

**Family Data and Research in the
Health and Retirement Study (HRS)**

Suzanne M. Bianchi

UCLA, Department of Sociology

Box 951551, 264 Haines Hall

Los Angeles, CA 90095-1551

bianchi@soc.ucla.edu

May 12, 2010

Paper prepared for the Data Monitoring Committee of the HRS and the National Institute
on Aging (NIA)

Summary Recommendations

The Health and Retirement study is a unique resource for the study of family relationships, particularly intergenerational family exchanges of time, money and coresidence. The data collection is high quality, innovative in its use of multi-methods and experimental modules for testing new content on family relationships. The HRS has been the source of a large number of papers and dissertations on family-related topics, including a number of highly cited articles on topics such as couple's coordination of retirement and financial transfers between older mothers and adult children.

In keeping with the charge to “provide critical assessments of the utility of the HRS data for addressing important research questions...and [provide] suggestions for improving the study's analytic potential particularly in light of expected developments in the coming decade” I offer the following four suggestions for making the HRS even more useful in the future for the study of family relationships. These are:

1) Reevaluate decisions about the family members on whom to gather information. In particular, the decision to only get information on siblings when there is a living parent may not be sufficient for the study of who gives or receives assistance later in the life course of HRS respondents and may partially explain why these data have not been used much.

2) Consider collecting more information directly from each spouse (and perhaps expanding the definition of “spouse” to include cohabiting partners as cohabitation is on the rise among cohorts now entering the HRS). The use of one spouse for reporting all family matters implicitly assumes a unitary decision-making model. Couple's closeness to kin, even children, and willingness to make transfers may actually diverge, especially among cohorts now being enrolled in HRS that are much more likely to have step- or unequally shared - kin ties.

3) Begin to experiment with interviewing adult children of HRS respondents and consider broadening the content of what is asked to include more on relationship quality and/or on everyday activities and exchanges. Currently, the HRS content is weighted toward ascertaining what the HRS respondent does for others, particularly in the realm of intensive time transfers such as care for grandchildren and care for disabled parents. This focus was perhaps sufficient when the HRS cohort was age 51-61 in 1992 but at ages 69 to 79, the current age of the HRS cohort, what the cohort receives in care from adult children takes on increased importance and this will grow as the cohort ages. Direct reports of exchanges of support and relationship quality from both generations would make the HRS an even more valuable resource for the study of intergenerational relationships than it already is.

4) Make the family data easier to access and use.

Background

The Health and Retirement Study (HRS) is one of a small number of longitudinal data collections in the United States that is extremely valuable for the study of family relationships and intergenerational ties, particularly intergenerational family exchanges of time, money and coresidence.¹ It is arguably the most important U.S. data collection for studying families at later stages of the life cycle, given its national representativeness, large sample size, expansive data collection on employment, income and assets, and its information on biomeasures, health status and health care.

A number of initial design decisions make the HRS particularly valuable for studying marital dyads and transfers among extended kin. First, the decision to obtain information on both spouses in married couples facilitates studies of the coordination of retirement of spouses and the responses of one spouse to (changes in) the health of the other spouse. The collection of data on both members of a couple also creates valuable parallelism in the information available on extended family members (e.g., information on parents-in-law as well as parents).² Second, the HRS made the important decision to gather information on *all* of a respondent's children, not just a "focal" child as was done in the National Survey of Family and Households (NSFH), the most widely used panel data set for the study of coresidence and transfers of time/care in family research.

¹ This section is informed by a review of major data sets for the study of intergenerational relationships that was carried out as part of an NICHD-funded project to assess data needs for the study of family change and diversity in the U.S. The evaluation of longitudinal data sets can be found in California Center for Population Research – On Line Working Paper CCPR-020-07 and is listed in the references under Bianchi et al. (2007).

² Data on parents may be reported with more accuracy than data on parents-in-law. According to Soldo and Hill (1995), there was more missing data on reports about parents-in-law in the first wave of the HRS data collection. Because the family respondent, usually the wife, reports this information, the quality of data on her kin may be better than on her husband's kin.

Knowing the characteristics of the full sibship is important in studying inequality in parental inter-vivos transfers and bequests.

