves of data now available, as well as data from two major supplements (i.e., CDS-I and CDS-II), an index that provides information about all categories of variables across the full archive of the panel is critical for allowing users to find and organize the data to be analyzed. Work to create such an index has been ongoing throughout calendar year 2005, and a version is currently being tested. We anticipate releasing a version early in 2006. Users will see a given domain of variables – income, health, or wealth, for example – and then be able to "unfold" this category to see all of the variables related to that domain. For a given specific variable -- for example, current employment status -- the index will list the years that the variable is available. The user can then click on the year to view the codebook for that specific year. An option will be added that will also allow the user to click on the year indicator for the given variable to add that variable for that year to their data cart.   In sum, the cross-year index integrated with the Data Center will allow users the option of "browsing" the entire PSID archive sorted by variable domains.

1.6.3. Family Identification Map File.
Because of its genealogical design, the PSID is one of the few nationally representative US datasets that can be used for intra- and intergenerational analyses. As described above, from its beginning the PSID has followed all 1968 family members and their descendants. When family members split-off and create their own separate family, the PSID interviews these new families as well as the original family. The numbers of sibling pairs, child-parent pairs, and grandparent-grandchild pairs are substantial. For example, tens of thousands of sibling pairs have been recorded in the PSID as of 2005. Of course, not all siblings are alive or reporting data in each and every wave of the PSID, therefore the number of siblings available for any particular analysis will be smaller. But even in the one wave 2003, there were thousands of sibling pairs who were heads or wives, which means that the full set of PSID data is collected on these siblings. These large samples will support a wide range of analyses for fairly narrowly defined samples, but the creation of these files is complex, and can be prohibitively so for some users.

With the goal of facilitating the use of these data to support complex models of family and life course development, in calendar year 2005 the PSID has created a Family Identification Map (FIM) application. FIM creates a customized file – i.e., "map file" - that contains the identification variables of the relatives an analyst wishes to examine. FIM also supplies SAS code that uses the map file and a file containing the variables the analysts wishes to examine to create a new data file that includes the variables for the relatives of interest. For example, an analyst could choose "biological grandparents" as the relative of interest. FIM would then generate a file that contains the IDs of all PSID sample members and each of their four biological grandparents. The user would then create an individual level data file that contains all of the variables they want for either the grandchild or the grandparent. PSID then provides SAS code that will merge these two files and create a data file in the shape desired by the analyst. The shapes available are "wide" – i.e., one observation per grandchild – and "long" – i.e., one observation per grandchild-grandparent pair. A beta version of FIM has been tested by several individuals in the research community, and we anticipate that FIM will be released for general use by the end of 2005.

1.6.4. Sensitive Data.
PSID first provided geospatial data under sensitive data contract on an exploratory basis in 1990. In addition to continuing to provide geospatial data, PSID now provides information on date, place, and age of death for known decedents, and will soon provide cause of death information. In addition, links to data about the specific school district and school are available from the National Center for Education Statistics Common Core of Data for use with the subsample of the PSID that participated in the Child Development Supplement.

Since 1990, demand for the sensitive data has increased, with 80 contracts currently active out of a total of 153 ever provided to users. And in just the past five years (2000-2005) we have approved 50 percent more contracts than in the entire decade of the 1990's. There are now more than 200 publications based on these data. As one example, PSID sensitive data can be used to construct geospatial life histories and examine psychosocial and physical exposures in much the same way that prior research has examined such exposures using occupational histories (Amick et al., 2002). For example, these data could assess cumulative risk to a variety of outcomes as a function of number of years residing near known sources of pollution such as high smog areas, airports, steel mills, coal-fired power plants, lead exposures, etc.

In order to maintain respondent confidentiality, these data are provided only under a contract between the University of Michigan and the institution of a qualified researcher. To obtain these sensitive data, a user must provide a CV, description of planned work which supports the appropriate use of the data, a data protection plan which meets a variety of security requirements, approval of human subjects review board, and an administrative processing fee of $750. A Ph.D. level staff person reviews all submissions weekly to ensure appropriate compliance. Applicants must complete their work within 3 years at which time they must return or provide proof that the original data file was destroyed, or renew their contract with new application materials.

Of all the sensitive data, the geospatial data have been available for the longest time period and are the basis of the vast majority of sensitive data contracts. The geocode match files include the identification codes necessary to link data from the PSID in each year to contextual data from secondary sources such as the census. We have worked to provide a set of accurate location indicators and to develop a system for geocoding new PSID addresses in the future. The geocoding process has been conducted using four different Censuses. Addresses from 1968-1985 were geocoded to 1970 and 1980 census identifiers; those from 1968-2001 were geocoded to the 1990 census identifiers; and the 2003 address information was geocoded to the 2000 census identifiers. The primary objective of this geocoding project has been to characterize the "neighborhoods" in which respondents live. Census tract is used as the approximation of neighborhood in tracted areas, Block Numbering Areas (BNAs) in blocked but untracted areas (i.e., small cities), and Enumeration Districts (EDs) in usually rural areas with neither tracts nor blocks.

We anticipate an increase in contracts to use the death file data. Two projects to increase the amount and quality of data available to researchers are currently underway. First, with support from NIA, we are now working toward providing a file from the National Center for Health Statistics National Death Index that includes cause of death information for known decedents. We plan to update

this information for every wave of data we collect. Second, a longer term project is now underway to determine the status of all approximately 10,000 individuals who have attrited from the PSID since 1969. We are searching for them first in the publicly available Social Security Death Index, and if we do not locate them there, they will be formally tracked by the Survey Research Center's operational branch that conducts respondent tracking. Data from the individuals that become known decedents will be sent to NCHS's National Death Index so that cause of death information can be added to their records and provided under sensitive data contract.

In the Spring of 2005, information obtained from the U.S. Department of Education National Center for Education Statistics (NCES) Common Core of Data (CCD) was linked to the CDS sample and made available to users under sensitive data contract. The CCD is a national database of all public elementary and secondary schools and school districts in the United States. The CCD survey annually collects data about all public elementary and secondary schools, all local education agencies, and all state education agencies throughout the United States. CCD contains three categories of information: general descriptive information on schools and school districts; data on students and staff; and fiscal data. Under sensitive data contract a variety of characteristics of the school and district may be linked to information about the child and used to support models of child development. These school- and district-level data may be used in combination with the geospatial data to include aspects of the neighborhood.

1.7. Funding and Administration

The PSID has been funded primarily by the National Science Foundation, as well as by a collection of federal agencies including the Office of Economic Opportunity, the Assistant Secretary for Planning and Evaluation of the Department of Health and Human Services; the Departments of Labor and Agriculture; the National Institute of Child Health and Human Development, and more recently, substantial funding has been received by the National Institute on Aging. Since 1982, the study has had an advisory Board of Overseers, created by NSF to foster input from the national community of scholars, researchers, and policymakers.

From its beginning in 1968 until 1989, the PSID was directed at the Survey Research Center, University of Michigan by James N. Morgan, now emeritus. From 1982-1989 responsibility for running the study was also shared by Greg J. Duncan (as co-director), Daniel H. Hill, and Martha S. Hill. Since 1989 and until 1995 Greg Duncan directed the study, with Martha Hill and James N. Lepkowski as co-directors. Frank P. Stafford became director of the study in 1995, with Sandra Hofferth as co-director until 2001. Since 2001 responsibility for running the study has been additionally shared by Robert F. Schoeni (as co-PI and Associate Director).

