Enhancing the Quality of Data on the Measurement of Income and Wealth

F. Thomas Juster, Honggao Cao, Mick Couper, Daniel Hill, Michael Hurd, Joseph Lutpon, Michael Perry, and James Smith

Enhancing the Quality of Data on the Measurement of Income and Wealth

F. Thomas Juster, Editor
University of Michigan

Co-authors

Honggao Cao
Wells Fargo

Mick Couper
University of Michigan

Daniel Hill
University of Michigan

Michael Hurd
RAND

Joseph Lupton
Federal Reserve

Mike Perry
University of Michigan

James Smith
RAND

January 2007

Michigan Retirement Research Center
University of Michigan
P.O. Box 1248
Ann Arbor, MI 48104
http://www.mrrc.isr.umich.edu/
(734) 615-0422

Enhancing the Quality of Data on the Measurements of Income and Wealth

F. Thomas Juster, Honggao Cao, Mick Couper, Daniel Hill, Michael Hurd, Joseph Lupton, Michael Perry, and James Smith

Abstract

Over the last decade or so, a substantial effort has gone into the design of a series of methodological investigations aimed at enhancing the quality of survey data on income and wealth. These investigations have largely been conducted at the Survey Research Center at the University of Michigan, and have mainly involved two longitudinal surveys: the Health and Retirement Study (HRS), with a first wave beginning in 1992 and continued thereafter every other year through 2004; and the Assets and Health Dynamics Among the Oldest Old (AHEAD) Study, begun in 1993 and continued in 1995 and 1998, then in every other year through 2006

This paper provides an overview of the main studies and summarizes what has been learned so far. The studies include; a paper by Juster and Smith (Improving the Quality of Economic Data: Lessons from the HRS and AHEAD, JASA, 1997); a paper by Juster, Cao, Perry and Couper (The Effect of Unfolding Brackets on the Quality of Wealth Data in HRS, MRRC Working Paper, WP 2006-113, January 2006); a paper by Hurd, Juster and Smith (Enhancing the Quality of Data on Income: Recent Innovations from the HRS, Journal of Human Resources, Summer 2003); a paper by Juster, Lupton and Cao (Ensuring Time-Series Consistency in Estimates of Income and Wealth, MRRC Working Paper, WP 2002-030, July 2002); a paper by Cao and Juster (Correcting Second-Home Equity in HRS/AHEAD: MRRC Working Paper WP 2004-081, June 2004); and a paper by Rohwedder, Haider and Hurd (RAND Working Paper, 2004).

Authors' Acknowledgements

Table of Contents

Introduction

Over the last decade or so, a substantial effort has gone into the design of a series of methodological investigations aimed at enhancing the quality of survey data on income and wealth. These investigations have largely been conducted at the Survey Research Center at the University of Michigan, and have mainly involved two longitudinal surveys: the Health and Retirement Study (HRS), with a first wave beginning in 1992 and continued thereafter every other year through 2004; and the Assets and Health Dynamics Among the Oldest Old (AHEAD) Study, begun in 1993 and continued in 1995 and 1998, then in every other year through 2004. The HRS and AHEAD studies were merged in 1998. Both HRS and AHEAD studies are currently in the field (2006).

At least 6 identifiable studies have been conducted on the quality of the asset and income data in the HRS and AHEAD datasets. In this paper, we examine these six studies in depth. The issues are:

- the use of unfolding brackets to convert "don't know" (DK) or "refuse" (RF) responses to amount questions into a set of categorical responses containing lower and upper bounds;
- an examination of the entry point bias issue that is associated with the use of unfolding brackets.
- the use of an improved survey module that integrates the measurement of income from assets with measurement of the assets themselves. The result is an estimate of capital income that, while it contains the usual measurement error, no longer contains a substantial bias;
- an attempt to improve the match between the periodicity of income receipt as measured by the survey question and by the actual event;
- the correction of substantial underestimates of assets in experimental measurements in the AHEAD 1993 survey, an experiment that turned out to involve confused wording in the financial asset section of the questionnaire as well as problems in other design features. These issues were first noted in Rohwedder, et al. (2004).
- the correction of underestimates of second-home wealth in AHEAD 1995 and HRS 1996 that resulted from a straight-forward skip-sequence error.

I. The Effect of Unfolding Brackets on the Quality of Wealth Data in HRS

Introduction

A characteristic feature of survey data on household wealth is the high incidence of missing data—roughly one in three respondents who report owning an asset are unable or unwilling to provide an estimate of the exact amount of their holding. A partial solution to that problem is to devise a series of questions that put the respondent's holdings into a quantitative range (less than x, more than x, or what?). These quantitative ranges are called unfolding brackets, and they represent a survey innovation that aims to improve the quality of wealth data by substituting range data for completely missing data.

In this part of the paper, we examine the effect of unfolding brackets on the quality of HRS wealth data. Special attention is given to the impact of unfolding bracket entry points on the distribution of asset holdings. Although there is a small positive relationship between mean asset holdings and entry point, there are many cases where that relationship does not hold. In general, our conclusion is that entry point bias problems are not a major concern in the evaluation of quality in the 1998 HRS wealth data.

Unfolding Brackets: Overview

One of the major innovations of the Health and Retirement Study (HRS) is the addition of an unfolding bracket question sequence for those respondents who own an asset but who are unwilling or unable to provide an estimate of the amount. (See Appendix 1 for the basic structure of an unfolding bracket question sequence.) The unfolding brackets idea originated in the wealth module of the Panel Study of Income Dynamics (PSID) in 1984, when a very short wealth sequence was first asked in an ISR/SRC survey. It turns out that the missing data rate (the R owns an asset but is not willing or able to provide a dollar amount) is very sizeable in both HRS and AHEAD—much larger than had proved to be the case for PSID. The typical missing data

4

rate in the HRS and AHEAD studies is of the order of the low thirty percent, a missing data rate that can be reduced to mainly single digits by using the unfolding bracket question sequence.

If it were the case that respondents who did not or could not provide point estimates of their asset holdings (or of other financial flows) did not behave differently, relative to demographic and other characteristics, than respondents who provided point estimates (continuous data cases), then how the missing data cases are treated would make relatively little difference. There would be no systematic bias associated with respondents placing themselves in an unfolding bracket category rather than reporting an exact data number. But if it turned out that missing data cases had values that were systematically high or low relative to personal characteristics of the respondent, then taking that into account might well make a substantial difference in estimates of the distribution of asset holdings, or in the mean levels of such holdings. It would be quite important to find an estimate of the size of that bias and to correct the data for it. In effect, if the imputation program used to convert missing data to imputed data produces the result that there is no systematic difference between continuous (exact) data and missing data, then the gains from using unfolding brackets would be miniscule. On the other hand, if it turned out that missing data cases were systematically very different than continuous data cases, then developing a proper imputation program that corrects for that bias would be quite important.

Initial exploration of this problem produced the not unexpected result that missing data cases were in fact quite different than continuous data cases, and that the appropriate adjustment would involve a substantial increase in the level of asset holdings. Two early papers made this point clearly. One was a paper by Juster and Smith, published in JASA in 1997, which adopted the strategy of imputing missing data cases by random draws from the bracket category that respondents placed themselves into. That is, if a respondent said that their asset holdings were more than $5000 but less than $50,000, an estimate of the respondent's holdings could be calculated by making a random draw from continuous data cases located in that particular bracket category—in this case, in the category $5000 to $50,000. Roughly the same results were obtained in another study, authored by Hurd and published in the Journal of Risk and Uncertainty in 1999. Both of these studies used data collected in Wave 1 of HRS and Wave 1 of AHEAD; the Hurd study also used Wave 2 HRS data.

Table 1 below shows the results of the imputations from these two studies. The top panel has mean values for each category of HRS 1992 asset holdings, while the second panel has HRS 1992 median values. The third panel has 1994 HRS data. The column labeled "RAND-H" represents work done on the imputation of asset holdings by RAND staff working with Hurd, while the category labeled "RAND-S" represents work done by RAND staff working with Smith.

Looking at the values in Table 1, it is quite clear that there is virtually no difference in the mean or median values for the categories labeled RAND-H and those labeled RAND-S. In all cases bracketed data cases yield a significantly higher mean and median value than continuous data cases, while the RAND-H and RAND-S estimates are essentially identical. The small differences that exist between the RAND-H data and the RAND-S data are probably due to the fact that the work done by Hurd treats Range Card cases as if they were continuous data cases, while the work done by Juster and Smith treats these cases as if they were unfolding bracket cases. This difference in treatment produces slightly higher values for RAND-H than for RAND-S because the Range Cards have substantially more detail in the highest categories than do the unfolding brackets—as a consequence, imputation using random draws is likely to produce a few very high values for the Range Card cases, and thus a higher mean.