The regularity of the HRS panel data collection – every two years - is a great strength when compared with many of the most commonly used data collections in family research. Although an every other year interview schedule may miss important employment and health events,³ other panel data sets used in family research have frequently had longer and often irregular intervals between the waves. This is true of the NSFH and the National Longitudinal Study of Adolescent Health (Add Health), to name two nationally representative panel data sets that are frequently used in family research. Long and uneven intervals between rounds of data collection increase the challenges in making causal attribution and identifying causal mechanisms and pathways. Although such challenges also exist for panel studies with regular data collection intervals like the HRS (and the PSID and NLSY), they are more tractable with a repeated, stable interview schedule.

Finally, the methodological experimentation in the HRS (e.g., the number of experimental modules with family content) and the use of mixed methods (e.g, the 2004 and 2006 Psychosocial Leave-Behind Participant Lifestyle Questionnaires) have added content to the HRS that is highly relevant to family researchers. The cohort design of the HRS offers considerable possibility for assessing family change, particularly given the demographic shifts that are occurring across cohorts in the study.

³ This was discussed in the last set of evaluation of the HRS (e.g., see Hayward (2002: 9) who described the issue of missing disability events with a 2-year window).

In the subsequent discussion, I review some of the family content and research that has been done to date and offer observations about what has been most successful and what might be done to enhance the panel uses of the HRS family data going forward.

Family Research with the HRS Data

The HRS website hosts a bibliography of 240 working papers, dissertations, and articles/book chapters under the topic of “Families and Transfers.” In fact, this appears to be the third most substantial area of research using the HRS – after employment (248 entries)/retirement (239 entries) and health (666 entries)/health care (211 entries). One might expect that a study called the Health and Retirement Study would have large numbers of publications in the two areas of health and labor force/retirement. But the extensive use of the data to study issues related to family life is a tribute both to the usefulness of the design for family research and to the importance of understanding family dynamics and the role they play in later life health and retirement decisions and general well-being.

In the discussion that follows, I first comment on the demographic factors that are changing the nature of families across cohorts of the HRS and that seem particularly relevant for assessing current and future data collection on families and potential data gaps in the HRS. Then I turn to the data and research on various family relationships.

Demographic Change in the Family

One of the advantages of the HRS design is the ability to compare cohorts of older individuals and family members at similar points in the life course. This advantage

is increasing with the routine introduction of new panels of respondents as they reach age 50 and as earlier cohorts are covered by more waves of data collection. The current sampling design and survey content were generated in preparation for the 1992 interview with the HRS cohort, a cohort born between 1931 and 1941 and age 51-61 at first interview. This cohort is roughly age 69-79 at the 2010 interview, with its oldest members having completed childbearing during the Baby Boom years and its youngest members having childbearing years that extended into the Baby Bust years of the 1970s.

The HRS is now beginning to enroll these Baby Boom Cohorts, with the addition of the EBB- Early Baby Boomers (born 1948-53) in 2004 and the MBB – Middle Baby Boomers (born 1954-59) in 2010. Their family experiences differ in important ways from the HRS cohort and the CODA cohort before them (i.e., those born in 1924-30 and who unfortunately were not enrolled until 1998 when they were already age 68-74 - well past age 50, in fact more the age of the AHEAD cohort enrolled in 1993). These new cohorts increase the complexity of accurately gathering data on extended kin but also enhance the usefulness of good data collection on extended family members for studying family change and health and well-being outcomes. For example,

- EBB and MBB cohorts have much higher rates of childlessness - estimates of 20% (or more) compared with 10% for the CODA/HRS cohorts, with rates of childlessness even higher for the highly educated (Dye 2007). Prior research with the AHEAD data suggests a positive correlation between childlessness and nursing home entry (Aykan, 2003).
- Much larger proportions of EBB and MBB cohorts delayed marriage and childbearing until age 30 (or later) (compared with HRS/CODA) which increases

the likelihood that these cohorts will be supporting adolescent and young adult children at the same time as their parents need care and as they begin to face decisions about retirement and face their own initial health crises.