2. Scientific Impact

In this section we report on several indicators of the usefulness of the data, including: number of published articles using the PSID, number of grants awarded by NIH and NSF to support research using the PSID, number of data-downloads from the PSID Data Center, and number of sensitive data contracts issued. After describing success of the PSID as measured by these quantitative indicators, we describe some of the specific scientific advances made possible by the PSID.

2.1. Publications Using the PSID

As of October 1, 2005, PSID staff had located 2,167 publications based on PSID-CDS data, including 1,432 published journal articles, 303 books and book chapters, and 402 dissertations. Today, on average, there is one publication using the PSID every 3.9 days. Moreover, the number of published journal articles has continued to increase (Figure 3).

Articles have appeared in the top economic and sociology journals and in journals from numerous disciplines. Over the past 10 years, the most common journal for articles using the PSID is the American Economic Review. Other leading journals in the "PSID Top 10" include the American Sociological Review, Journal of Marriage and the Family, Review of Economics and Statistics, Demography, and Journal of Political Economy. Articles have appeared in many journals from a variety of scientific disciplines, including demography, health, medicine, child development, and psychology.

The PSID's user community is very broad. At almost all major research universities in the U.S., there is at least one scholar analyzing the data. Many young scholars have begun their careers analyzing PSID data, with 402 dissertations written using the PSID. And these new users are also spread throughout the nation. For example, there were 49 different educational institutions represented among the 94 dissertations completed between 2000 and October of 2005.

2.2. Grants Awarded by NSF and NIH Using the PSID

Although the vast majority of social science research in the U.S. is conducted without grant support, a substantial share is. In this section, we report on the amount of such research supported by two of the major sponsors of research using the PSID – NIH and NSF – both of which have searchable on-line data bases containing the abstracts of funded projects.

The NIH Crisp database was used to search all awards made by NIH agencies to use or collect PSID data. The search showed that a total of 133 unique awards by NIH were made between 1972 and 2005. Note that this total is an underestimate

because the database searches abstracts only; thus an award that did not use 'PSID' or 'Panel Study of Income Dynamics' in its abstract was not counted. A search of the NSF Fastlane database shows that a total of 66 unique awards had been made for analysis or data collection of the PSID data as of October 10, 2005. Of the 66 awards, about 78% (n=52) were made for secondary data analysis and 21% (n=14) were made specifically for data collection.

2.3. Data Downloads and Web Site Activity

The PSID can be downloaded from the publicly available and free web site. Information on general website activity was monitored and collected starting in calendar year 2004. During 2004, there were 1,251,718 total hits to the PSID website from 52,613 unique users (i.e., unique IP addresses). These totals exclude users within the Institute for Social Research, including PSID staff. User activity specific to the actual downloading of data was also assessed. There are two ways to download data, both through the PSID Data Center. The first way is to create customized datasets directly from the Data Center by selecting various types and years of data and variables. The second way is to download complete data files that are compressed in zip packages. In 2004, 12,756 data downloads were made directly through the data center, and 11,830 zipped downloads were made, for a grand total of nearly 25,000 data downloads. Across both types of downloading, data sets were created by 4,874 unique users (i.e., unique IP addresses). These figures indicate that the PSID is used frequently and by a large number of distinct users.

2.4. Specific Scientific Advances Made Possible Using the PSID

While it is difficult to fully and accurately describe all of the scientific contributions made possible with the PSID, we touch on some of the major areas of discovery. A list of additional scientific questions and social theories that have been investigated with the PSID is provided in Table 3. A complete bibliography is provided on the PSID web site: http://psidonline.isr.umich.edu/Publications/Bibliography/default.aspx

2.5. Intergenerational Studies.

The PSID has supported unique contributions to intergenerational research. The long time-series provide a rare opportunity for researchers to construct family and individual experiences through the life stages, from birth to childhood to early adulthood to old age and death. In addition, the PSID data provide many sibling pairs for researchers to address the long-standing issue of unobserved heterogeneity. This body of literature helps to understand how early family events, poverty, welfare receipt, and early human capital investment affect achievements in adulthood after controlling for a wide range of observed and unobserved family and neighborhood characteristics (Haveman et al. 1997; Solon et al, 2000). The fact that both parents and their adult children are interviewed in the PSID has supported extensive analyses of the intergenerational transmission

of economic status as measured by earnings and income (e.g., Solon, 1992), wealth (Hurst and Charles, 2003), and consumption (Mulligan, 1997, and Aughinaugh, 2001). More recently the data have been used to examine the transmission of health status across generations (Davis, McGonagle, and Schoeni, 2005).

PSID data have been used for some important work on intergenerational transfers. Altonji et al. (1996; 1997) examined theories of parental helping behavior and *inter vivos* assistance. They found that money transfers tend to reduce inequality in household incomes and that time transfers are only weakly related to income differences. Among parents and in-laws, the richer set of parent is more likely to give money and less likely to receive money. Richer siblings give more to parents and receive less. In contrast to the implications of simple exchange models of transfers, there is little evidence in the cross section or in the analysis using siblings that parental income or wealth raises time transfers from children or that time transfers are exchanged for money transfers. Work along this line includes that of Jayakody (1998), Wilhelm (1996), and Couch et al. (1999). Furstenberg et al. (1995) examined the effects of divorce on intergenerational transfers of money and time and found no evidence that divorced fathers who paid child support are more likely to be involved in intergenerational transfers than those who did not pay child support. These results support a growing body of evidence that marital disruption is altering the organization of kinship in American society. When men relinquish ties to their children during childhood, they rarely resume those ties in later life. Smock and Manning (1997) found that the characteristics of nonresident parents are central to understanding levels of child support.

2.6. Health and Aging.
In the health literature, researchers have found that marriage decreases men's and women's risks of mortality, even taking positive selection into marriage into account (Lillard and Waite, 1995). Other work has examined the contribution of employment factors and personal health behaviors on mortality (e.g., Wolfe and Haveman, 1983; Haveman et al., 1994). Physical and environmental hazards lead to deterioration in men's and women's health status, net of other factors, and smoking has been found to be consistently harmful to health. Recent research in this area has focused on factors associated with childhood obesity (Gibson, 2004; Vandewater et al., 2004). Taking advantage of the rich set of income data over a long period of observations, a number of studies have examined the effects of employment characteristics, income, and income dynamics on mortality (McDonough et al., 1997; 1999; McDonough and Amick, 2001; McDonough and Berglund, 2003; Amick et al., 2002).

The PSID has been used to show that among senior Baby Boomers, there is a large group with little in the way of household wealth accumulation by 1994.

Whether this reflects their 'grasshopper' approach toward household finances or whether they were just unlucky 'ants' is not fully clear. The PSID has also been used reconcile the apparent puzzle of a low saving rate but yet high wealth (Juster et al., 2001). The PSID has been used to test the life-cycle consumption model in numerous studies. Disparities in wealth and disability across socio-demographic groups, and differences in the SES-health gradient over the life course have also been examined.