Table 1. The Impact of Unfolding Brackets on Estimates of the Level and Distribution of Wealth

A. Mean Values, HRS Wave 1 Data (000)

Asset Component	RAND-H[1]			RAND-S[2]		
	Continuous	Bracket	Bracket Δ	Continuous	Bracket	Bracket Δ
Real Estate	149	219	+70	129	222	+93
Business/Farm	168	294	+126	166	349	+183
IRAs	45	45	0	44	56	+12
Stocks/Mutual Fund	59	73	+14	57	74	+17
Corporate Bonds	48	73	+25	47	68	+21
Ck/Sv/MM Acct	16	21	+5	16	23	+7
CDs/T-bills/Gov sv bd	27	45	+18	27	48	+21
Transportation	13	22	+9	13	18	+5

B: Median Values, HRS Wave 1 Data (000)

Asset Component	RAND-H			RAND-S		
	Continuous	Bracket	Bracket Δ	Continuous	Bracket	Bracket Δ
Real Estate	45	75	+30	42	70	+28
Business/Farm	25	95	+70	24	98	+74
IRAs	20	25	+5	20	30	+10
Stocks/Mutual Fund	18	20	+2	17	23	+6
Corporate Bonds	12	20	+8	14	24	+10
Ck/Sv/MM Acct	5	5	0	5	7	+2
CDs/T-bills/Gov sv bd	8	10	+2	10	10	0
Transportation	7	10	+3	8	10	+2

C: HRS Wave 2 RAND-H Data (000)

Asset Component	Median			Mean		
	Continuous	Bracket	Bracket Δ	Continuous	Bracket	Bracket Δ
Real Estate	50	90	+40	98	229	+131
Business/Farm	55	75	+20	112	197	+82
IRAs	28	30	+2	55	60	+5
Stocks/Mutual Fund	26	25	-1	66	74	+8
Corporate Bonds	20	20	0	69	69	0
Ck/Sv/MM Acct	5	10	+5	16	30	+14
CDs/T-bills/Gov sv bd	8	20	+12	24	64	+40
Transportation	8	10	+2	12	18	+6

---

[1] From Michael D. Hurd, "Anchoring and Acquiescence Bias in Measuring Assets in Household Surveys," *Journal of Risk and Uncertainty*, 1999.
[2] From Juster and Smith, "Improving the Quality of Economic Data: Lessons from the HRS and AHEAD," *JASA*, 1997.

It might be useful to spell out exactly why there are a set of cases derived from Range Cards in a study where the missing data estimates are basically derived from unfolding brackets. The reason that there are Range Card cases in this study is that the original HRS design was based on measures developed for the PSID. In the PSID, housing values are asked about before either assets or income, and missing data on housing values was obtained from Range Cards rather than from unfolding brackets.[3] Since the 1992 HRS survey was a personal interview survey, it was feasible to use a Range Card for missing data cases on house value. Thus the respondent had physical control of the Range Card while the housing section was being administered, and some respondents continued to use the Range Card when the survey shifted to other forms of assets. Of the roughly 30% of cases with missing data that had to be imputed, roughly six percentage points are cases where Range Cards were used rather than the unfolding bracket sequence. The Hurd paper uses these Range Card cases after converting them to continuous data cases (using random draws of continuous data cases falling in each of the specific Range Card categories).

There are other characteristics of the bracket data than need to be taken into account in any imputation process, and these seem to have been handled somewhat differently in the Juster and Smith paper than in the Hurd paper. For example, it is unambiguously clear that missing data cases that represent refusals (REF) are really quite different than missing data cases where respondents say they don't know (DK). One major difference is that REF cases show a different distribution among bracket categories than DK cases, and the imputation process produces substantially higher mean and median values for REF cases than for DK cases.[4] Another major difference is that REF cases typically do not complete the unfolding bracket sequences but continue to refuse, while the DK cases generally go through the unfolding bracket sequence.[5]

---

[3] The Range Card that is used for both the HRS and the PSID consisted of 10 categories denoted by a letter (A through J), with amount categories as follows: A = Less Than $500, B = $500-1000, C = $1001-2500, D = $2501-10,000, E = $10,001-50,000, F = $50,001-250,000, G = $250,001-999,999, H = $1 Million - $9,999,999, I = $10 Million - $100 Million, J = More than $100 Million.

[4] This analysis is based on REF or DK cases where the original response was a DK or REF, but the response to the next (bracket) question was one of the bracket categories. That is, if a DK or REF response was followed by the selection of a bracket category, the imputation was based on a random draw from continuous data cases falling into that bracket category. Cases where the only response is a DK or REF are imputed by selecting a random draw from cases where there is both a DK or REF response and a subsequent bracket selection.

[5] About 40% of REF cases are followed by a bracket response, while about 90% of DK cases are followed by a bracket response.

Unfolding Bracket Bias

In recent years, analysis of the unfolding bracket categories and their relationship to the continuous data category has undergone a substantial change. What has basically taken place is that some researchers have become persuaded that various types of potential biases in the treatment of unfolding bracket cases need to be corrected if the data are to be regarded as unbiased (Hurd, 1999; Soest and Hurd, 2003). The kinds of considerations that these researchers worry about are known as "entry point" or "anchoring" bias, or as "acquiescence" bias. The entry point phenomenon is basically concerned with what difference it makes where the unfolding bracket categories are entered--on the low side (e.g., "is it less than $2500, greater than $2500, or what?"), on the high side (e.g., "is it less than half a million, more than half a million, or what?"), or somewhere in the middle (e.g., "is it less than $125,000, more than $125,000, or what?"). Depending on where the respondent enters into this bracket sequence, entry point bias would mean that the distribution of responses would be shifted toward the initial entry point. That is, if the initial entry point is the lowest possible bracket category, the true distribution of assets will be higher than the imputed distribution because the question sequence will generate a bias in the direction of the entry point.

The second type of bias, acquiescence bias, is associated with a respondent preference to agree with the way the question is framed by the survey designer—e.g., is it more than $25,000? More than $50,000? In this type of question sequence, one possible answer is "yes", and it is widely thought that questions of that type produce biased responses because respondents are more apt to say "yes" than not to say "yes"—a yea-saying bias. We do not examine acquiescence bias in this paper because the question wording was changed in HRS 1996 to a balanced version that eliminated the possibility of acquiescence bias (e.g., is it less than x, more than x, or what?).

There are some characteristics of entry point bias that represent what seem to us puzzling features of the data. The theory underlying the psychology that generates these types of biases is that the way the question is framed will influence the way the question is answered. A number of well known and highly regarded papers by Kahneman and Tversky (e.g., Tversky and Kahneman, 1974, 1981; and Kahneman and Tversky, 1986) examine this framing bias. It must be the case that this type of bias is much more important, and clearly more common, when we

are dealing with questions that the respondent does not or may not know the answer to. For example, it is not difficult to understand why there might be an entry point bias if the survey question was something like: "How many African tribes are there in the continent of Africa?" and if the respondent said "don't know", that question might be followed by one that said: "Are there more than 50 such tribes, less than 50 such tribes, or what?" Since the interviewer, and the respondent, can be presumed to know absolutely nothing about the true number of tribes in the continent of Africa, it would not be surprising if there were substantial bias in favor of producing a number that was close to the number specified in the question, on the grounds that the questionnaire designer knew what was a foolish question and what was not, while the respondent didn't know either and was best off relying on the implicit judgment of the interviewer and the question designer.

But what if the question, as in the case of HRS and AHEAD, has to do with checking, saving, or money market accounts, which the respondent must know quite a lot about, but may not be perfectly certain about the exact amounts in those accounts? It is hard to believe that respondents who say they own checking accounts, saving accounts, or money market accounts, wouldn't know approximately the amount of assets in those accounts—whether the accounts add up to "more than $50,000, less than $50,000, or what?" The major difficulty in answering this question is very likely to be that the respondent doesn't know how to interpret "accounts". Over the last decades or so, there has been a veritable explosion of financial instruments that have an accounts flavor, and a typical respondent who has a large number of such accounts might be unclear about which ones should be counted and which ones should be ignored.[6]

How difficult is it to demonstrate that there really is entry point bias, and that this bias needs to be taken care of before the data can be shown to be an unbiased representation of the true distribution of assets? The idea of entry point bias, as noted above, is that low entry points produce estimates of amounts that are biased downward, high entry points produce estimates of

---

[6] There must be many households where the answer to this question is simple and straightforward and where the entry point makes absolutely no difference. Take a household that owns only a single checking account, has no saving accounts, no money market accounts, and no other assets. Is it really plausible to suppose that it matters whether the first question in the sequence asks whether such an account adds to up to less or more than $1000, the next question asks about less or more than $25,000, and the third question asks about less or more than $125,000? It is hard to see why an estimate of the amount in the respondent's checking account is going to be affected by which of those three numbers ($1000, $25,000, or $125,000) shows up first in the question sequence.

amounts that are biased upward, and entry points in the middle produce estimates that have relatively modest bias. If that were the case, one would expect to find that the mean value of assets of a particular type should show an increase from entry point one (on the low side) and entry point two (in the middle), and there should also be increases in the mean value of assets when moving from entry point two (in the middle) to entry point three (on the high side). That is, entry point bias basically says that the respondent will be moved toward the entry point in responding to any question about assets where the respondent lacks perfect certainty about the amount. Finally, picking an entry point around the mean or median may well give better results than picking an entry point at either end of the distribution.

Entry Point Bias: The Empirical Evidence

There have been enough data generated by a variety of entry point experiments in both the HRS and AHEAD survey designs so that we can look at the actual results of entry point differences. Entry point bias ought to mean that going from entry point one (low) to entry point two (higher than entry point one) would show an increase in the mean, and going from entry point two to entry point three (highest) would also show an increase in mean value. If, on the other hand, entry point bias is not present, we should find that the difference in means between entry points one and two or two and three is basically a random process and is just as likely to show a decrease as an increase.