- The EBB and MBB cohorts will have more siblings but fewer children than the HRS/CODA cohorts. We know how these changes in the relative size of generations will affect the availability of care for older adults, particularly the availability of informal family care. More siblings may reduce the burden on any given individual for parental care or it may just result in more “shirkers” of their “filial obligation.” Similarly for children: perhaps all one needs is one responsible child late in life but the likelihood of having that “responsible child” may be correlated with family size in unexpected ways.
- There is little question that EBB and WBB cohorts of women will have higher rates of (lifetime) labor force participation and more accumulated work experience by age 50 than HRS/CODA cohorts of women. Other things equal, this should add complexity to joint retirement decision, increase financial assets later in life, and perhaps also reduce time and caregiving reserves in extended families.
- The EBB and MBB cohorts have experienced much higher rates of nonmarital childbearing, marital disruption, and informal cohabiting unions before and after marriage (compared with HRS/CODA), thereby increasing the complexity of collecting complete and accurate data on children (more step-children) and other kin.

- Family demographic work that has primarily focused on younger families suggests that there may be great heterogeneity in family patterns by race and class in future cohorts of older individuals. Highly educated EBB and WBB cohort members tended to delay marriage and parenthood and had more stability in their marriages than their less educated peers who began childbearing earlier, often outside marriage, with considerable family instability. High rates of non-marriage and nonmarital fertility characterize large segments of the African American EBB and WBB cohorts, and to a lesser extent, Hispanic members of these cohorts. Understanding health disparities at older ages may well require understanding earlier life patterns, given findings from existing studies with the HRS that family disruption negatively affects the flow of care and support between the generations (Pezzin and Shone, 1999).

An important issue going forward for the HRS is to assess where the current design remains adequate for capturing the complexity and consequences of these family changes and where experimentation and modification in data collection on the family may be needed to adequately assess the health and well-being of older Americans as they age.

Marriage, Marital Dissolution and Couple Decision Making

The HRS has been a rich source for studies of joint labor force participation of couples and also the coordination of retirement of married couples. Out of curiosity, I used Google Scholar to get citation counts on some of the articles in the HRS bibliography. For example, David Blau's (1998) *Journal of Labor Economics* article on the "Labor Force Dynamics of Older Married Couples" has been cited 339 times. This

suggests high impact of the work that has been done with the HRS on couple decision making surrounding employment and retirement. A relatively large number of articles in the bibliography have looked at the joint health status and outcomes of married couples, using the HRS.

One thing that was suggested in the 2002 evaluation was to increase the HRS content on marital quality and relationship quality in general (see especially Bumpass (2002) and Ryff (2002)). It is clear that a major effort was undertaken in 2004 and 2006 to augment psychosocial data collection (following suggestions in the Ryff (2002) memo). What is not so clear is whether content on relationship quality will make it into the HRS in an ongoing basis. Without more content on relationship quality and these psycho-social dimensions, many social scientists other than economists and demographers, will not invest in using the HRS. An amazing number of papers on couples (and on intergenerational parent-child ties) still use the first two waves of the NSFH, despite the fact that these data are now over two decades old! The rich content on relationship quality and time allocations by both members of a couple (e.g. to housework) keeps the NSFH two-wave panel data in use and no other data collection – HRS nor PSID – has added sufficient content to compete.

HRS has taken a large step in this direction with the 2004 and 2006 Psychosocial Leave-Behind Participant Lifestyle Questionnaires. I easily found information on the content of the questionnaires in the useful guide by Clarke et al. (2008) but I did not find it easy to locate research or evaluation articles based on these data and I do not think their availability is widely known in the family research community.

What content needs to be collected from both members of a couple, versus from one member who reports about both, is a design feature that also may need to be revisited in light of the increased family complexity noted above. It may be important to engage in new experimental work on when proxy reporting can be used with confidence and when it cannot. Much of what we know in the HRS about family transfers and family members is from the wife's perspective in married couples.

Siblings (and Parents) of the HRS Respondent

Larry Bumpass (2002) noted in his review of the family data in 2002 that the extensive data on HRS siblings had not been used much. This seems to still be the case and it also seems likely that, as Bumpass suggested, this reflects at least in part the decisions made at the start of the HRS about how and when to collect sibling data.