2.7. International Comparisons.

PSID data have also played an increasingly important role in international comparative research. Several recent studies have used PSID and European panel data to conduct comparative analyses of economic indicators in the US with other countries such as Germany and Belgium (Van Kerm, 2004), the Netherlands (Headey et al., 2004), and Sweden (Evertsson and Nermo, 2004). Couch and Dunn (1997) used PSID and German panel data to calculate comparable measures of intergenerational correlation of earnings, hours, and education in the two countries. A remarkable similarity exists across the two countries in the correlation of earnings and of annual work hours of fathers. There is a stronger correlation for daughters and mothers in the U.S. than in Germany, which may be due to the greater labor market integration of women in the U.S. Blau and Kahn (1996) reported a higher level of wage inequality in the U.S. than in nine other OECD countries. They found that the greater overall U.S. wage dispersion primarily reflects the substantially greater compression at the bottom of the wage distribution in the other countries. Duncan et al. (1995) examined poverty and social-assistance dynamics in North American and European countries. Bjorklund and Jantti (1997) used the PSID to study intergenerational income mobility in Sweden and U.S. Recently released cross-national equivalent files prepared by Burkhauser and colleagues will further facilitate international comparisons of findings using panel data from the US (PSID), Germany (GSEOP), Britain (BHPS), and Canada (SLID).

2.8. Demographic Trends and Behaviors.

The PSID continues to provide long-term histories of marriage, childbirth, and living arrangement data to contribute to the understanding of these demographic trends and their effect on the socioeconomic well-being of families and individuals. Haveman et al. (1997; 2001), Corcoran and Kunz (1997), and Foster et al. (1998a; 1998b) examined the costs and social consequences of teenage pregnancy and premarital childbirth. Blank (1998a; 1998b), Powell and Parcel (1997), and Hill et al. (1999; 2001) investigated the living arrangement patterns and the effects of family structure and children's early adult achievement and women's early trajectories. Brines and Joyner (1999) attempted to explain what unites cohabiting partners over time. They tested their theory and found support for the hypothesis that cohabiters are more likely to remain together under conditions of equality.

Many researchers have studied the relationship between marriage or marital dissolution and individuals' earnings and labor supply behavior (Nakosteen and Zimmer 1997; Ono, 2003; Vanderklaauw, 1996), health (Lillard and Waite, 1995; Lillard and Panis, 1996), and intra-household resource allocation (Ono, 1998). Moffitt et al. (1998) used four data sets, including the PSID, to investigate the extent and implications of cohabitation and marriage among U.S. welfare recipients. They found weak evidence in support of the idea that AFDC provided incentives for recipients to cohabit. Whittington and Alm (1997) examined the effect of income taxation on divorce and Gray (1998) examined divorce-law changes, household bargaining, and married women's labor supply. Charles and Stephens (2004) examined the relationships between job displacement, disability, and divorce.

2.9. Neighborhood Effects.
Utilizing neighborhood data linked to the rich family and individual information collected in the PSID, researchers have studied the effects of community over the life course. South (1999; 2001a; 2001b) and South and Crowder have a series of papers (1997; 1999; 2000) showing that the level of neighborhood poverty weakens the relationship between childbirth and marriage for women. Education and marriage increase the likelihood of leaving poor tracts, while age, home ownership, and receiving public assistance reduce it (Crowder and South, 2003). Residential segregation by race and poverty status and the supply of new housing in the metropolitan areas also influences the likelihood of moving between distressed and non-distressed neighborhood. Foster and McLanahan (1996) examined whether neighborhood conditions affect a young person's chance of finishing high school.

Sucoff and Upchurch (1998) found that compared with living in a racially mixed neighborhood, living in a highly segregated neighborhood is associated with a 50% increase in the rate of a premarital first birth, regardless of neighborhood socioeconomic status. Quillian (1999) found that the migration of the non-poor away from moderately poor neighborhoods has been a key process in forming new high-poverty neighborhoods. He also found that neighborhoods with increasing black populations tend to lose white population rapidly. Harris (1999) examined whether the housing prices are lower in neighborhoods with high concentrations of black residents and, if so, whether this is due to racial discrimination. He found clear evidence that property values do respond to racial composition. However, housing in neighborhoods with a high percentage of black residents is less valuable not because of an aversion to blacks per se, but rather because people prefer affluent, well-educated neighbors, and these traits are more common among whites than blacks.

A series of papers by Solon, Page, and co-authors use two unique features of the PSID – the clustered sampling in 1968 and the following of "split-offs" into

adulthood -- to re-examine the claim that neighborhood effects are large. Specifically, they estimate the correlation in educational outcomes among siblings and among neighbors and find that the former is much larger than the latter, and that neighborhood correlations in general are quite small; these findings suggest that there is little room for neighborhood effects to play a large role in educational attainment and labor market earnings. At the same time, recent work using the same approach finds that neighborhood effects on health status may be substantial (Johnson and Schoeni, 2006).

2.10. Child Development.

With the addition of the 1997 and 2002/2003 Child Development Supplements (CDS), the PSID has begun to make substantial contributions to the child development literature with current, high-quality data based on a national sample. Recent work includes examination of children's time use (e.g., Sandberg and Hofferth, 2001), fathers' involvement with children (Marsiglio et al., 2000), the connection between TV, media and obesity (Vandewater et al., 2004), the effects of welfare reform on children (e.g., Hofferth et al, 2000), and determinants of child achievement (e.g., Conley and Yeung, forthcoming; Davis-Kean, forthcoming; Hofferth and Reid, 2002; Loveless, 2003). We anticipate continued growth in this area with the recent release of the 2002/2003 CDS.

2.11. Influences of Family History.

There are several key findings in the demographic literature that could not have been obtained without PSID's longitudinal data following parents and children over a long period of time. These include the finding that parental divorce is harmful for kids (e.g., McLanahan and Sandefur, 1994), that teenage childbearing is likely to be harmful for the young mothers (e.g., Hoffman et al., 1993), and that neighborhoods have effects on children's development independent of family characteristics (Brooks-Gunn et al., 1993). Studies of intergenerational mobility have also been possible because of the longitudinal nature of the PSID.

2.12. Income and Balance Sheet Dynamics.

The PSID has contributed to a redefinition of the notion of poverty from one of a static, long-term state, to that of dynamic, cyclical phases. For example, Rank and Hirschl (2001) used data from PSID to demonstrate that nearly 2/3's of all Americans and 90% of all African-Americans will live below the poverty line at some point during their lives. It has been shown that compared with the older generation, the younger generation has experienced a slowdown in the attainment of middle-class status and financial independence from their parents (Duncan et al., 1995). For example, 42% of young men who turned 30 in 1989-1992 had attained middle-class earnings, compared with 60% of those who turned 30 in 1977-1988. This decline in upward mobility has been uniform across all demographic groups. PSID was the basis of research showing that income inequality has grown in the U.S. and upward economic mobility for the lowest

income groups has not increased 1988-1998 (Bradbury & Katz, 2002).   In addition, the PSID is now being used to address the puzzle of the low savings rate in the United States. Recent work based on the PSID wealth data indicates a high rate of wealth mobility (Hurst et al. 1998; Hurst and Lusardi, 2004).

3. Key Strenghts of the PSID

While the key strengths of the PSID have been described throughout this paper, there are several specific features of the PSID that should be specifically highlighted.   These features include consistently high response rates, the longevity of the data collection, a sample that is nationally representative and genealogically-based, content domains that are broad and recurring, and innovative supplements.

