

**Appendix S**  
**Details of the Sample and Data**

**for**

**Intergenerational Transfers and the Prospects for Increasing Wealth Inequality**

Stephen L. Morgan  
Cornell University

John C. Scott  
Cornell University

## Selection of the Sample

This study is based on six waves of the Health and Retirement Study (HRS), 1992 through 2002. The HRS is a nationally representative longitudinal sample of 12,652 respondents that focuses on persons aged 51 to 61 in 1992 and their spouses or partners. The HRS collects extensive information regarding demographic, employment, pension, health, family structure, and financial characteristics of age-eligible respondents and their spouses or partners. The HRS survey design is a multistage area probability sample of households that is augmented by oversampling of African-Americans, Hispanics, and residents of the state of Florida.

The HRS is made up of four sub-samples, but we make use of one in creating two age-based cohorts as discussed in the main article. The sub-sample<sup>1</sup> we use is the initial HRS cohort, which was born between 1931 and 1941 and who were household residents in the U.S. in the spring of 1992. This cohort was interviewed first in 1992 and subsequently every two years. From this sub-sample, we create two cohorts for study, one aged 51 to 53 and another aged 59 to 61 in 1992.<sup>2</sup>

Our primary unit of analysis is the individual, but we also utilize the household-level data that is provided by the HRS. Generally, all age-eligible persons and their spouses, if any, were interviewed in each household, and in some cases certain household-level information was also obtained from age-ineligible persons.<sup>3</sup>

---

<sup>1</sup> In 1998, an additional survey and samples were merged into the HRS with the result that there are a total of 26,728 observations as of 2002. The second subsample is the Study of Assets and Health Dynamics Among the Oldest Old (AHEAD), which consists of U.S. household residents who were born in 1923 or earlier. The third subsample is the Children of the Depression, who are those born between 1924 and 1930. The fourth subsample is the War Baby cohort, born between 1942 and 1947, were household residents in the U.S. in the spring of 1992, who, at that time, did not have a spouse or partner born before 1924 or between 1931 and 1941.

<sup>2</sup> We also make use of the RAND HRS Data files (RAND), a cleaned, processed, and streamlined collection of variables derived from the HRS. While not officially part of the HRS, the RAND files were created by the RAND organization and HRS staff and are publicly available through the HRS website. However, RAND is not a complete reproduction of all HRS variables. For a limited number of variables, particularly those relating to occupational history, pension coverage, and family transfers, we rely on the original HRS data files. In addition, we excluded 109 observations that overlap with the AHEAD survey that was incorporated into the RAND and HRS datasets in 1998 because the RAND files code these observations according to the AHEAD identification coding scheme.

<sup>3</sup> The selection of respondents was somewhat complicated. After selection of a household, the interviewer determined the number of age-eligible persons within that household. If the selected household contained only one age-eligible person or if there were two age-eligible persons who were married/partnered to each other, no respondent selection procedure was required. The single person or both partners were designated as the financial unit to be interviewed. If there was more than one age-eligible person and they were not married or in an equivalent relationship, the interviewer objectively selected a single eligible respondent to be interviewed. If the single age-eligible person had a spouse, the spouse was also designated for the HRS person interview whether or not the spouse was age-eligible. Unmarried age-eligible respondents were automatically designated the “R1” or primary household respondent. In the case of a married couple, the persons who considered themselves more knowledgeable about the family’s assets, debts, and retirement were designated the “R1” or primary respondent and the spouse became the “R2” or secondary respondent. In a married couple household financial unit, the “R1” respondent was not necessarily age-eligible. It is also important to keep in mind that often spouses, children or others in the household were interviewed in

***Creation of Cohorts and Attrition.*** Our study focuses on comparisons of two cohorts that consist of 3-year age groups, those persons aged 51 to 53 and 59 to 61 in 1992. We chose these age groupings by balancing concerns over creating groups with enough distinction from each other in order to be meaningful and having enough observations for practical statistical analysis. Both self-employed persons and respondents working for someone else are included as are those persons working and those retired or otherwise out of the labor force in 1992. From the original 1992 sample of 12,652 observations, the 51 to 53 year-old cohort consists of 2,869 respondents. With attrition, this cohort has 2,216 respondents by the time the group reaches the ages of 59 to 61 in the year 2000. The cohort of those aged 59 to 61 in 1992 consists of 2,320 respondents. Table S1 provides sample attrition and death patterns between 1992 and 2000 for the full HRS sample.

[INSERT TABLE S1 ABOUT HERE]

Given that additional observations are added to HRS cohorts to some degree over time,<sup>4</sup> our cohorts are determined at the stage of initial interview. The cohorts are not augmented by additional persons of the same age in later waves, and the cohorts suffer attrition over time. Approximately between 5 and 10 percent of the survey is lost to attrition from one wave to the next.