Soldo and Hill (1995) state that in developing the family transfer data for the 1992 data collection on the HRS cohort, priority was given to collecting the information that might impinge on labor supply. Hence, the HRS did not go the route of the Wisconsin Longitudinal Study (WLS) or the NSFH in asking questions about transfers of time to provide things like home maintenance and repairs, or assistance with transportation, or emotional support – the thinking being that these were often things that were given when one “had time” and also could be scheduled for non-work time. Soldo and Hill (1995) also state that, based on PSID analysis of what children gave to parents, children did not give much and hence, the decision was made not to ask about time transfers from children to HRS respondents (e.g., time received). Instead, time the HRS respondent gave to select others was captured, again with an eye toward what types of assistance

might interfere with paid work. Data collection on time transfers was limited to childcare for grandchildren and help to disabled parents. Because help to disabled parents might be shared with siblings, information on siblings was ascertained when at least one of the HRS respondent's parents was alive.

These decisions made sense in light of the study beginning in 1992 with an initial HRS cohort then age 51-61 (born 1931-41). However, it seems likely that older members of this initial cohort had a higher likelihood of no living parents and thus no information on siblings. Also, the ineligibility for collection of sibling data was not random but correlated with mortality in the HRS respondent's parental generation. Thus, it seems likely, given the correlation between mortality and socioeconomic status, that highly educated HRS respondents are more likely to have sibling information than less educated respondents. If correct, this likely dampens usage of the sibling information except in limited cases where the focus of analysis is on transfers to parents.

What may have been a very reasonable and practical decision for the initial interview of mostly healthy 51-61 year olds in 1992 becomes less justifiable as the HRS moves forward in time and becomes a longitudinal panel study of older Americans. By the 2010 interview, the HRS cohort is now age 69-79, much more likely to be needing care than giving care than in 1992 (when they were 51-61), and much more likely to be widowed. Apart from their children, their closest relatives who might be providing assistance (or to whom they are likely giving assistance) are their siblings. Hence, as the HRS cohort ages, it becomes more important than initially to have full information on siblings and on siblings' health status. Whereas parents may have been the ones whose need for care impinged most on HRS respondents when they were age 51-61, at ages 69-

79 this has likely shifted to siblings who need care (or are available to assist with care and companionship, particularly for older women). We do not know very much about siblings in later life such as when they cooperate, or remain in contact, or how correlated outcomes are for them. HRS is potentially the best vehicle to explore these topic. Also, sibsize is changing – the EBB and WBB cohorts will have more siblings than HRS and CODA cohorts – and more siblings than future cohorts. By not getting full information on siblings, the HRS is missing an opportunity to fill an important gap in the family and caregiving literature and missing an opportunity to contribute to the emerging research on the “dynastic” correlations in health and well-being more generally.

One final comment about the decision to focus on the types of care that interfere with labor supply – i.e., the intensive amounts of caregiving to parents and grandchildren. As a cohort ages, time spent doing things for others can be thought of in the broader sense of connecting older people to kin and friends in the community and as a potential barrier to social isolation and associated (mental) health decline. Hence, there may be great value in revisiting the kinds of questions (and research) on assistance to family and friends with questions asked in surveys like the NSFH and WLS. The goal might be to experiment with whether the inclusion of a broader set of “helping” questions might not only enhance content of interest to family researchers but also enhance understanding of later life social support and “social connectedness” that might be related to health and well-being.

Children, Parents and Intergenerational Transfers

As noted above, one of the unique features of the HRS is that it gathers information about all children. These data have been a challenge to use because they have not been organized in a user friendly format and are not currently part of the data files that RAND distributes (though I am told that plans are underway to produce user friendly family files and incorporate these data into the RAND data distribution system by next year).⁴

As with the HRS research on couples, there are a number of widely cited papers that use data on parents and adult children in the HRS. For example, Kathleen McGarry's (1999) research on inter vivos transfers and bequests (published in the *Journal of Public Economics*) has been cited 170 times according to Google Scholar. Her paper with Bob Schoeni (1995) on transfers in the *Journal of Human Resources* has been cited 250 times.

Work that examines the relationship between marital disruption in the parent generation and transfers to and from children, such as that of Pezzin and Shone (1999) in *Demography*, also has high citation counts (cited 96 times according to Google Scholar). This work shows that divorce interrupts the flow of support from children to their elderly fathers. If anything, remarriage exacerbates this situation. Parents also favor biological children over step-children and children give more support to biological parents than step-parents. These are important findings that need to be replicated and updated with cohorts where marital disruption was more common (the EBB and WBB cohorts just now coming into the HRS).