The data in Appendixes 2 and 3 show the distribution of bracket cases for those who responded DK or REF when asked about the amount of money in the various asset categories. The HRS 1998 sample was used in the analysis. Appendix 2 has counts of households in the various bracket categories, and has a complete set of tabulations for each of the ten net worth components. These include real estate properties, businesses and farms, IRAs, stocks and mutual funds, corporate bonds, checking/savings/money market accounts, government saving bonds/CDs/T-bills, transportation vehicles, other assets, and debts. These tabulations are organized by entry point, which varies from asset to asset and is pre-determined according to an algorithm described in Hill (1999).

Parallel to Appendix 2, Appendix 3 shows the mean values for each bracket category, along with the mean for all the cases corresponding to each entry point and the mean for all the households who responded DK or RF.

The data in Appendix 3 are based on the unweighted means for asset owners. For example, the unweighted means for those who own a real estate asset, and who responded DK when asked about the amount of their real estate asset, is $168,006 for those with a low entry point ($2,500), $205,737 for those with a medium entry point ($125,000), and $238,004 for those with a high entry point ($500,000). The data also show that the mean values of their real estate assets increase going from the low to middle entry point, and from the middle to the high entry point. This pattern shows up for the DK cases, for the REF cases, and for the sum of the two types of cases.

Table 2 below details the incidence of asset increases (+) or decreases (-) for respondents in each of the possible entry points for each of the ten net worth components in the HRS study. DK responses are distinguished from REF responses. Thus, REF respondents showed an increase in Real Estate assets between entry points 1 and 2 for those who refused to give an amount of their Real Estate holdings; these respondents also showed an increase in Real Estate assets between entry points 2 and 3.

Table 2 Increases (+) and Decreases (-) in Mean Asset Values as a Function of Response Bracket Entry Points, Where 1 is the lowest of the Entry Points, 3 is the Highest

	DK		REF		DK, REF	
ASSET:	1-2	2-3	1-2	2-3	Σ+	Σ-
Real Estate	+	+	+	+	4	0
Business/Farm	+	+	-	+	3	1
IRAs	-	+	+	-	2	2
Stocks/Mutual Fnd	+	-	-	+	2	2
Corporate Bonds	-	+	-	-	1	3
Ck/Sv/MM Acct	+	+	+	+	4	0
CDs/T-bills/Gov sav bds	+	-	-	+	2	2
Vehicle	+	-	+	-	2	2
Other Assets	-	+	+	-	2	2
Debts	+	-	-	+	2	2
Σ+	7	6	5	6	24	
Σ-	3	4	5	4		16
Σ++		3		2		5
Σ Other		7		8		15

The summary statistics at the bottom of Table 2 indicate that, of the ten net worth components, increases in the means between entry points one and two or two and three (for DK respondents) can be found in seven or six cases, while decreases show up in three or four cases. For REF cases, increases show up in five of the ten categories between entry points one and two, and in six categories between entry points two and three. What if we ask a somewhat more demanding question—do differences in means between entry points one, two and three follow the pattern where both entry points 1-2 and 2-3 always show increases? In that test, DK cases show up as continuous increases in three of the net worth categories (Real Estate, Business/Farm, and Checking/Saving and Money Market accounts), while the other seven categories do not show continuous increases. For the REF cases, two asset categories show continuous increases (Real Estate and Checking/Saving and Money Market accounts) while eight do not. Of the sum of the DK and REF cases, five show continuous increases as entry points increase, fifteen do not.

The data in Table 3 summarizes the results shown in Appendixes 2 and 3, and examine the consistency of the differences in mean values for the three entry points selected for each of the assets. A strong entry point bias would show up as a consistent increase in the means for each asset as we move from entry point one to entry point two, and from entry point two to entry point three. For example, owners of Real Estate show up as having entry point bias because the mean values show consistent increases from the lowest entry point to the middle point and then to the highest point. Thus the highest entry point (designated as H) also shows the highest mean (designated as 3), and the lowest entry point (L) shows the lowest mean (1). But in IRAs, the lowest entry point (L) shows the highest mean (3).

Table 3 compares the rank order of means, for all net worth components and for the four types of financial assets—stocks and mutual funds, checking/savings/money market accounts, corporate bonds, and CDs/T-bills/government saving bonds—for respondents who entered the bracket sequence from a DK response to the amount question, and the rank order of means for respondents who entered the bracket sequence from a REF response to the amount question.

Table 3. Entry Point Rank Order L, M, H (low, middle, high)
For Asset Owners classified as Don't Know (DK) or Refuse (REF)

	DK Entry Point			REF Entry Point		
	L	M	H	L	M	H
Real Estate	1	2	3	1	2	3
Business/Farm	1	2	3	2	1	3
IRAs	3	1	2	1	3	2
Stocks/Mutual Fd	1	3	2	3	1	2
Corporate Bonds	2	1	3	3	2	1
Ck/Sv/MM Acct	1	2	3	1	2	3
CDs/T-bills/Gov sv bds	1	3	2	2	1	3
Vehicles	1	3	2	1	3	2
Other Assets	3	1	2	1	3	2
Debts	1	3	2	1	2	3
Entry point						
observed	15	21	24	16	20	24
predicted	10	20	30	10	20	30
Observed-predicted	+5	+1	-6	+6	0	-6
Financial Assets						
observed	5	9	10	9	6	9
predicted	4	8	12	4	8	12
Observed-predicted	+1	+1	-2	+5	-2	-3

Note: Financial assets include stocks, or stock mutual funds, checking/savings/money market accounts, corporate bonds, and CDs/T-Bills/government saving bonds.

Overall, these data suggest that entry point bias has some influence on the responses to the asset questions, but the influence is modest and entry point selection may not be a major source of bias. The financial asset patterns, especially those for REF cases, do not show any systematic relation between entry point and mean. While the theory calls for the highest mean to be associated with the highest entry point, and the lowest mean associated with the lowest entry point, the quantitative differences in the entry point patterns for REF cases are effectively zero—summing the rank order values for the lowest and highest entry points shows them to be equal.

The analysis so far has been concerned with relatively crude measures of association—comparisons of means, the direction of change (up or down), and so forth. It seems useful to apply a somewhat more rigorous statistical tests to these data, in order to determine whether any clear cut statistical signals come across from the analysis. For this purpose, we pooled together all the ten types of net worth data for those who either gave a don't know answer to the question or refused to give an answer at all. We estimated a set of simple regression models of asset level on asset type, a don't know/refusal dummy (DK/RF), dummies for two entry point categories, and interactions between DK/RF and entry points. The results (Table 4) suggest that there were no statistically significant differences in the mean value of assets between DK and RF responses, or among different entry point categories.

Results were not drastically different when the models were estimated for each type of net worth component separately (Table 5). Of the ten individual models (nine assets and debt), entry point effects appeared only in the models for checking/savings/money market accounts and debts. In these two models, the lowest entry points were generally associated with low asset values compared to the other entry points. The DK/RF effect showed statistical significance only in the model for debts.

Is there a refinement of the entry point bias model that is more consistent with the data than the original entry point bias model? Several features of the data suggests a useful modification of the original model as it applies to the analysis of asset holdings. These modifications are basically driven by noting the degree of certainty associated with the response patterns.

Table 4. Effects of Entry Point and Missing Value Type on HRS 1998 Asset Holdings In Pooled Data Models

	Baseline Model	Full Model
Don't Know (DK)	-	-2.94 (-0.35)
Low-Entry-Point (L)	-	-7.21 (-0.81)
Middle-Entry-Point (M)	-	1.89 (0.15)
Low-Entry-Point x DK	-	-4.50 (-0.39)
Middle Entry-Point x RF	-	-7.99 (-0.53)
Real Estate	203.84** (9.34)	204.01** (11.10)
Business/Farm	282.55** (12.19)	282.32** (10.15)
IRAs	57.26** (3.04)	56.55** (11.36)
Stocks/Mutual Funds	157.65** (8.71)	156.94** (6.86)
Corporate Bonds	75.50** (3.07)	75.15** (7.44)
Ck/Sv/MM Acct	14.43 (0.86)	13.84** (5.65)
CDs/T-bills/Gov Sv Bnd	36.71 (1.91)	36.15** (10.54)
Vehicles	2.51 (0.14)	2.60 (1.34)
Other Assets	36.30 (1.56)	36.19** (7.36)
Constant	11.80 (0.76)	17.92** (2.56)
Adjusted R^2	.040	.041

Note: The dependent variable was (the imputed asset value)/1000. The omitted (reference) groups were Refusal (RF), High-Entry-Point (H), and Debts. The "cluster" option was used when the models were estimated, with a cluster variable "HHID" + "FSUBHH". In the "Full Model", not all the possible interaction terms were included because of collinearity. The joint effect of entry points was statistically insignificant ($F=.71$). $N=11,723$. t-values in parentheses. **=$p<.01$. *=$p<.05$.