As described in some detail above, the PSID is the nation's longest running, nationally representative panel study with information collected on the same families for 34 waves, 1968-2005.   Reinterview response rates have been a consistent 96-98%. Having a long panel of data improves the precision of the measurement as multiple measures are collected within the same families as well as multiple family members over a period of many decades. The extended time series of data allows the study of economic behavior through changing conditions, such as wage variations for different populations during the course of business cycles.   In addition, the longitudinal data facilitate the conduct of cohort analysis as persons from one time period to another may be compared. These data also facilitate developmental analysis, as early experiences may be used to predict longer term outcomes, such as the prediction of earnings from early schooling experience.

 With analysis weights, the PSID data are nationally representative of U.S. families.   Thus, results based on analysis of PSID data may be used to make statements about the U.S. population as well as about specific demographic groups such as those broken out by age, gender, race, and income categories. These nationally representative data have had a significant impact on the development and setting of national public policy in areas such as financial asset ownership, welfare reform, childcare, and the economic policy.

The original family members and their descendants have been PSID sample members since 1968. Having a genealogical panel opens up many opportunities to illuminate human and social behavior within the context of family, including conducting intergenerational analyses to examine similarities and differences between family members from different generations, understand how parental and grandparental experiences and behaviors affect the experiences and behaviors of current generations, and study sibling pairs to understand the relative contributions of family/social environment and genetic influences.

The PSID has collected data on a wide array of social, demographic, economic, geospatial and psychological topics, supporting unique multidisciplinary research. In 2005, the 75 minute interview collected data on: employment, earnings, income from all sources, food expenditures, housing, geospatial data, health and health insurance, educational expenditures, marriage and fertility, participation in government programs, vehicle ownership; wealth and pensions, and philanthropy. Many of these areas have been included in the instrument since 1968. Hundreds of additional variables that fall into other domains have been collected in various waves throughout the history of the PSID. All of the data are available on PSID's web-based data center: www.psidonline.isr.umich.edu

4. Limitations of the PSID

There are several limitations of the PSID. The first limitation is the cumulative attrition that results in part from the longevity of the panel. Despite the fact that the reinterview response rates obtained in the PSID are unprecedented at 96-98%, attrition nonetheless accumulates over time. Over the course of nearly 4 decades, substantial attrition can lead to data that are no longer representative of the U.S. population of households. Tables 4 and 5 summarize the evolution of the sample since 1968. Of the 18,192 individuals in the sample in 1968, 5,282 were alive and interviewed in 2001. The remainder either died, were explicitly dropped from the study in 1997, or attrited.

Cumulative attrition is an issue that is highly salient to the investigators of the project and is under constant review. A sampling statistician is part of the investigative team and evaluates the statistical properties of the sample with respect to its success in representing the U.S. population of households. This evaluation led to the introduction in 1997 of the post-1968 immigrant refresher sample to update the sample. In addition, the statistician reviews the sampling weights that are created for each wave of data collection against a range of key demographic variables to ensure that the data continue to represent the population of family units in the U.S.

A second limitation is the periodicity of PSID's data collection. As described above, for budgetary reasons, PSID went from collecting data annually from 1968-1997, to biennially 1999-2005. Interviewing families every-other-year is particularly disadvantageous for the collection of some information, especially that which is known to be fairly dynamic, such as income and employment. The gathering of high quality, detailed income data is the hallmark of PSID. To address the limitation of having such detail only biennially, PSID implemented a series of questions asking respondents about their income two years ago, in addition to asking detailed questions about income in the prior year. In addition, the Event History Calendar described above was implemented to enhance recall over the two-year period for transitions in the area of employment, a related life

domain marked by frequent change.

Targeting the family head and wife for data collection to the exclusion of other family members is a third limitation. Until 1997, the PSID did not collect much information about children beyond the data gathered about the family in which the child resided. This limitation was addressed in 1997 with the introduction of the first wave of the Child Development Supplement (CDS-I), and again in 2002, with data collected on these same children 5 years later in the second wave (CDS-II). In addition, until 2005, there was scant information collected on most family members until they became PSID heads and wives themselves. Scientists are becoming increasingly aware of the fact that the years from 18-24 are critical for life span development, as it is during this period that major investments are made in education, and careers are planned and initiated. Moreover, these educational and occupational transitions are now typically made while the young adults are still dependent on their parents, long before they become family "heads" and "wives" themselves. To address this limitation, a pilot study was launched in 2005 to collect information in a variety of domains that are collected in parallel in the PSID, such as education, employment, career aspirations, and social and psychological well-being, from the children who participated in the CDS, who had turned at least 18 by 2005. A proposal to continue this data collection for the CDS children who will reach age 18 by 2007 is currently under review. It is hoped that funding can be secured to continue to collect data from all the children at age 18 who participated in the CDS until they become part of the PSID main sample as heads and wives of their own family units. This will provide rich information on these respondents from early (CDS-I) and middle childhood (CDS-II), through adolescence and young adulthood (TA), and throughout their adult life as they participate in the PSID.

A fourth weakness is derived from the limits of the types of data that can be collected over the telephone. Since PSID conducts its interviews via CATI, all information is based on self-report. As there is growing recognition of the importance of data on physical health obtained from actual measurement, the PSID has been evaluating the idea of augmenting its data collection with face-to-face interviews that include obtaining biomarker data (as well as reviewing biomarker data that can be collected through means other than face-to-face). The genealogical design of the PSID implies that such data would be collected from siblings, parents, grandparents, children, and grandchildren, allowing researchers to look at the variation of these factors within the family and over the life course. PSID broadened its collection of measures of health status in 1999 (e.g. BMI, chronic conditions, activities of daily living, instrumental activities of daily living), health behaviors (e.g. smoking history, alcohol use, exercise), access to health care (health insurance coverage and medical expenditures), and psychological distress (2001 and 2003). The idea of augmenting the PSID with biomarker data was reviewed by the PSID Board of Overseers in the mid 1990's.

At that time it was decided to add more in the domain of health and health behavior, but the decision was to stop short of extensive biomarker collection. Recently, in October 2004 and again in November 2005, an ongoing panel has been examining the feasibility of including biomarker information in the PSID. Since the PSID Board's evaluation in the mid late 1990's, the range of assessment possibilities has increased and the cost of such measures in a population sample has in many cases fallen substantially.

A fifth weakness is that new immigrants to the US are not continually represented in the sample. While the sample of post-1968 immigrants added in 1997/1999 incorporate immigrants who arrived between 1968 and 1999, a large number of immigrants have arrived in the US since 1999 and the PSID cannot be used to asses their outcomes.