One observation in 2002 was that relatively few papers used the HRS longitudinally to study families. With longer panels, this is changing. For example,

⁴ Personal communication with Kathleen McGarry.

Zissimopoulos and Smith (2009) use the full panel of the HRS and information on the children of HRS respondents to estimate trends and inequality in parental transfers of money to children. They find that parental giving to children is more unequal in larger than smaller families but that overtime, giving tends to equal out across children. Hence the length of the observation period matters. The HRS is beginning to have sufficiently long coverage – up to 8 waves are used by Zissimopoulos and Smith for 16 years of lifespan coverage – to begin to address important questions of intergenerational transfers over larger portions of the older life span.

Altonji & Villanueva (2003) use the HRS to estimate the percent of parents' wealth that will be passed on to children. For these estimates, they require only the number of children – not the full detail on children that is captured in the HRS.

However, one could envision extensions of this work that utilized characteristics of the children. Both the Altonji and Villanueva (2003) and the Zissimopoulos and Smith (2009) studies suggest that the direct transfer of income to children is a relatively small component of the intergenerational correlation in well-being, with the direct transfers for education perhaps the most important transfer that parents make. These conclusions need to be tested against similar analyses from panel data from the PSID, where children are observed over their early lives when educational investments are made, and with more recent cohorts of the HRS as the number of interviews accumulate. With the EBB and WBB cohorts, there is a much greater incidence of delayed childbearing than for the CODA or HRS cohorts – and a greater incidence of delayed nestleaving on the part of children. There has also been more family disruption and hence the role of parental transfers to children needs much further interrogation. This requires continuing to collect

information on all children, making these data more accessible and easier to use, and, if possible, enhancing data collection to make sure that information about step-children is well-captured.

Aykan (2003), using two waves of the AHEAD cohort, finds that childlessness is associated with increased risk of nursing home entry. This is a particularly important finding, given that the EBB cohort begun in 2004 (and the MBB cohort scheduled for data collection in 2010) will have much higher rates of childlessness than the (youngest) AHEAD, CODA and (oldest) HRS cohorts. To assess the relative risks associated with being childless, requires information for comparison on those with children of varying family sizes, including what those children do for their parents.

An indication of the growing interest in studying transfers within a “dynastic” framework was evident at the recent PAA meetings in Dallas, where the Presidential Address by Robert Mare highlighted this topic and where there were 7 sessions on intergenerational transfers and relationships. Few of the papers used the HRS data but one of the major criticisms of PAA papers that did not was the lack of information on full sibships, something HRS has. Without information on all a parent’s children, it is difficult if not impossible to study inequality in “within-family” transfers.

Kathleen McGarry, Judith Seltzer and I have recently begun to use the data on children and parents in the HRS for a project on geographic mobility and proximity of parents and children later in life. We have had two skilled graduate student programmers working with the sibling data and they appear to be of high quality, though it has taken the knowledge of a veteran HRS user (McGarry) and considerable programming time to organize the data into usable analysis files.

The HRS cohort, the cohort followed the longest, is now moving into the years when the likelihood is increasing that members will be net receivers rather than givers of time/care. That is, over the span of data collection from 1992 to the present, survivors of the initial HRS cohort are moving into the age ranges that characterized their own parents when the study began and when so many questions were asked about what they did for their parents. Given the roster of children, this would be an ideal time to begin to plan to actually interview their children, perhaps for a random subset of the HRS cohort, and using an instrument that is very similar to what was used for HRS respondents when asked about the help they gave their parents. One could envision, for example, tests for demonstration effects – do those HRS respondents who assisted their parents generate greater levels of assistance from their children, other things equal?

Gathering more systematic information on what HRS children report doing for their parents (HRS respondents) seems at least worthy of pilot work in the very near future before the HRS cohort becomes too old and dies out. The HRS cohort is now at ages akin to those of the AHEAD cohort in 1993 and many of the types of analyses originally done with the AHEAD cohort need to be repeated with the HRS cohort. A sampling design that linked interviews with HRS respondents with interviews with their children, as children approached eligible ages for inclusion in the study, is a two-generation model worth thinking about.