Table 5 Effects of Entry Points and Missing Value Types on HRS 1998 Asset Holdings In Single Asset Models

| | DK (*t*-value) | Entry Point | | Entry Point and DK/RF Interaction | | Joint Effect of Entry Points (*F*-value) | Joint Effect of Entry Points and Interactions (*F*-value) | *N* |
		L (*t*-value)	M (*t*-value)	L*DK (*t*-value)	M*RF (*t*-value)			
Real Estate	-1.07	-.79	-.74	.21	.20	.58	.72	571
Business/Farm	.32	-.01	-.10	-.42	-.29	.01	.33	454
IRAs	-.08	-.51	-.86	.68	1.30	.50	1.01	1191
Stocks/Mut Fund	.03	.16	.78	-.17	-.52	.32	.24	1560
Corporate Bonds	.77	1.48	-.98	-1.40	1.22	1.57	.80	371
Chk/Sav/MM Acct	-.73	-2.74**	-.38	.29	-1.29	3.82*	4.77**	3147
CDs/T-bills/Gov sv bds	-.82	-.74	.61	.34	-1.74	.46	1.03	1053
Vehicles	-.36	-1.30	-1.38	.38	.77	1.80	3.05*	2378
Other Assets	.20	-.33	-.17	.25	.50	.07	.18	444
Debts	-3.24**	-2.46	-.01	1.92	-1.68	3.03*	1.78	554

Note: DK = Don't Know. L = Low-Entry-Point. M = Middle-Entry-Point. The omitted (reference) groups were Refusal (RF), and High-Entry-Point (H). "Joint Effect of Entry Points" denotes an *F*-test that the coefficients on L and M are both zeros. "Joint Effect of Entry Points and Interactions" denotes an *F*-test that the coefficients on L, M, L*DK, M*RF are all zeros. **=$p<.01$. *=$p<.05$.

First, it appears to be the case that holdings of real assets are more consistent with the original entry point model than holdings of financial assets. The reason may be that the market values of real assets (the two most important being Real Estate assets and Business/Farm assets) are subject to more uncertainty than holdings of other assets. The greater uncertainty in turn might be due to the greater market volatility of these assets.

Second, it appears to be the case that REF respondents are much more random in the pattern of their mean asset holdings than DK households. That result is probably explained by the fact that REF respondents are not uncertain about the value of their asset holdings, but are simply unwilling to reveal them. In contrast, DK respondents, almost by definition, are very likely to be uncertain about the value of their holdings.

Next, the asset category of Checking, Saving, and Money Market accounts tends to show asset holding patterns that are consistent with the original entry point model. As noted earlier, the reason may be the uncertainty associated with the definition of "account," which may confuse many respondents who have multiple accounts and are unclear about which ones to include.

Finally, the fact that the debt category shows a significant relation to both the DK variable and the entry point variables may be due to the way in which the debt variable was measured. Each of the asset questions had a potential debt component. The specific asset question was: "If you sold all those and paid off anything you owed on them, about how much would you have?" The specific debt question was: "Aside from any debt that you have already told me about, do you have any outstanding debt?" It would not be surprising if many respondents didn't remember how they handled the asset-linked debt component, with the result that the explicit debt question might be quite unreliable.

In summary, while a visual inspection of the mean assets produced using brackets suggests some entry point bias, the multivariate models show that this bias does not reach traditional levels of statistical significance. In part this is due to the relatively small number of cases for which brackets are used. This suggests that the modest bias associated with entry point is a small component of measurement error, which is dominated by variance rather than bias.

Unfolding Brackets and Data Quality: An Overview
Data quality is an issue of longstanding concern among researchers interested in wealth accumulation (Curtin, Juster, and Morgan 1988; Juster and Smith, 1997; Ferber 1959; Lansing, Ginsberg, and Braaten 1961). Recently, available wealth data have proliferated, as many surveys have incorporated wealth modules into studies whose major objectives were quite different than the measurement of wealth or savings. In this paper we argue that some relatively simple survey extensions may significantly improve the quality of household economic data. The survey extensions are "follow-up brackets" - bracket categories offered to respondents who initially refused or were unable to provide an exact value for their assets or income. Brackets represent partial responses to asset questions and can significantly reduce uncertainty about the actual value.

Applied to wealth modules, these extensions originated in the Panel Study of Income Dynamics (PSID) and were used extensively in the Health and Retirement Study (HRS) and the Asset and Health Dynamics Among the Oldest Old Study (AHEAD). Their value is clearest in surveys with relatively short wealth modules. Although application of this methodology to surveys mainly concerned with wealth risks alienating respondents with an excessive number of follow-up questions, wealth surveys with extensive modules might be able to use brackets successfully by tailoring brackets to a limited number of specific assets or by using them judiciously. Use of follow-up brackets appears to provide a partial remedy to deal with non-ignorable non-response bias, a critical problem with economic survey data. Our estimates indicate that wealth imputations based on this methodology are typically higher by a factor of two compared to conventional "hot-deck" imputations made without these brackets. In the two surveys that we examine extensively, the failure to use brackets understated population estimates of non-housing wealth by 19% among those in their 50s and by 9% among those over 70. The effect of this methodology on behavioral models has yet to be assessed.

Background

Assets are notoriously poorly reported on surveys. Non-response is pervasive, and other evidence (Curtin et al. 1989) suggests that the values may also be reported with errors. Although many prominent surveys have included wealth modules, their quality has been viewed with skepticism, due partly to large numbers of missing values.

Three types of cognitive problems may help explain why missing-data rates are so high for many forms of household wealth. First, the respondent may simply not know the answer to the question, particularly if the answer requires adding together several different accounts or placing a value on hard-to-measure assets like a business. Second, the respondent may have a rough idea of the amount but assumes that the interviewer wants a very precise figure, which the respondent cannot provide. Third, the respondent may refuse to disclose the value of assets because he or she regards it as too personal or intrusive.

These considerations may help explain why some wealth components are subject to higher missing-data rates than others. For example, many individuals are quite inactive investors. They may have a much better idea of the amount in their checking account than in an account

20

reflecting their common stock holdings. These households buy stock infrequently, do not check the price with any regularity, and have only a very general notion of their value. In contrast, households with checking accounts get a monthly statement from banks, which is often used to monitor expenditures. Housing equity offers another interesting contrast. Respondents are more willing to respond to questions about the market value of their homes than to questions about their financial assets, possibly because they may feel that anyone, including the interviewer, is able to make a pretty good guess about how much their quite-visible home is worth.

Survey designers have tried various ways to mitigate the missing data problem in financial variables. One strategy, discussed in the early methodological literature (Ferber 1959; Juster, 1977), was to encourage respondents to reduce missing data by providing exact data from financial records. But records were often inaccessible and almost always incomplete, so additional information was always necessary. Another technique, used extensively in early waves of the Surveys of Consumer Finances (SCF), gives respondents a range card with letters corresponding to quantitative intervals.

The SCF Range Card categories are: A. Less than $500; B. $500-$1,000; C. $1001-$2,500; D. $2,501-$10,000; E. $10,001-$50,000; F. $50,001-$250,000; G. $250,001-$999,9999; H. $1 Million-$9,999,999; I. $10 Million-$100 Million; J. More than $100 Million.

These various methods of mitigating missing-data problems all have pluses and minuses. First, any method of following up "don't know" or "refuse" responses is time-consuming and runs some risk of annoying or badgering the respondent. Second, follow-ups that take the form of range cards can be used effectively only in personal interview surveys. The reason is that while the respondent can look over a range card and select the most appropriate value in a personal interview situation, the respondent in a telephone survey has to listen to a complete description of range card categories being read off over the phone by the interviewer-a procedure that many respondents will try to short cut because they find it annoying. Third, unfolding bracket questions provide a uniform stimulus and are generally easy to answer, but are necessarily limited to placing values into relatively few categories. Finally, failure to probe for exact answers may result in some loss of exact answer data.

The HRS and AHEAD methodology involved two main survey features. First, unfolding brackets (is the amount more than x, less than x, or what, and if more than x, is it more than y, less than y, or what?) placed the respondent's asset into one of a set of categories; second, interviewers were told not to extensively probe "don't know" or "refuse" responses, but rather to proceed to the first question in the unfolding bracket sequence. The design philosophy was that dropping the usual practice of probing for exact answers would shorten the survey and minimize chances of annoying respondents. The loss of data quality resulting from losing some exact answers (either by not probing or by learning to provide ranges rather than exact amounts) would hopefully be smaller than the gain resulting from converting completely missing data into categorical data. In HRS wave 1, the strategy used in the 1984 and 1989 PSID wealth modules was adopted, where unfolding brackets were used for financial assets and debts, but range cards were used for housing assets and were also a possibility (as a range card category volunteered by the respondent) in the financial asset module. In later waves where telephones were the primary interviewing medium (AHEAD 1 and 2, HRS 2 and 3), range cards were not used, and all questions about assets used unfolding brackets.

Missing Values and Data Quality

This section documents the ability of follow-up brackets to limit the effects of initial non-response. Table 6 shows the prevalence of item non-response in the HRS and AHEAD asset modules; exact data non-response is shown in column 3 of this table. Housing yields the lowest non-response rates, with less than 5% of HRS respondents not providing an exact home value and almost twice as many unwilling or unable to specify the size of the mortgage. Missing values are considerably more frequent in the financial and tangible asset categories, often on the order of 30% or more. For example, 1 in 3 HRS business or common stock owners had initial non-responses on the value of their businesses or stocks. In most cases, a larger fraction of AHEAD households than HRS households would not give an exact value to their assets. Among financial asset owners, 32% of AHEAD (28% of HRS) households did not report the exact amount in their checking and savings accounts. In general, item non-response ran about 4-8 percentage points larger in AHEAD than in HRS. Because most AHEAD respondents are at least 70 years old and many are in their 80s, reasonable caution in the face of a stranger, minor forgetfulness, or other mild cognitive problems may account for AHEAD's somewhat higher item non-response rates.

22

Where severe cognitive problems were discovered, the likely outcome was use of a proxy respondent.

Non-response to asset questions is commonplace in all household surveys with wealth modules, and these problems are not unique to HRS and AHEAD. For example, 38% of the owners of common stock did not provide an exact value to the amount question in the 1986 Survey of Income and Program Participation (SIPP); the comparable figure for the 1983 Survey of Consumer Finances (SCF) was 25%. Roughly one-third of respondents in each of these surveys did not provide an exact amount for the value of their businesses.