Figure 1
Steady State
Panel Schematic

(A2) Attrition Stock

(m1)

(a2)   (f2)

(S) Split-off Stock

(m2)

Split-off Flow (s)   (f1) *Followed Split offs*

(N) **Original Sample** (National, SEO)

(n) **Refresher Sample** (Equal-Probability, Immigrants)

(d) **Demographic Inflows** (Births, Marriages)

(P) **Active Panel** (Family, Individual)

(m3)

(M) **Mortality**

Attrition (a1)   (r1) *Recontact*

(A1) **Attrition Stock**

(m4)
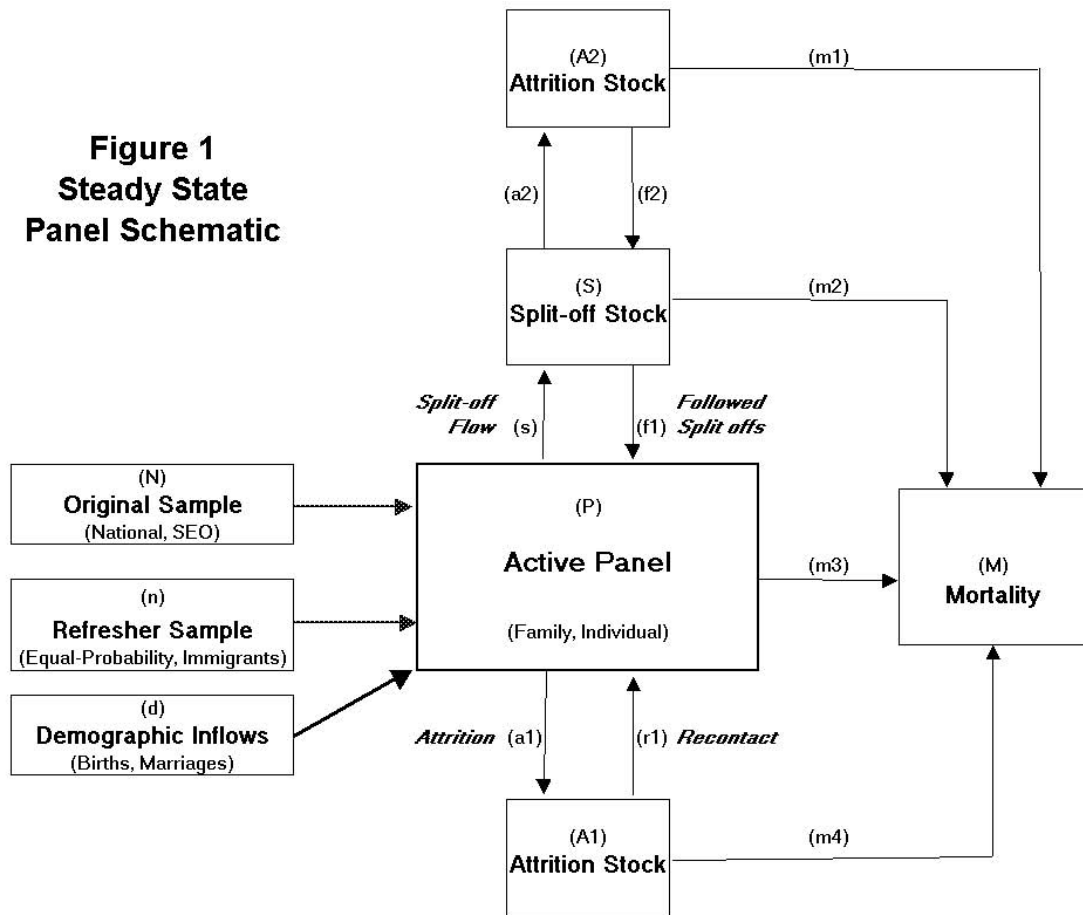
Figure 2. Example of Event History Calendar Used in the 2003 PSID

Figure 3. Number of published journal articles using the PSID within
consecutive 3-year periods: 1972-2004

*Cumulative number of journal
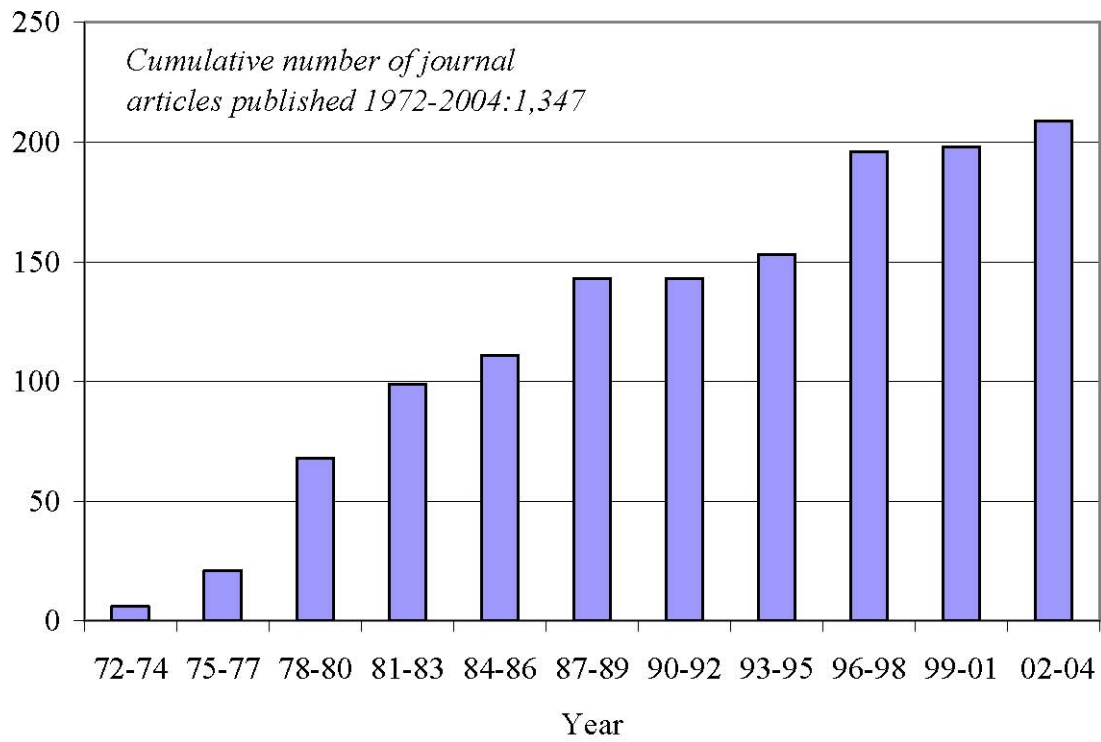articles published 1972-2004:1,347*

Year

Table 1. PSID Response Rates and Sample Sizes: 2003-2005

Sample	Non-Split-offs		Split-offs		Total # of Interviews
	Response Rate (%)	Number of Interviews	Response Rate (%)	Number of Interviews	
2003 Actual					
Core re-interview	97	6554	83	561	7115
Core re-contact	65	200	56	15	215
Core subtotal	95	6754	82	576	7330
Immig. re-interview	94	459	61	36	495
Immig. re-contact	51	42	43	3	45
Immigrant subtotal	88	501	59	39	540
Total	*95*	*7255*	*80*	*615*	*7870*
2005 Actual					
Core re-interview	98	6756	89	516	7272
Core re-contact	61	194	60	3	197
Core subtotal	96	6950	88	519	7469
Immig. re-interview	93	500	74	45	545
Immig. re-contact	42	27	-	0	27
Immigrant subtotal	88	527	74	45	572
Total	*95*	*7477*	*87*	*564*	*8041*

Table 2. PSID 2005 Instrument Content Domains and Section Lengths

Section	Components	Length in Minutes
A	Housing, Utilities, Computer Usage	4.9
B, C, D, E	Employment	20.2
F	Housework, Food Expenses, Food Assistance	2.6
F	Transportation, Education Expenses, Expenditures	6.9
G	Income	6.6
H	General Health	1.1
H	Health Conditions, Behaviors, Expenses	10.7
J	Marriage/Fertility	1.1
K, L	New Head/Wife Background	2.3
M	Philanthropic Giving and Volunteering, Religiosity, Help Received	6.8
P	Pensions	3.5
R	Off-Year Income, Welfare	2.8
W	Wealth and Active Savings	5.2
	Total Questionnaire Length	**74.7**
	Family unit composition update and split-off information, IW close, respondent address and payment information	8.2
	Total Respondent Burden	**82.9**