The cumulative attrition across all five waves for the original HRS sample and for the aged 51-53 cohort is around 22 to 23 percent while the aged 59-61 cohort exhibits a slightly larger loss of 26 percent by 2000. Relatively more of the 59-61 cohort was lost to death as compared to the 51-53 cohort.

The loss in terms of attrition does not appear to be spread evenly over all socioeconomic groups. Persons of higher socioeconomic status appear to “survive” the survey waves more so than those of lower economic status.

Table S2 provides the unweighted mean years of fathers’ education for the original HRS sub-sample and for each age group in 1992 and for those remaining in 2000. In each case, the mean of fathers’ education increases over time, indicating that those respondents who came from a lower socioeconomic background were more likely to drop out of the survey over time. Similarly, the percent nonwhite also declines from 1992 and 2000 as the overall sample and cohorts undergo attrition, as Table S3 shows, likely reflecting racial and ethnic disparities in health.

[INSERT TABLES S2 AND S3 ABOUT HERE]

Table S4 provides the means for fathers’ education over the sample waves but using HRS sampling weights, and these results still show an increase in mean level of educational attainment over the course of the HRS surveys. Thus, using the 1992 sampling weight

---

addition or in place of the primary respondent where appropriate or necessary. HRS provides both person-level and household sampling weights for these reasons.

<sup>4</sup> In 1998, the HRS made two general changes. The AHEAD survey, which began in 1993, was folded into the HRS, and two new cohorts were added (the Children of the Depression and the War Babies cohorts).

adjusts for the oversample of individuals of low socioeconomic status but not for selective attrition.

[INSERT TABLE S4 ABOUT HERE]

As described in the main article, with reference to Table 1, we construct weights that allow us to reweight the younger cohort so as to balance attrition patterns across the two cohorts.

### **Missing Data**

As shown in Table A1 (see Appendix A, also on this website), missing data patterns are substantial for some variables, especially for the data on intergenerational transfers (in large part because not all HRS respondents have children). In contrast, the wealth measures presented in Table A1 appear to have no missing data. This is somewhat misleading because the wealth variables are composites of underlying variables for which RAND imputed missing values in the process of constructing the composite variables. Since wealth is our dependent variable in much of our analysis, our reliance on imputed data is somewhat of a concern.<sup>5</sup> Our position is that the RAND imputations were accomplished with considerable skill (as random draws from informative predictive distributions, modeled in considerable detail). And, yet, for dynamic longitudinal analysis, the imputations are somewhat limited because wealth in subsequent years of the survey is not imputed based on past levels of wealth. Some of the true over-time dependence of wealth is not represented fully in the imputed data.

### **Coding of Variables**

**Wealth and Income.** Our primary outcome variable is total household wealth, and we use the RAND dataset for our primary measures of wealth. Tables A1 (in Appendix A) and Table 2 in the main article provide descriptive statistics on total household wealth and other explanatory variables for two of the age groups at the same point in the lifespan, those aged 59-61 in 1992 and 59-61 in 2000 (the latter being between the ages 51 and 53 in 1992).

Total net wealth equals assets (primary residence, other real estate, vehicles, businesses, individual retirement accounts, stocks, bonds, checking, certificate of deposits, savings bonds, treasury bills, and other savings) minus liabilities (primary residence mortgage, other real estate debt, and all other debt). Net financial wealth includes stock, bonds, checking accounts, certificates of deposits, savings bonds, treasury bills, and other savings minus other debt (aside from real estate debt). Retirement plan wealth includes the value of accounts in defined contribution plans, individual retirement accounts, and Keogh accounts (Retirement wealth includes amounts held in defined contribution

---

<sup>5</sup> Virtually all studies of wealth accumulation (see Wolff 1998) have had to rely to some extent on imputed wealth data. It has proven nearly impossible to measure exhaustively all of the underlying dimensions of wealth, in part because of gaps in the knowledge and memory of respondents as well as concerns about guarantees of anonymity.

retirement plans, which is based on an HRS Researcher Contribution file. The defined contribution data contains imputations for pension data from self-reported data on current jobs. The data set may be downloaded from the HRS Public File Download Area at <https://ssl.isr.umich.edu/hrs/> (free registration required)). Net value of housing is the value of primary residence less real estate debt. Total other wealth equals the sum of the net value of real estate (other than the primary residence), vehicles, and businesses.

Total household income is the sum of all income in the household, including individual earnings, capital income, pension income, Social Security disability/SSI/retirement income, unemployment compensation, other government transfers, and other sources of income such as insurance. Total household earnings are the sum of the wage and salary income of the respondent and spouse, if any. Total non-wage income is the difference between total household income minus total household earnings.