This picture of large amounts of missing data changes dramatically if the categorical data obtained from unfolding brackets are considered. The value of brackets depends first on whether they induce sufficient numbers of respondents to provide range responses. Some believe that non-respondents to asset questions are hard-nut cases, reluctant for privacy reasons to reveal their asset values. In this common view of non-response as dogmatic refusal, the cost of countering the initial non-response with more probing is thought to be high and the yield in new information low. But our experience with HRS and AHEAD suggests that persuading non-

Table 6. Response Rates (percent of total)[7]

Variable	No Asset (1)	Exact Data Report (2)	Exact Data Missing (3)	Owners Only Range Card (4)	Owners Only Unfold Brackets (5)	Owners Only No Information (6)	(4) + (5)/ 3 (7)
HRS							
House*	28	96	4	1	n/a	3	.25
1st Mortgage	55	92	8	1	n/a	7	.12
Business Equity	82	68	32	5	20	7	.76
Other Real Estate	75	74	26	6	15	5	.81
IRAs & Keoghs	58	73	27	5	14	8	.78
Stocks/Mutual Funds	70	67	33	6	19	9	.73
Corporate Bonds	92	69	31	6	12	13	.43
Ck/Sv/MM Accounts	18	72	28	5	14	8	.78
CDs/ T-bills/Gov sv bd	73	70	30	6	14	10	.68
Vehicles	0	86	14	3	9	2	.83
Other Assets	83	71	29	5	15	8	.62
Other Debts	60	86	14	3	n/a	11	.21
AHEAD							
House*	29	78	22	n/a	20	2	.91
1st Mortgage	89	86	14	n/a	13	2	.87
Other Real Estate	80	74	26	n/a	21	4	.84
Business Equity	95	59	41	n/a	36	5	.88
IRA & Keoghs	83	74	26	n/a	19	7	.73
Stocks/Mutual Funds	79	55	45	n/a	37	8	.82
Corporate Bonds	92	59	41	n/a	31	10	.76
Ck/Sv/MM Acct	24	68	32	n/a	25	7	.78
CDs/T-bills/Gv sv bd	77	62	38	n/a	28	10	.74
Vehicles	0	83	17	n/a	15	2	.88
Other Assets	88	70	30	n/a	25	6	.81
Other Debts	85	86	14	n/a	12	2	.86

*Refers to house or apartment (not ranches, farms, or mobile homes).

respondents to provide bracketed responses is often easy. To illustrate, Table 6 separates missing-data responses on HRS and AHEAD into three subcategories: categorical data obtained from a range card, unfolding brackets, and the residual - cases where the respondent refused to provide any information at all. The proportion of all missing data converted to range card or unfolding bracket responses is shown in the last column, and is often of the order of 70% for HRS and 80% for AHEAD respondents.

Although we cannot know what information might have been obtained by direct probing, both surveys showed a substantial reduction in the amount of completely missing information with the unfolding bracket technique. For example, the brackets converted a 33% item non-response for stocks in HRS to only 9% of cases for which we have no information on value. In many financial asset categories, brackets reduced HRS item non-response (defined as no information) by 75%. Because we have only a partial response to a question and not an exact value, this reduction in item non-response is not the same as eliminating item non-response entirely for these cases. But although knowing that a value lies within some prespecified range does not equal knowing an exact value, it is extremely valuable for imputation.

Table 6 shows that brackets were even more successful in decreasing item non-response in AHEAD. For example, brackets converted a 45% full-item non-response in stock value to only 8% of cases with no information on value. In general, full item non-response (no information on value) in both surveys ends up mostly in single digits after the brackets are offered.

While providing some information about the distribution of asset values, a legitimate concern is whether unfolding brackets reduce the probability of reporting exact data. Unfolding brackets might encourage respondents to avoid the difficult cognitive task of counting up asset values in favor of the simpler one of providing "yes" or "no" answers to various threshold amounts. Although these reactions are plausible, our evidence from these surveys actually goes in the opposite direction. We examined respondents who used unfolding brackets in the early parts of the survey to see whether they were also more likely to use brackets in answering questions in the later part of the survey. The answer is no- in fact, just the reverse is true. For all assets, respondents who use brackets early tended to provide exact responses later. Our speculation is

that respondents learned from the bracket questions that a rough approximation to asset value was of sufficient accuracy, and used that insight to provide exact answer data (often in round numbers) later in the survey.

The HRS and AHEAD survey design also sheds some light on the motivation for non-response. In the initial question sequence, respondents who did not give an asset value were separated into two categories: those who refused to respond [refusals (REF)], and those who said that they did not know [don't knows (DK)]. This is an important distinction, not only for the eventual success rate in converting completely missing data into bracket responses, but also in estimating the distribution of the unknown-asset values. Although some respondents are reluctant to reveal the value of their assets, others may simply be unsure of precise values, an uncertainty that will often be converted to a bracket response, and is therefore extremely valuable for imputation.

Table 7 provides some insight into this issue by listing the distribution of HRS and AHEAD cases originally recorded as "DK" or "REF" on asset questions. Respondents who went completely through the bracket sequence are labeled complete bracket. Those who went partly through the bracket sequence, but refused at some later point, are called partial bracket. Finally, those who refused to respond to any of the bracket questions are labeled DK or REF. Data are shown separately for those who originally responded DK and for those who originally responded REF.

The data show a substantial difference in willingness to provide bracket responses for original DK and REF responses. Almost 90% of initial DK responses provided either complete or partial bracket data; the great majority - typically 80% or more - gave complete bracket information. In contrast, more than half of those initially responding REF on a specific item refused to provide any additional information about that asset; only about 40% on average provided complete bracket information. Perhaps some respondents who are unsure of precise values may initially be polite refusals; these respondents are willing to provide some information about asset values with the follow-up brackets. This marked contrast in the behavior of DK and REF responses suggests that the two need to be handled separately when imputations are being done.

Table 7. Bracket Response Distributions, HRS Data (% of Total)[8]

Asset Type	Don't Know Response			Refusal Response		
	Complete Bracket	Partial Bracket	DK REF	Complete Bracket	Partial Bracket	DK REF
Real estate	84	6	10	40	11	49
Business equity	83	4	14	41	7	53
IRA and Keoghs	82	7	10	42	12	46
Stock/Mutual Fund	82	6	12	34	12	54
Corporate Bonds	71	5	24	19	7	74
Ck/Sv/MM Acct	85	7	8	45	11	46
CD's/T-bills/Gov sv bds	80	10	11	35	8	58
Vehicles	89	3	8	32	3	65
Other	83	5	13	28	5	67

Imputation of Missing Values: Methods

Follow-up bracket questions persuaded many initial non-respondents to provide ranges for their asset values. Without brackets, imputation would treat these converts as if they had the same assets as exact-answer respondents with similar personal attributes. It turns out that for both HRS and AHEAD, the exact-answer cases in wave one are heavily weighted toward the lower end of the asset value distribution, whereas REF and DK cases are weighted more toward the upper end. As one example, just 8% of HRS households giving exact answers had business equity in excess of $500,000, compared to 19 (22) % of those who gave initial don't know (refusal) responses, but who answered the bracket question sequence. In general, based on respondents who eventually used the brackets, REF cases are weighted more toward the upper end of the amount distribution than DK cases. [9]

Bracket Respondents

One way to establish the information value of brackets is to estimate missing values as though

the bracketed data were not available. Accordingly, we imputed values under two assumptions

for respondents who placed their assets within brackets. The first (brackets used) recognizes that

the correct value must lie within self-reported limits and that only respondents with assets within

those limits should serve as potential donors. The second (brackets ignored) uses the

Table 8. The Effect of Ignoring Brackets for Imputation of Missing Non-Housing Values Among Respondents Providing Bracketed Responses[10]

ASSET	HRS Brackets		AHEAD Brackets	
	Used	Ignored	Used	Ignored
Mean Values				
Real Estate	221,676	129,098	146,149	107,472
Business	348,600	165,986	219,580	99,872
IRAs, Keoghs	56,415	44,357	55,110	52,608
Stocks/Mutual Funds	74,736	56,982	104,694	74,866
Corporate Bonds	67,846	47,447	90,208	54,275
Ck/Sv/MM Acct	23,409	16,014	21,648	20,750
CD's/T-bills/Gov sv bnd	47,665	27,253	34,823	39,852
Vehicles	18,079	12,539	6,606	6,141
Other Assets	78,711	41,885	21,671	29,684
Other Debts	-7,118	-8,630	-5,481	-4,949
Average Value	75,647	45,287	45,522	35,593
Median Values				
Real Estate	69,678	42,123	62,840	48,940
Business	98,000	24,260	110,400	15,960
IRAs, Keoghs	30,000	20,080	24,500	23,224
Stocks/Mutual Funds	22,928	17,017	39,340	27,760
Corporate Bonds	24,220	14,340	41,540	31,060
Ck/Sv/MM Acct	6,672	5,000	7,780	6,320
CD's/T-bills/Gov sv bnd	10,000	9,760	13,440	19,400
Vehicles	10,000	7,800	4,272	2,024
Other Assets	20,000	13,980	8,050	9,950
Other Debts	-2,918	-2,544	-2,000	-1,554

---

[10] From Journal of the American Statistical Association. December 1997. Juster and Smith: Improving Economic Data.

conventional procedure - all exact-answer respondents serve as potential donors. In both cases a full list of personal attributes is used in the imputation algorithm. Table 8 shows means and medians (averaged over 25 iterations) for each nonhousing asset. The row labeled "average value" contains the weighted average of individual asset values where the weights are the fraction holding each asset among all bracketed assets.