Table 3. List of Additional Areas of Investigation Supported by PSID Data

- Game-theoretic interactions among family members;
- Theories of the impact of technical change on earnings and life-course mobility;
- Models of spatial equilibrium and neighborhood effects;
- Bequests and altruism tested with measures of savings, wealth, and transfers;
- Models of intergenerational mechanisms in social and economic life;
- Life-course stress and mortality theories tested with long-term health and mortality measures;
- Theories of social acceptance or 'stigma' tested with models used to understand the rise in bankruptcy;
- Changes in welfare policies and incentives for saving by low income families;
- Long-term effects of economic conditions of the family during early childhood;
- Family adaptation to change;
- How labor markets adjust and workers make commitments to large durables during different stages of the business cycle;
- How strong the spending stimulus is from the wealth effect of stock market gains;
- Whether the wealth effect can explain much of the low U.S. savings rate;
- Participation of younger and less educated employees in company-sponsored pension plans;
- Long-term effects of health capital as indexed by healthful activities, nutrition and maintaining proper weight on economic well-being;
- How consumers select among the bewildering array of vehicle types;
- How income and wealth mobility in the US compare with mobility in other countries;
- What share of total wealth is in the form of pensions and whether it protects those who would otherwise save little for retirement;
- Whether Baby Boomers really are bigger spendthrifts than earlier cohorts and if so, whether this is because they expect transfers from their parents;
- Multidisciplinary and life-course research on the roles that the family, school and neighborhood play in fostering or hindering children's learning at home and at school;
- Social and economic determinants and consequences of caring for elderly parents;
- Time use studies as indicators of social and economic well-being of children;
- The effect of early human capital investment on children's development;
- Patterns, causes, and consequences of union formation behavior; and
- Family processes as mediators of how socio-economic characteristics, neighborhood factors, and public policies affect children's cognitive, emotional, and physical well-being.

Table 4. Entry and Exit of Individuals in the PSID

| Year | Total* | Exited the PSID | | | | Entered the PSID | | | | |
		Total	Left	Died	Dropped	Total	Born into PSID	New Sample	Nonsample	Re-entrants
1968	18230	0	0	0	0	0	0	0	0	0
1969	17212	1849	1832	17	0	831	275	21	535	0
1970	17350	725	642	83	0	863	375	7	481	0
1971	17590	616	511	105	0	856	337	0	507	12
1972	18051	569	475	94	0	1030	391	1	608	30
1973	18236	731	603	128	0	916	355	1	536	24
1974	18396	751	637	114	0	911	368	0	519	24
1975	18623	700	593	107	0	927	382	2	520	23
1976	18768	850	741	109	0	995	422	2	525	46
1977	18998	819	707	112	0	1049	435	2	578	34
1978	19140	784	688	96	0	926	369	0	505	52
1979	19443	788	712	76	0	1091	469	2	576	44
1980	19747	910	810	100	0	1214	521	0	634	59
1981	19796	928	814	114	0	977	441	0	486	50
1982	20112	746	662	84	0	1062	427	1	584	50
1983	20327	924	814	110	0	1139	440	0	619	80
1984	20393	1000	882	118	0	1066	459	0	537	70
1985	20680	982	871	111	0	1269	481	0	685	103
1986	20437	1235	1121	114	0	992	368	0	554	70
1987	20486	1044	932	112	0	1093	442	0	553	98
1988	20506	1030	914	116	0	1050	418	0	532	100
1989	20451	1105	991	114	0	1050	381	0	562	107
1990	20745	898	786	112	0	1192	372	0	598	222
1991	20770	964	851	113	0	989	401	0	508	80
1992	21145	1090	956	134	0	1465	356	0	719	390
1993	22311	1493	1355	138	0	2659	457	11	1102	1089
1994	24511	1551	1396	155	0	3751	418	518	1056	1759
1995	23929	1439	1241	198	0	857	258	45	497	57
1996	23810	1247	1148	99	0	1128	351	41	615	121
1997	19761	6695	1287	125	5283	2646	295	1717	471	163
1999	20515	1536	1377	159	0	2290	607	336	897	450
2001	21400	1401	1189	212	0	2286	664	72	1004	546
2003	22292	1499	1283	216	0	2391	699	70	1112	510
Totals				3795	5283		13434	2849		

Number ever in the PSID: 54,727

*Total is the number of people with positive sequence number, and it equals the number in the prior wave minus exits plus entrants.
Totals are excluded from some columns because the number is not meaningful due to multiple spells.

Table 5. Composition of the PSID: Individuals with the PSID Gene

	Individuals in the PSID					
Year	Total	SRC	Immigrant	Original Sample Member	Head or Wife/"Wife"	Heads+Wives in PSID >= 5 Years	Individuals in PSID >= 5 Years
1968	18192	9454	0	18192	7878	0	0
1969	16324	8261	0	16028	7118	0	0
1970	16121	8181	0	15460	7160	0	0
1971	16073	8111	0	15091	7238	0	0
1972	16051	8119	0	14697	7330	6260	14607
1973	15950	8065	0	14279	7445	6268	14428
1974	15879	8022	0	13891	7562	6352	14374
1975	15842	8012	0	13530	7668	6426	14306
1976	15743	7987	0	13078	7707	6483	14165
1977	15701	8004	0	12689	7769	6583	14066
1978	15696	7971	0	12401	7842	6715	14044
1979	15683	7985	0	12036	7995	6814	13958
1980	15720	8032	0	11664	8075	6883	13876
1981	15749	8035	0	11365	8098	6938	13890
1982	15790	8072	0	11108	8160	7033	13891
1983	15833	8099	0	10813	8217	7163	13919
1984	15824	8090	0	10498	8212	7228	13965
1985	15802	8089	0	10168	8258	7221	13891
1986	15601	8008	0	9810	8186	7181	13760
1987	15537	8020	0	9488	8164	7167	13679
1988	15495	8015	0	9209	8160	7151	13675
1989	15391	7969	0	8914	8106	7204	13664
1990	15443	8028	0	8760	8198	7220	13643
1991	15416	8032	0	8504	8178	7248	13677
1992	15557	8113	0	8452	8319	7258	13625
1993	15917	8334	0	8217	8376	7089	13283
1994	17846	9067	0	8624	8944	6988	13256
1995	17532	8919	0	8275	8802	6949	13261
1996	17434	8879	0	8018	8730	7020	13521
1997	14850	8841	1659	5703	7169	5330	10806
1999	15044	8904	1679	5469	7281	5536	11417
2001	15388	9107	1638	5282	7494	5344	11091
2003	15745	9197	1674	5108	7761	5174	10857

References

Altonji, J. , & Dunn, T. (1996, November). Effects of Family Characteristics on the Return to Education. *The Review of Economics and Statistics.* 78, 692-704.

Altonji, J. , Hayashi, F., & Kotlikoff, L. (1997). Parental Altruism and Inter Vivos Transfers: Theory and Evidence. *Journal of Political Economy.* 105(6), 1121-1166.