***Wealth and Income: Adjustment for Inflation.*** The wealth and income variables we analyzed were converted to constant 2000 dollars using the Personal Consumption Expenditures Deflator (PCED) of the Bureau of Economic Analysis (U.S. Department of Commerce), which is now favored over the traditional CPI-based adjustment. See in particular NIPA Table 2.3.4 (July 2006) that contains the deflator of 85.824 for nominal 1992 dollars.<sup>6</sup>

***Pension Plans.*** Much of the HRS is concerned with pension and retirement plans, and it provides a great deal of information on respondents interaction (or the lack thereof) with pension plans. Based on HRS information, we created dummy variables that indicate whether the respondent is in a defined benefit pension plan (DB plan) or a defined contribution retirement plan (DC plan). If a respondent stated that he or she was included in a workplace pension plan, the HRS asked respondents if the plan was (1) a DB-type plan, (2) a DC-type plan, or (2) both. Coding of our plan-type dummy variables assigns a “1” for the DB (DC) plan if they indicated they were in either a DB (DC) plan or a combination of DB and DC plans.

***Amounts of transfers to children*** – Given the focus of our paper on transfers from HRS respondents to their children, we compare the intergenerational transfers of two cohorts who are at similar ages, and in another part we focus on the cumulative transfers of the younger cohort (those workers aged 59 to 61 in 2000). Creating variables that measured the amount of transfers was problematic chiefly, but not wholly, due to the fact that the HRS did not ask consistently worded questions regarding transfer amounts at each survey wave. We describe the changes in question wording in the main article of the article.

We also discuss how we resolved the alternative recall period for these two different questions. Here, we discuss this issue again, but first we recount the full set of other problems that we encountered when constructing transfer amount variables and how we approached such problems.

---

<sup>6</sup> Table 2.3.4 can be accessed on the web (as of July 2006) at <http://www.bea.gov/bea/dn/nipaweb/TableView.asp?SelectedTable=64&FirstYear=2004&LastYear=2006&Freq=Qtr>

Child-level data – One issue has to do with the structure of the data itself. Data on transfer amounts to children are stored in child-level files. In such files, each child is an observation with a unique personal identifier, and the data on transfers from respondent-parents is an attribute attached to each child-level observation. In order to merge this information with the main data, which is organized at the respondent and household levels, we restructured the child information in order to create a household-level file. In other words, the new file has each household as the unit of observation with transfers, if any, to each child within the household as a separate variable. Child-specific transfer variables are then summed to create an aggregate amount for each household. We then merged this household-level file into the main dataset using the household identifier.

Refusals and Don't Knows – The 2000 data contained a small number of those respondents who refused to answer or did not know about transfers while the 1992 data did not contain such answers. We coded the 2000 transfer variables as missing if the respondent refused to answer or did not know.

Resident versus nonresident children, grandchildren – In the 1992 wave of the HRS, the survey asked the transfer questions only of those households with nonresident children (nonresident children include those who were not living at home or temporarily away at school). In contrast, the 2000 questionnaire asks the transfer questions of all households with children. As a result, the number of nonzero responses and the amount of transfers themselves in 2000 are significantly larger than the 1992 responses. In addition, the 2000 question asks about transfers to children or grandchildren. In order to make meaningful comparisons, we impute the values for the 1992 respondents who have children at home or temporarily away at school. The 1992 transfer variables were imputed using 1992 variables on respondent's race, gender, education, wealth, income, earnings, the probability of a large bequest, work and self-employment status, job tenure, occupation and industry, number of children, number of children at home, whether parents were living, whether in a coupled relationship, and parents' education. Table S5 compares the effect of imputation for those aged 59-61 in 1992 in the number of households making transfers as well as the amount of transfers

[INSERT TABLE S5 ABOUT HERE]

For the issue of grandchildren in the 2000 survey, the questions on amounts asks who is the parent of the grandchild, thus crediting grandchild transfers to the child of the respondent. This may or may not be an issue due to the ambiguity of transfers in general. For example, the 1992 respondents may have responded affirmatively if they made transfers to child for the benefit of grandchildren, but the questions do not probe respondents' intentions about the ultimate recipient of the transfer. While the 2000 data could be inflated as a result of the expanded question wording, we did not make any adjustments.

Threshold dollar amount – As noted above, the preliminary question in both the 1992 and 2000 surveys is whether the respondent provided financial assistance of at least \$500 to

the respondent's child(ren). An affirmative answer then led to the follow-up question on amounts. However, using the price index for personal consumption expenditures, \$500 in 1992 is worth \$583 in 2000 dollars. To ensure that the same types of respondents are included in our comparison, we recode the transfer amounts to 0 for those persons in the 2000 survey who transferred less than \$583 in 2000 (subject to the further adjustment for the timeframe of the question, which is discussed below).