The quantitative differences produced by these two imputation methods are substantial, especially for HRS households. For example, we estimate a mean HRS business asset of $348,600 when brackets are used, with a standard deviation of that mean across the 25 iterations of $21,546. This estimate is well in excess of the mean business asset of $165,986 when bracket information is ignored. In virtually every case, the differences in means in Table 8 are well in excess of the standard errors of these estimates. Mean HRS nonhousing imputations are 67% higher when brackets are used than when brackets are ignored. The difference from using brackets appears somewhat greater for tangible than for financial assets; our estimate of mean business equity among HRS (AHEAD) respondents is more than $182,000 ($120,000) greater when the brackets are used in imputation. Although not trivial, the bias is considerably smaller in AHEAD; our estimated average asset value using brackets was 29% higher than when brackets were ignored. Because these discrepancies are as great with medians, the higher mean values are not simply the consequence of a few very high values.

There are many plausible reasons for this difference between the HRS and AHEAD surveys. Most important, given the age difference between the samples, is that there are fewer AHEAD respondents with extremely high asset values. Second, relative to their total portfolio, AHEAD respondents have fewer assets in categories, such as business equity, where the bias is particularly large. Finally, HRS respondents use both unfolding brackets and range cards, whereas only unfolding brackets were used in AHEAD. The difference between using and ignoring brackets was larger with range cards. For example, average nonhousing asset values were about 50% higher for those who used unfolding brackets than for exact data responses, compared to about 100% higher for respondents who answered using range cards. The reason may be that range cards contain many more categories than unfolding brackets, especially at very

high asset values. Thus it is possible that the unfolding bracket imputations may still understate respondents' asset values.

Final Non-response Imputations

More accurate estimates of missing data for respondents who gave bracketed responses are only part of the gain from the use of brackets. The indirect benefit is that bracketed respondents provide a more relevant donor pool for final non-response cases. Table 9 lists imputed mean values for all "final non-response cases" using two alternative donor pools. The first, more conventional pool consists of respondents who provided exact answers to asset questions. This pool corresponds to that used by many survey organizations when they conduct their imputations. In contrast, the second pool uses as donors only respondents who gave bracketed

Table 9. Imputation of Mean Non-Housing Values for "Final Non-Responses"[11]

	HRS Donor Pool		AHEAD Donor Pool	
	Exact Data Responses	Bracket Responses	Exact Data Responses	Bracket Responses
Other Real Estate	109,449	226,308	91,108	165,454
Equity in Business	280,105	413,221	43,429	251,780
IRAs, Keoghs	37,554	61,272	46,554	47,555
Stocks/Mutual Funds	63,258	98,055	75,891	153,968
Corporate Bonds	45,681	51,747	50,322	92,842
Ck/Sv/MM Acct	16,823	24,585	20,880	23,571
CD's/T-bills/Gov sv bd	24,805	46,259	38,399	33,204
Vehicles	11,209	20,684	7,774	5,697
Other Assets	51,683	69,992	36,883	20,479
Other Debts	-6,665	-7,170	-6,211	-4,635
Average Value	44,185	72,118	40,297	57,156

---

[11] From Journal of the American Statistical Association. December 1997. Juster and Smith: Improving Economic Data.

Table 10. Value of Assets in Full HRS and AHEAD Samples[12]

	HRS		AHEAD	
	Imputations Using Brackets	Imputations Ignoring Brackets	Imputations Using Brackets	Imputations Ignoring Brackets
Housing	75,459	75,864	66,882	66,705
Non-Housing	162,253	136,904	100,583	91,694
All Assets	237,712	212,768	167,465	158,399

responses. We believe that the latter is more representative of the final non-response cases, because they share an initial reluctance to answer asset questions. If anything, the pool of bracketed respondents will still understate asset values of the final non-responses, who are even more reluctant than bracket respondents to reveal their assets.

Table 9 demonstrates how critical the correct donor pool may be. The value of the average HRS (AHEAD) nonhousing asset is approximately 63 (42) % larger using bracketed responses than exact answer responses as donors. Once again, the largest understatements occur in both surveys in the tangible asset categories (business, farms and other real estate). For example, business equity in HRS is higher by roughly $130,000 if we use the donor pool of unfolding bracket responses instead of the conventional donor pool of exact answer responses.

Complete Sample Imputations

Although brackets make a substantial difference when imputing missing data cases, the impact is obviously attenuated in the full sample, which includes respondents who gave exact answers to asset questions or who did not possess the asset. Table 10 summarizes the effect of using brackets on total asset values for the complete HRS and AHEAD samples. In the full HRS sample, mean non-housing wealth is 19% higher using bracket pools. This approximately $25,000 in additional wealth is equivalent to ignoring all wealth in stocks, mutual funds, and checking, saving, and money market accounts. The size of the discrepancy in the full sample varies with the type of asset. Although there is little difference in housing equity, the use of

---

[12] From Journal of the American Statistical Association. December 1997. Juster and Smith: Improving Economic Data.

brackets increases total business and real estate equity by 37% and total financial assets by 17%. Because they vary systemically with age, these discrepancies may affect our views on such basic questions as the adequacy of saving for future retirement. Table 10 shows that mean non-housing wealth is 9% larger when brackets are used in AHEAD. Although this is a non trivial effect, it is much smaller than the 19% reported for HRS. On the basis of our estimates for these two samples, wealth imputation without brackets may understate by roughly 10% the asset holdings of those in their 50s relative to those 70 and older. Because the relative size of wealth in these two age groups is a critical part of any test of the life-cycle hypothesis, (Deaton, 1992), our results suggest that additional work on the size of the bias across age groups should be conducted.

Extensions

Although unfolding brackets can improve the quality of financial data, research on their optimal design and implementation is just at the beginning stages. These issues are complex and in need of additional research; their potential importance is briefly sketched here.

The HRS and AHEAD unfolding bracket questions all had a common format where the initial bracket question is phrased: "Is it more than x?" But there are alternative ways to phrase the question, with some obvious possibilities being: "Is it x or more?"; or "is it more than x, less than x, or what?" The distinction in these three questions is whether or not the rounded number specified by x is associated with a "yes" or a "no" response (if the question is "more than x," then the rounded number calls for a "no" response), and whether the respondent can indicate that their asset holdings are just about the same amount as the rounded number. Based on analyses of some experimental data from HRS and AHEAD, there is little difference in the "x or more" and "more than x" versions, but the balanced question (is it more than x, less than x, or what) provides a somewhat different distribution of responses, with about 5-10% of respondents reporting that "about equal to x" is the correct answer.

Conclusions

The first part of this paper has investigated survey techniques used in the HRS and AHEAD surveys. These techniques - follow-up bracket responses - reduce the implications of initial non-response to wealth questions and narrow uncertainty about precise asset values. Because initial

levels of item non-response in HRS and AHEAD are similar to those obtained in other household surveys, follow-up brackets may also lower the pervasiveness of complete item non-response in other surveys.

The potential value to other household surveys of follow-up brackets goes beyond simply reducing non-response. Our evidence suggests that missing wealth data involves nonignorable response bias, and that follow-up brackets provide a partial remedy to this problem. For example, our estimates imply that household surveys may distort the age-wealth profile by understating wealth in the preretirement years relative to the postretirement years by 10%. Even if there were no effect on nonignorability, range brackets undoubtedly produce efficiency gains as the size of the imputation error is reduced. One must be careful in extrapolating our results to other household surveys that differ in many ways besides the use of brackets. But we think that our results are strong enough to recommend that multipurpose surveys with relatively short wealth modules try follow-up brackets to mitigate a serious problem of nonignorable non-response. In fact, based largely on the HRS and AHEAD experience, the new 1996 National Longitudinal Survey of Youth has already incorporated an extensive use of brackets in its wealth module.[13]

II. Underestimates of Income from Assets-Part I
Introduction

There has been concern about the reliability of survey estimates of income and wealth ever since such measures began to be collected systematically in the 1940s and 1950s (Sudman and Bradburn (1974); Radner (1982)). Obtaining accurate and unbiased household wealth measures has been problematic due to the reluctance of the extremely wealthy to participate in social science surveys at all, and the widespread prevalence of item non-response to wealth questions in particular. Ironically, using new survey innovations, there has been considerably greater progress in mitigating problems for wealth measurement than for income. For example, given the extreme skew in wealth distributions, the bias resulting from the substantially higher non-response rates among very wealthy households has been dealt with in the various Surveys of Consumer Finances conducted since 1983 by the use of special sampling frames (such as tax files) that

---

[13] Aughinbaugh, Alison; Gardecki, Rosella M., "The Performance of Bracketing Techniques on the Collection of Income Data in the NLSY79", January 2005.

over-sample the super-wealthy.[14] Similarly, the growing use of unfolding bracket techniques to handle missing data problems have resulted in reduced measurement error and lower bias due to non-ignorable item non-response to wealth questions (Juster and Smith (1997)). To date, no parallel progress has been documented for income measurement.