Amick III, B.C., McDonough, P., Chang, H., Rogers, W.H., Pieper, C.F., & Duncan, G.J. (2002). Relationship Between All-Cause Mortality and Cumulative Working Life Course Psychosocial and Physical Exposures in the United States Labor Market from 1968 to 1992. *Psychosomatic Medicine.* 64(3), 370-381.

Belli, R., Lee, E., Stafford, F., & Chou, C., (2004). Calendar and Question-List Survey Methods: Association Between Interview Behaviors and Data Quality. *Journal of Official Statistics*. 20(2), 185-218.

Belli, R., Shay, W., & Stafford, F. (2001). Event History Calendar and Question List Survey Inverviewing Methods: A Direct Comparison. *Public Opinion Quarterly.* 64, 45-74.

Bjorklund, A., & Jantti, M. (1997). Intergenerational Income Mobility in Sweden Compared to the United States. *The American Economic Review.* 87(5), 1009-1018.

Blank, S. (1998a). Hearth and Home: The Living Arrangements of Mexican Immigrants and U.S.-Born Mexican Americans. *Sociological Forum.*13(1), 35-57.

Blank, S., & Torrecilha, R. (1998b). Understanding the Living Arrangements of Latino Immigrants: A Life Course Approach. *International Migration Review.* 32(1), 3-19.

Blau, F., & Kahn, L. (1996). International Differences in Male Wage Inequality: Institutions Versus Market Forces. *Journal of Political Economy.* 104(August), 791-837.

Bradbury, K., & Katz, J. (2002). Are Lifetime Incomes Growing More Unequal? Looking at New Evidence on Family Income Mobility. *Regional Review.*12(4, Q4), 2-5.

Brines, J., & Joyner, K. (1999). The Ties that Bind: Principles of Cohesion in Cohabitation and Marriage. *American Sociological Review.* 64, 333-355.

Brooks-Gunn, J., Duncan, G., Klebanov, P., & Sealand, N. (1993). Do Neighborhoods Influence Child and Adolescent Development?. *American Journal*

of Sociology. 99(2), 353-95.

Charles, K., Danziger, S., Li, G., & Schoeni, R. (2004). Comparison of Estimates of Spending in the PSID and CEX. Working Paper, University of Michigan.

Charles, K., & Stephens, M., Jr. (2004). Job Displacement, Disability, and Divorce. *Journal of Labor Economics.* 22(2), 489-523.

Conley, D. & Yeung, J. (2005). Black-White Differences in Occupational Prestige: Their Impact on Child Development. *American Behavioral Scientist.* 48(9), 1229-1249

Corcoran, M., & Kunz, J. (1997). Do Unmarried Births Among African-American Teens Lead to Adult Poverty? *Social Service Review.* 71(2), 274-287.

Couch, K., & Dunn, T. (1997). Intergenerational Correlations in Labor Market Status. *The Journal of Human Resources.* 32(1), 210-232.

Couch, K., Daly, M., & Wolf, D. (1999). Time? Money? Both? The Allocation of Resources to Older Parents. *Demography.* 36(2), 219-232.

Crowder, K., & South, S. (2003). Neighborhood Distress and School Dropout: The Variable Significance of Community Context . Social Science Research. 32(4), 659-698.

Davis, M.M, McGonagle, K.A., Schoeni, R.F. (under review). Relative Contributions of Grandparental and Parental Obesity to Risk of Childhood Obesity: Three Generations of American Families.

Davis-Kean, P. (Forthcoming). The Influence of Parent Education and Family Income on Child Achievement: The Indirect Role of Parental Expectations and the Home Environment. Journal of Family Psychology.

Duncan, G., Bjorn, G., Hauser, R., Gunther S., Jenkins, S., Messinger, H., et al. (1995). Poverty and Social-Assistance Dynamics in the United States, Canada, and Western Europe. In Katherine McFate (eds.). *Poverty, Inequality, and the Future of Social Policy: Western States in the New World Order.* New York: Russell Sage

Duncan, G., Smeeding, T., & Rodgers, W. (1995). Household Income Dynamics in the 1970s and 1980s.

Foster, E., & McLanahan, S. (1996). An Illustration of the Use of Instrumental Variables: Do Neighborhood Conditions Affect a Young Person's Chance of

Finishing High School?. *Psychological Methods*. 1(3), 249-260.

Evertsson, M., & Nermo, M. (2004). Dependence within Families and the Division of Labor: Comparing Sweden and the United States. *Journal of Marriage and the Family.* 66(5), 1272-1286.

Foster, E., & McLanahan, S. (1996). An Illustration of the Use of Instrumental Variables: Do Neighborhood Conditions Affect a Young Person's Chance of Finishing High School? *Psychological Methods* 1(3), 249–60.

Foster, E., & Furstenberg, F., Jr. (1998a). Most Disadvantaged Children: Who Are They and Where Do They Live? Journal of Poverty. 2(2), 23-47.

Foster, E., Jones, D., & Hoffman, S. (1998b). The Economic Impact of Nonmarital Childbearing: How Are Older, Single Mothers Faring? *Journal of Marriage and the Family*. 60(1), 163-174.

Furstenberg, F., Jr., Hoffman, S., & Shrestha, L. (1996). The Effect of Divorce on Intergenerational Transfers: New Evidence. *Demography*. 32(3), 319-333

Gouskova, E., & Schoeni, R. (2002a). Comparison of Family Income in the Panel Study of Income Dynamics and the March Current Population Survey: 1968-2000, from http://psidonline.isr.umich.edu/Guide/Quality/psid_vs_cps.pdf .

Gouskova, E., & Schoeni, R. (2002b). Analysis of the Quality of the Health Data in the   PSID,   from http://psidonline.isr.umich.edu/Guide/Report_On_Health_QsV2.pdf
Gray, J. (1998). Divorce-Law Changes, Household Bargaining, and Married Women's Labor Supply. *The American Economic Review* 88(3), 628–42

Harris, D. (1999). Property Values Drop When Blacks Move In, Because...: Racial and Socioeconomic Determinants of Neighborhood Desirability. *American Sociological Review* 64, 461–79.

Haveman, R., Wolfe, B., Kreider, B., & Stone, M. (1994). Market work, wages, and men's health. *Journal of Health Economics.* (13), 163-182

Haveman, R., Wolfe, B., & Peterson, E. (1997). Children of Early Childbearers as Young Adults. In Kids Having Kids: Costs and Social Consequences of Teen Pregnancy. Washington DC: *Urban Institute Press.* 257-284.

Haveman, R., Wolfe, B., & Pence, K. (2001). Intergenerational Effects of Nonmarital and Early Childbearing. in: Wu, L., & Wolfe, B. (2001). *Out of Wedlock: Causes and Consequences of Nonmarital Fertility* (p 287-316). New York: Russell Sage Foundation.

Headey, B., Headey, S., Muffels, R., & Janssen, C. (2004). Who Enjoys the Fruits of Growth? Impact of Governments and Markets on Living Standards in Germany, the Netherlands and the U.S.A., 1987-1996. *Social Indicators Research.* 65(2), 125-144.

Hill, M.S. (1992). The Panel Study of Income Dynamics: A User's Guide. Newbury Park, CA: Sage Publications.

Hill, M.S., & Yeung, W.J. (1999). How Has the Changing Structure of Opportunities Affected Transitions to Adulthood?. In A. Booth, A. Crouter & M. Shanahan, (Eds.), *Transitions to Adulthood in a Changing Economy: No Work, No Family, No Future?.* (pp. 3-39). Westport CT: Greenwood Publishing Group, Inc.