Timeframe – Each question on transfer amounts asks for the amount of the transfer over a period of time, but this time period varies across the HRS surveys. In 1992 and 1994, the questions are framed for transfers over the past 12 months while the 1996 through 2000 waves are posed in terms of the past 2 years (or since the last time of the last interview). Thus, our 1992 and 1994 data may be understating the amount of transfers relative to amounts reported for 1996 through 2000. To address this issue, we inflate the amounts of transfers in 1992 and 1994 by a factor of 2, which is necessary to raise the mean level of transfers in those years to a level roughly comparable to the mean transfers in 1996 through 2000. (This issue intersects with the threshold issue discussed above and that we resolve by recoding the transfers to 0 for those persons in the 2000 survey who transferred less than \$1,166 (2 times \$583) in 2000.) As we note in the main article, it is possible that this doubling of the 1992 and 1994 transfers may lead to a slight over-estimation of the transfer amounts in 1992 and 1994.

**Table S1. Number of Respondents and Attrition Patterns for the Health and Retirement Surveys from 1992 through 2000**

	1992		1994		1996		1998		2000	
	N	%	N	%	N	%	N	%	N	%
<u>All Ages in HRS 1992 Sample</u>										
Respondents in sample	12,652	100	11,424	90.29	10,775	85.16	10,242	80.95	9,630	76.11
Respondents missing	0	0	1,054	8.33	1,370	10.83	1,616	12.77	1,873	14.80
Respondents deceased	0	0	174	1.38	507	4.01	794	6.28	1,149	9.08
<u>Aged 51-53 in 1992</u>										
Respondents in sample	2,869	100	2,585	90.10	2,449	85.36	2,337	81.46	2,216	77.24
Respondents missing	0	0	250	8.71	337	11.75	406	14.15	474	16.52
Respondents deceased	0	0	34	1.18	83	2.89	126	4.39	179	6.24
<u>Aged 59-61 in 1992</u>										
Respondents in sample	2,320	100	2,092	90.17	1,960	84.48	1,828	78.79	1,707	73.58
Respondents missing	0	0	162	6.98	226	9.74	292	12.59	320	13.79
Respondents deceased	0	0	66	2.84	134	5.78	199	8.62	293	12.63

*Notes:* In this table, “missing” includes both those who are alive and did not respond in a particular wave and those persons of whom it is not known whether they are alive or dead.

*Source:* HRS, 1992-200



**Table S2. Mean years of education for respondents' fathers as reported in 1992 (std. dev. in parentheses), overall and by age groups, in 1992 and 2000**

	1992 sample	2000 sample
Full HRS 1992 Sample		
Father's education	8.946 (4.009) n=11,066	9.050 (3.998) n=8,534
Aged 51-53 in 1992		
Father's education	9.207 (4.027) n=2,554	9.267 (3.996) n=1,986
Aged 59-61 in 1992		
Father's education	8.526 (4.079) n=1,973	8.631 (4.069) n=1,468

*Source:* HRS, 1992-2000

**Table S3. Percentage nonwhite, overall and by age groups, in 1992 and 2000**

	1992	2000
Full HRS 1992 Sample		
Pct. nonwhite	20.55	19.18
	n=12,652	n=9,630
Aged 51-53 in 1992		
Pct. nonwhite	21.78	20.57
	n=2,869	n=2,216
Aged 59-61 in 1992		
Pct. nonwhite	20.30	19.15
	n=2,320	n=1,707

Source: HRS, 1992-2000

**Table S4. Weighted mean years of education for respondents' fathers (std. dev. in parentheses), overall and by age groups, in 1992 and 2000**

	1992	2000
Full HRS 1992 Sample		
Father's education	9.244 (3.879) n=11,066	9.349 (3.868) n=8,534
Aged 51-53 in 1992		
Father's education	9.542 (0.078) n=2,554	9.606 (0.088) n=1,986
Aged 59-61 in 1992		
Father's education	8.792 (4.079) n=1,973	8.870 (4.069) n=1,468

Source: HRS, 1992-200

**Table S5. Descriptive statistics for financial transfers to children, by age groups 59-61 in 1992 and 59-61 in 2000**

Variables	Aged 59-61 in 1992 before imputation			Aged 59-61 in 1992 after imputation			Aged 59-61 in 2000		
	n	Mean	Std. Dev.	n	Mean	Std. Dev.	n	Mean	Std. Dev.
Financial Transfers (all households with children)	2,135	952	5,349	2,135	1,647	6,407	2,183	2,877	9,665
Financial Transfers (only households that made transfers)	399	5,094	11,501	624	5,637	10,867	837	7,504	14,458

Note: All dollar values have been converted to 2000 dollars using the Personal Consumption Expenditures Deflator. The 1992 transfer variables were imputed using 1992 variables on respondent's race, gender, education, wealth, income, earnings, the probability of a large bequest, work and self-employment status, job tenure, occupation and industry, number of children, number of children at home, whether parents were living, whether in a coupled relationship, and parents' education.

Source: HRS, 1992-2000