In this part of the paper, we attempt to remedy this situation by evaluating two survey innovations aimed at improving income measurement. These innovations are (1) integrating the question sequences for income and wealth, which may elicit more accurate estimates of income from capital than has been true in the past, and (2) examine changes in the periodicity over which income flows are measured, which may provide a closer match between what the survey respondent knows best and the periodicity contained in survey measurement. These innovations have been introduced into both the Health and Retirement Study (HRS) and the study of Asset and Health Dynamics Among the Oldest Old (AHEAD). Based on the results reported below, the potential return in quality of income measurement from these innovations is substantial.

Bias in Income Reporting

Questions about income rank among the most difficult to answer in household surveys (Sudman and Bradburn (1974); Coder and Scoon-Rogers (1995)). Besides any reluctance respondents may have in revealing information they consider private and sensitive, significant cognitive issues exist that may make it difficult for respondents to accurately report their incomes. Especially when asked about the incomes of other family members, the respondent's knowledge about the actual income amounts may be quite limited. Some incomes are received on an irregular basis so that accuracy of reports may depend on how soon after the last receipt the survey questions are asked. Similarly, the dollar amounts involved may be variable from period to period, or taxes and other expenses may or may not be deducted. Finally, respondents may be asked to report their incomes over a time span that is different than how their incomes are received or remembered. These factors may result both in a significant bias (typically under-reporting) or in mis-reporting or random measurement error.

---

[14] See for example, Kennickel, U.S. Bureau of Labor Statistics, et al.

Table 11 gives some indication of the extent of income under-reporting by comparing Current Population Survey (CPS) estimates of various types of income with external benchmarks. Across all income sources, CPS income reports are 89 percent of the benchmark, indicating an 11 percent under-report on average. However, there exists considerable variation around that average. There is little bias in CPS estimates of wage and salary incomes, which are 98 percent of the benchmark. Social Security Income contains a bit more bias (reported income is about 95 percent of the benchmark), but appears to be less understated than the other major source of retirement income, private pensions. But private pensions may be a case where the benchmark is too high since it includes lump sum withdrawals and rollovers to other accounts such as IRAs and Keoghs. Excluding such lump sum payments places the CPS pension income at about 84 percent of the benchmark (Woods (1996); Schieber (1995)). By far, the most severe under-reporting occurs in interest and dividends, where CPS reports are about half the external benchmarks. Even when these income sources are reported without bias, there remains the problem of substantial measurement error in reports of amounts (Ferber (1966), Moore, Stinson, and Welniak (1997)).

Our research will rely on data from three well known surveys; the Health and Retirement Study (HRS), the Asset and Health Dynamics of the Oldest Old (AHEAD), and the Current Population Surveys (CPS). HRS and AHEAD are both longitudinal surveys with data collected every other year. Both surveys obtained extensive information about the economic situation of the households, including a complete accounting of asset stocks and income flows. In addition to housing equity (with separate detail for the first and second home), assets are separated into the following categories in HRS and AHEAD: other real estate; vehicles; business equity; IRAs or Keoghs; stocks or mutual funds; checking, savings, or money market accounts; CD's, government savings bonds, and treasury bills; other bonds; trusts and estates; other assets; and other debt. Similarly, separate questions were asked in both surveys about a long list of income sources for both the respondent and spouse: wages and salaries, self-employment income, tips and bonuses, unemployment compensation, workers compensation, Social Security income, Supplemental Security Income, private pension income, welfare and disability income, and

Table 11. CPS Income as a Percent of Independent Sources [15]

Wages and Salaries	98.2
Social Security and Railroad Retirement	94.8
Interest	51.3
Dividends	42.9
Net Rents and Royalties	81.3
Private Pensions and Annuities	70.6
All Income	89.2

Derived from Current Population Reports Consumer Income Series P-60. Money Income of Households, Families, and Persons in the United States. Department of Commerce, Bureau of Census. Numbers produced here are averages of Volume No 180 and 184.

veterans' benefits or military pensions. In addition, questions are asked at the household level about rental income, income from business, interest and dividends, annuities, and food stamps.

There are two specific enhancements implemented in HRS and AHEAD aimed at improving the quality of income measurement. These are the integration of questions about income from assets with questions about the assets from which such income is derived, and the use of periodicity questions that for certain income sources more closely match the frequency with which such income is received.

We discuss these enhancements in detail below. HRS and AHEAD income and asset modules are given to the knowledgeable financial respondent- the eligible respondent most knowledgeable about the household's financial situation. Especially in AHEAD, proxy respondents are occasionally used if the financial respondent is not physically able to respond or suffers from severe cognitive problems. Because the integration of asset and income questions took place between the second and third waves of HRS and the first and second waves of AHEAD, across-wave comparisons of reports of income from capital provide a convenient way

[15] From "Enhancing the quality of data on income: recent innovations from the HRS." Michael Hurd, F. Thomas Juster and James P. Smith. In the Journal of Human Resources Summer 2003; 38(3): 758-72

of evaluating the impact of this integration. Since AHEAD did not vary the periodicity of income reporting, on that issue we turn to other surveys for comparison.

The Current Population Surveys (CPS) are the most widely used source to monitor labor force participation and income changes by year in the United States, and thus represent a useful standard of comparison to HRS and AHEAD. CPS conducts interviews each month with the number of households interviewed varying from 47,000 to 57,000 during the 1990s (Current Population Reports). CPS households are interviewed for four successive months, are not interviewed for the next eight months, and then are interviewed once again for four successive months. Annual incomes from a large variety of sources are obtained during the March interview. Consequently, although CPS is normally not thought of as a panel, approximately half the respondents are interviewed about income across two adjacent March interviews.

Since no questions are asked about the value of household assets, the CPS cannot be used to evaluate the merit of integrating asset and income questions. As noted above, CPS does ask questions about a long list of income sources using varying reporting periodicities. CPS questionnaires are typically answered by one household member who may or may not be the most knowledgeable about its financial affairs.

The Measurement of Income from Assets

Table 11 indicated that the most serious under-reporting of income takes place in measures of income from capital. The cleanest case is interest and dividend income, where the underlying sources of the income flows- holdings of common stock, bonds, CDs, checking and savings accounts, money market funds, etc.- are apparently more likely to be reliably reported by the household than the income generated from these assets. A comparison of the fraction of households who report holding an asset and the fraction who report receiving any interest or dividend income from that asset strongly suggests that survey estimates of income from assets are badly underestimated. In the typical survey, the fraction of households reporting interest or dividend income is much smaller than the fraction reporting ownership of assets that might yield an interest or dividend income flow. To illustrate, 75 percent of HRS wave 2 households report holding some financial assets, but less than 30 percent report having any interest or dividend income.

In light of this gross inconsistency in income and asset reports, in the third wave of HRS and the second wave of AHEAD we revised the way some of the income questions were asked. Essentially, we created a merged asset and income module in which questions about particular types of assets were followed immediately by questions about the receipt of any income from that asset. The key to this entire sequence is the way in which income-yielding assets are handled. The question sequence we developed asked first about ownership of the asset; for those households reporting ownership, we then asked about the value of the assets; we next asked whether any income was received from the asset and, if so, about the periodicity and whether or not about the same amount was received every period. For households reporting ownership, value, some income, and a monthly periodicity, with about the same amount received every month, the idea was to calculate last year's income from the periodic amount and the periodicity. For households reporting that the amount received every period wasn't always the same, we branched to a question about the amount of income received from the asset in the prior calendar year. This question sequence was used for the four types of financial assets included on HRS and AHEAD (checking, savings, and money market accounts; CDs, government savings bonds and Treasury Bills; stocks and mutual funds; and bonds), as well as for real estate investment equity and business and farm equity.

Comparison of results from this new way of asking about income from assets (used in HRS 3 and AHEAD 2) with estimates of income from assets produced by the conventional survey methodology (as reflected by HRS 2 and AHEAD 1) show dramatic differences in income amounts reported. Table 12 highlights the impact by listing mean income and the value of asset holdings by source in HRS 2 and 3 and AHEAD 1 and 2. The effects of the integration are

Table 12. Weighted Means of Assets and Income of HRS and AHEAD[16]

Categories	HRS-3	HRS-2	AHEAD-2	AHEAD-1
Asset Values, Four Financial Flows	73,139	56,771	91,929	50,766
Income from Four Financial Flows	3,218	1,502	6,740	2,991
Real Estate Value	49,527	41,700	25,591	24,231
Rental Income	2,592	1,564	1,399	554
Asset Value, Own Business or Farm	22,064	28,839	NA	NA
Income from Own Business or Farm	3,456	2,603	NA	NA
Total Non-housing Asset Values, $	144,730	127,310	117,520	82,010
Total Income from Assets, $	9,266	5,669	8,138	3,545

---

[16] From "Enhancing the quality of data on income: recent innovations from the HRS." Michael Hurd, F. Thomas Juster and James P. Smith. In the Journal of Human Resources Summer 2003; 38(3): 758-72

substantial. Between HRS 2 and HRS 3, income from financial assets, real estate investments, and business and farm equity combined increased from $5,669 a year to $9,266 a year. The percentage growth in income from financial assets was even larger. While the integration of asset and income questions affected all income sources, the impact was largest in income amounts from the four financial assets (a greater than two-fold increase from $1502 to $3218), and smallest in income from business and farm (a 32 percent increase). Following the integration of the asset and income questions from AHEAD-1 and AHEAD-2, income from financial assets increased at an even larger rate for the AHEAD sample (from about $3.5 thousand to over $8 thousand).