Barriers to Entry for Family Scholars

Why is the HRS – and the family data in particular – not more widely used in the family research community? From visiting the HRS website and also the RAND website,

I think the HRS staff has done an extensive amount of work to make the HRS accessible and to take seriously the suggestions from the 2002 review. However, it remains the case that the HRS is a complex data set and growing more complex with the addition of each new cohort. (The mere fact that the HRS data documentation notes that there are 39 files to download, potentially, might scare off more than a handful of potential users!)

Because the HRS is so complex, anyone hoping to use the HRS must make a sizable initial investment in data analysis. It also takes considerable investment to figure how to use the HRS data wisely. Thus, one must likely see several potential research papers – or a dissertation – to motivate use of the data. The sizable number of dissertations on interesting family topics listed in the bibliography on the HRS website is one of the most promising indicators of the (future) value of the HRS family data.

Streamlining the family data and making it even easier to link family members' information across waves seems a high priority, given the considerable effort that has gone into collecting the data. The data are unique and the HRS has now been in existence long enough that there are beginning to be papers making innovative use of the panel data to address topics such as parental investment in children.

My understanding is that this work is underway to improve access to the family data, in collaboration with RAND. Let me just end by noting how important I think this effort is and I look forward to the inclusion of family data in the RAND-HRS system.

References

- Altonji, Joseph G. & Villanueva, Ernesto. 2003. "The Marginal Propensity to Spend on Adult Children." *NBER Working Paper 9811*.
- Aykan, H. 2003. "Effect of Childlessness on Nursing Home and Home Health Care Use" *Journal of Aging and Social Policy*. 15:33-53.
- Bianchi, S., Evans, V.J., Hotz, V.J., McGarry, K. & Seltzer, J. 2007. "An Assessment of Available Data and Data Needs for Studying Intergenerational and Intra-Generational Relationships and Behaviors." *California Center for Population Research On Live Working Paper Series, CCPR-020-07*. Los Angeles, CA: UCLA.
<http://papers.ccpr.ucla.edu/papers/PWP-CCPR-2007-020/PWP-CCPR-2007-020.pdf>
- Blau, D. M. 1998. "Labor Force Dynamics of Older Married Couples" *Journal of Labor Economics*. 16:595-629.
- Bumpass, L. 2002. "HRS/AHEAD Family-Related Research." Review Paper Prepared for the Data Monitoring Committee (October).
<http://hrsonline.isr.umich.edu/index.php?p=show&hfyle=dmc2002>
- Philippa C., Fisher, G., House, J., Smith, J., & Weir, D. 2008. "Guide to Content of the HRS Psychosocial Leave-Behind Participant Lifestyle Questionnaires: 2004 & 2006." Documentation Report v2.0 (December 2008).
- Dye, J. L. (2008). "Fertility of American Women 2006." *Current Population Reports, P20-558*. Washington, DC: U.S. Census Bureau.
- Hayward, M. 2002. "Using the Health and Retirement Survey to Investigate Health Disparities." Review Paper Prepared for the Data Monitoring Committee (October).
<http://hrsonline.isr.umich.edu/index.php?p=show&hfyle=dmc2002>
- McGarry, K. 1999. "Inter Vivos Transfers and Intended Bequests" *Journal of Public Economics*. 73:321-51.
- McGarry, K. & Schoeni, R. F. 1995. "Transfer Behavior in the Health and Retirement Study: Measurement and the Redistribution of Resources within the Family" *Journal of Human Resources*. 30:Suppl. p.S184-226.
- Pezzin, L. E. & Schone, B. S. 1999. "Parental Marital Disruption and Intergenerational Transfers: An Analysis of Lone Elderly Parents and Their Children" *Demography*. 36:287-97.
- Ryff, C. 2002. "HRS Review: Psychosocial Variables." Review Paper Prepared for the Data Monitoring Committee (October).
<http://hrsonline.isr.umich.edu/index.php?p=show&hfyle=dmc2002>

Soldo, B. J. & Hill, M. S. 1995. "Family Structure and Transfer Measures in the Health and Retirement Study: Background and Overview" *Journal of Human Resources*. 30:0 Suppl. p.S108-37.

Zissimopoulos, J. & J.P. Smith. 2009. "Unequal Giving: Monetary Gifts to Children and Across Countries Over Time." RAND Labor and Population Working Paper WR-723 (December).