Hill, M.S., Yeung, W.J., & Duncan, G. (2001). Childhood Family Structure and Young Adult Behaviors. *Journal of Population Economics.* 14(2), 271-299.

Hofferth, S., Smith, J., McLoyd, V., & Finkelstein P. (2000). Achievement and Behavior Among Children of Welfare Recipients, Welfare Leavers, and Low Income Single Mothers. *Journal of Social Issues.* 56(4), 747-773.

Hofferth, S. & Reid, L. (2002). Early Childbearing and Children's Achievement and Behavior Over Time. *Perspectives on Sexual and Reproductive Health. 34*(1), 41-49.

Hoffman, S., Foster, E., & Furstenberg, F., Jr. (1993). Reevaluating the Costs of Teenage Childbearing. *Demography.* 30(1), 1-13.

Hurst, E., & Lusardi, A. (2004). Liquidity Constraints, Household Wealth, and Entrepreneurship. *Journal of Political Economy.* 112(2), 319-347.

Hurst, E., Luoh, M., & Stafford, F. (1998). The Wealth Dynamics of American Families, 1984-94. *Brookings Papers on Economic Activity.* (1), 267-337.

Jayakody, R., (1998). Race Differences in Intergenerational Financial Assistance. *Journal of Family Issues.* 19(5), 508-533

Johnson, R., & Schoeni, R. (2006). Health Dynamics and the Evolution of Health Inequality over the Life Course: The Importance of Neighborhood and Family Background. Working Paper, University of Michigan.

Lillard, L., & Waite, L. (1995). 'Til Death Do Us Part: Marital Disruption and Mortality. *American Journal of Sociology.* 100(5), 1131-1156.

Lillard, L., &. Panis, C. (1996). Marital Status and Mortality: The Role of Health. *Demography* 33(3), 313–27.

McDonough, P., & Berglund, P. (2003). Histories of Poverty and Self-Rated Health Trajectories. *Journal of Health and Social Behavior.* 44(2), 198-214.

McDonough, P., & Amick, III, B. (2001). The Social Context of Health Selection: A Longitudinal Study of Health and Employment. *Social Science and Medicine.* 53(1), 135-145.

McDonough, P., Williams, D., House, J., & Duncan, G. (1999). Gender and the Socioeconomic Gradient in Mortality. *Journal of Health and Social Behavior.* 40(1), 17-31.

McDonough, P., Duncan, G.J., Williams, D., & House, J. (1997). Income Dynamics and Adult Mortality in the U.S., 1972-1989. *American Journal of Public Health.* 87(9), 1476-1483.

McLanahan, S., & Sandefur, G. (1994). *Growing Up with a Single Parent: What Hurts, What Helps.* Cambridge, MA: Harvard University Press.

Moffitt, R., Reville, R., &. Winkler, A. (1998). Beyond Single Mothers: Cohabitation and Marriage in the AFDC Program. *Demography* 35(3), 259–78.

Mulligan, C., (1997). The Intergenerational Dynamics of Consumption, Earnings, Income, and Wealth. *Parental Priorities and Economic Inequality.* Chicago and London: The University of Chicago Press.

Mulligan, Casey B. Borrowing Constraints and the Persistence of Inequality. . *Parental Priorities and Economic Inequality.* Chicago and London: The University of Chicago Press; 1997.

Nakosteen, R., & Zimmer, M. (1997). Men, Money, and Marriage: Are High Earners More Prone than Low Earners to Marry? *Social Science Quarterly* 78(1), 66–82.

Ono, H. (1998). Husbands' and Wives' Resources and Marital Dissolution. *Journal of Marriage and the Family* 60(3), 674–89.

Powell, M., & Parcel, T. (1997). Effects of Family Structure on the Earnings Attainment Process: Differences by Gender. *Journal of Marriage and the Family.* 59(2), 419-433.

Quillian, L. (1999). Migration Patterns and the Growth of High-Poverty Neighborhoods, 1970–1990. *American Journal of Sociology* 105(1), 1–37.

Rank, M., & Hirschl, T. (2001). Rags or Riches: Estimating the Probabilities of Poverty and Affluence Across the Adult American Life Span. *Social Science Quarterly.* 82, 651-669.

Rank, M., & Hirschl, T. (2001). The Occurrence of Poverty across the Life Cycle: Evidence from the PSID. *Journal of Policy Analysis and Management.* 20(4), 737-755. Sandberg, J., & Hofferth, S. (2001) Changes in Children's Time with Parents, U.S. 1981-1997. *Demography.* 38(8), 423-436.

Smock, P., & Manning, W. (1997). Nonresident Parents' Characteristics and Child Support. *Journal of Marriage and the Family.* 59(4), 798-808

Solon, G., Page, M., & Duncan, G. (2000). Correlations Between Neighboring Children in Their Subsequent Educational Attainment. *The Review of Economics and Statistics.* 82(3), 383-392.

Solon, G. (1992). Intergenerational Income Mobility in the United States. *The American Economic Review.* 82(3), 393-408.

South, S. (1999). Historical Changes and Life Course Variation in the Determinants of Premarital Childbearing. *Journal of Marriage and the Family.* 61, 752-763.

South, S. (2001a). The Variable Effects of Family Background on the Timing of First Marriage: United States, 1969-1993. *Social Science Research.* 30(4), 606-626.

South, S. (2001b). The Geographic Context of Divorce: Do Neighborhoods Matter? *Journal of Marriage and the Family.* 63(3), 755-766.

South, S., & Crowder, K. (1997). Escaping Distressed Neighborhoods: Individual, Community and Metropolitan Influences. *American Journal of Sociology.* 103(1), 1040-1084.

South, S., & Crowder, K. (1999). Neighborhood Effects on Family Formation: Concentrated Poverty and Beyond. *American Sociological Review.* 64(1), 113-132.

South, S., & Crowder, K. (2000). The Declining Significance of Neighborhoods? Marital Transitions in Community Context. *Social Forces.* 78(3), 1067-1099.

Sucoff, C., & Upchurch, D. (1998). Neighborhood Context and the Risk of Childbearing Among Metropolitan-Area Black Adolescents. *American Sociological Review* 63(4), 571–585.

Van Kerm, P. (2004). What Lies Behind Income Mobility? Reranking and Distributional Change in Belgium, Western Germany and the USA. *Economica.* 71(282), 223-239.

Vanderklauw, W. (1996). Female Labor Supply and Marital Status Decisions: A Life-Cycle Model. *Review of Economic Studies* 63(2), 199–235.

Vandewater, E., Shim, M., & Caplovitz, A. (2004). Linking Obesity and Activity Level with Children's Television and Video Game Use. *Journal of Adolescence.* 27(1), 71-85.

Whittington, L & Alm, J. (1997). 'Til Death or Taxes Do Us Part: The Effect of Income Taxation on Divorce. *The Journal of Human Resources* 32(2), 388–412.

Wilhelm, M., (1996). Bequest Behavior and the Effect of Heirs' Earnings: Testing the Altruistic Model of Bequests. *The American Economic Review.* 86(4), 874-92.

Wolfe, B., & Haveman, R. (1983). Time Allocation, Market Work, and Changes in Female Health. *The American Economic Review.* 73(2), 134-140.