The failure to report interest or dividend income using the conventional survey format, while in an absolute sense related to the size of asset holdings, appears to apply throughout the full range of asset holdings. Table 13 provides the relevant data for HRS 2 and 3, dividing the sample into asset categories ranging from none to more than a quarter of a million, and then sub-dividing income into categories starting with none and going up to $25,000 or more. Examine first the relationship between asset holdings and income flows for the sum of the four financial assets contained in the surveys. Ninety percent plus of households in HRS 2 who report a small amount of financial assets ($1-$2499) also report zero interest or dividend income. In contrast, 63 percent of HRS 3 households in the same asset group report zero interest or dividend income.

But the most dramatic results occurs among those with a great deal of these assets. For example, 31 percent of HRS 2 households who had more than $250,000 of financial assets still reported that they received no income at all from these assets. That result is not plausible and indicates that without tying the income questions to the presence and amount of the asset there is a substantial understatement of the prevalence and level of income from assets. The integration of the asset and income questions resulted in a substantial decrease in the inconsistency between asset and income reports. For example, in HRS 3 among those with more than $250,000 in financial assets, only 3 percent did not report any income from this source.

It is not surprising if people with a few dollars of interest or dividend yielding asset holdings report that they had zero interest and dividend income. It is quite surprising that many people

with more than a quarter of a million dollars of financial asset holdings report zero interest or dividend income when the question is asked in the conventional format relative to what they report when the question is asked in the merged format. We believe that the better quality income reports are obtained with the merged format because a respondent has just been asked to think about the existence and size of asset holdings. This merged format makes it difficult to report zero income having just reported substantial asset holdings. Whatever the explanation, the merged income/asset format produces a dramatic improvement in the reporting of income flows from assets.

Table 13. Distribution of Income from Assets[17]

Financial Asset Holdings

| | *A. Interest or Dividend Income from Four Financial Assets* | | | | | | | |
	Total	None	< $50	$50-	$250-	$1K-	$5K-	> $25K
HRS-3								
None	1243	97.2	0.8	0.8	0.9	0.2	0.1	0.0
$1 - 2499	1351	63.1	17.2	11.6	6.5	1.3	0.4	0.0
$2500 - 9999	956	27.0	15.6	28.8	19.6	8.5	0.5	0.1
$10K - 49,999	1520	10.0	6.8	17.6	29.8	32.1	3.6	0.1
$50K - 249,999	1275	6.7	2.0	4.0	8.8	43.2	31.8	3.5
> $250K	371	3.0	0.8	1.1	1.1	16.7	48.8	28.6
Total N	6716	38.2	7.8	11.4	12.7	17.9	9.7	2.3
HRS-2								
None	1322	98.5	0.2	0.5	0.5	0.2	0.2	0.0
$1-2499	1294	91.8	2.1	3.1	1.6	1.4	0.1	0.0
$2500 - 9999	1123	76.6	2.0	8.8	8.5	3.7	0.5	0.0
$10K - 49,999	1703	60.0	1.1	7.0	16.4	12.6	2.5	0.4
$50K - 249,999	1217	43.1	0.9	2.6	10.9	26.9	14.3	1.2
> $250K	278	30.6	0.7	2.5	6.1	15.1	30.9	14.0
Total N	6937	71.8	1.2	4.4	7.9	9.3	4.5	0.9

---

[17] From "Enhancing the quality of data on income: recent innovations from the HRS." Michael Hurd, F. Thomas Juster and James P. Smith. In the Journal of Human Resources Summer 2003; 38(3): 758-72

Real Estate Holdings

	Total	None	< $50	$50-	$250-	$1K-	$5K-	> $25K
				B. Rental Income				

HRS-3

	Total	None	< $50	$50-	$250-	$1K-	$5K-	> $25K
None	5153	99.8	0.0	0.0	0.0	0.1	0.1	0.0
$1 - 2499	22	77.3	0.0	0.0	4.6	13.6	4.6	0.0
$2500 - 9999	123	86.2	0.0	1.6	0.8	7.3	4.1	0.0
$10K - 49,999	483	64.0	0.0	1.2	1.5	20.1	13.0	0.2
$50K - 249,999	641	40.6	0.0	0.3	0.3	16.2	38.1	4.5
> $250K	294	27.9	0.0	0.3	1.0	8.5	29.9	32.3
Total N	6716	88.1	0.0	0.2	0.2	3.7	6.1	1.9

HRS-2

	Total	None	< $50	$50-	$250-	$1K-	$5K-	> $25K
None	5299	95.5	0.0	0.0	0.6	2.3	1.6	0.1
$1-2499	50	88.0	0.0	0.0	0.0	8.0	4.0	0.0
$2500 - 9999	141	90.8	0.0	0.7	2.1	4.3	2.1	0.0
$10K - 49,999	539	73.1	0.0	0.7	2.4	13.2	10.2	0.4
$50K - 249,999	666	51.5	0.0	0.6	1.8	15.3	26.4	4.4
> $250K	242	51.7	0.0	0.0	0.8	5.8	25.6	16.1
Total N	6937	87.8	0.0	0.1	0.9	4.6	5.5	1.1

Business and Farm Assets

	Total	None	< $50	$50-	$250-	$1K-	$5K-	> $25K
				C. Income from Own Business or Farm				

HRS-3

	Total	None	< $50	$50-	$250-	$1K-	$5K-	> $25K
None	5966	98.9	0.0	0.0	0.0	0.3	0.5	0.3
$1-2499	24	33.3	8.3	0.0	0.0	16.7	33.3	8.3
$2500 - 9999	117	31.6	1.7	0.0	2.6	7.7	29.9	26.5
$10K - 49,999	117	32.5	0.0	0.9	3.4	16.2	24.8	22.2
$50K - 249,999	361	33.0	0.0	0.6	3.3	12.5	26.6	24.1
> $250K	131	21.4	0.0	0.0	0.8	3.8	20.6	53.4
Total N	6716	91.3	0.1	0.0	0.3	1.5	3.4	3.5

HRS-2

	Total	None	< $50	$50-	$250-	$1K-	$5K-	> $25K
None	6009	95.9	0.0	0.3	0.4	1.0	1.6	0.8
$1-2499	34	64.7	0.0	0.0	0.0	20.6	11.8	2.9
$2500 - 9999	74	74.3	0.0	1.4	5.4	8.1	10.8	0.0
$10K - 49,999	226	72.6	0.0	0.4	3.1	6.2	9.3	8.4
$50K - 249,999	416	64.7	0.0	1.2	2.2	7.9	15.4	8.7
> $250K	178	57.9	0.0	1.1	2.3	6.7	11.2	20.8
Total N	6937	91.9	0.0	0.4	0.7	1.9	3.0	2.1

There are also some income distribution consequences to the enhanced reporting of income from capital. Income from assets tends to be held by wealthier households so that under-reporting of this income source may simultaneously understate the extent of income inequality in the population. This issue is examined in Table 14 which stratifies households into quintiles by the amount of their total household income in HRS1, and within each quintile, lists the amount of total capital income reported in HRS 2 and HRS3. While HRS 3 numbers indicate that much more capital income is reported in the aggregate, the increased reporting of income from capital had very little impact on those households in the bottom fifth of the income distribution whose income declined relative to incomes in all other quintiles. In contrast, those households in the top quintile registered an increase in capital income of over $7000 between HRS2 and HRS3. In general, the size of the increase in capital income between waves 2 and 3 grew across income quintiles. This pattern implies that the absolute income gap of the well-to-do relative to the poor is understated by conventional survey methods of measuring household income.

The Effect of Income Periodicity

The second survey innovation we evaluate concerns the time span or periodicity over which income is reported. For simplicity, many surveys have respondents report all income sources in the same periodicity even though periodicity and regularity of payments may vary a great deal by source. Yet, especially for income sources which are not variable, respondents may know and answer best if the question refers to the time interval at which they normally and most recently receive that income. (Tourangeau, Rips, and Rasinki (2000)). When respondents are requested to report in a periodicity different than that of usual receipt, we may be asking them to perform quickly some difficult cognitive and computation tasks. The value of a specific periodicity may be highest for income flows that tend to continue indefinitely, to change slowly (perhaps due to a COLA adjustment), and to arrive with uniform periodicity (typically a month).

Given these specifications, the most likely income flows to gain from alternative periodicities may be income sources generally received by older and retired households. The most common source in this category is Social Security benefits, which are received monthly, are adjusted annually for Cost of Living changes, typically do not have taxes withheld, and involve withholding only to the extent that respondents select Medicare Part B as an option (more than

90 percent do). In this case, asking for the amount of last month's Social Security check may produce better estimates of Social Security income than asking, as is the usual case, for Social

Table 14. Weighted Means of Capital Income Flows by HRS-1 Total Household Income Quintiles[18]

HRS-1 Total Household Income		Weighted Means		
Quintile	Mean Value	HRS-2 Capital Income	HRS-3 Capital Income	Change in Capital Income
First	9,886	1,652	2,003	351
Second	25,428	2,107	4,366	2,259
Third	40,762	3,571	5,371	1,800
Fourth	59,660	5,018	10,193	5,175
Fifth	116,397	16,757	23,956	7,199

Table 15. Percentiles of Differences in Annual Social Security Income[16]

	AHEAD	CPS			
	1994-1995	*1992-1993*		*1996-97*	
Percentile			All	Monthly 1[a]	Monthly 2[b]
95	1,563	3,415	3,799	2,682	2,167
90	863	1,965	1,948	1,271	1,134
75	208	545	435	301	256
50	-57	46	-36	-49	-47
25	-263	-405	-540	-369	-310
10	-807	-1,973	-1,921	-1,161	-1,034
5	-1,578	-4,062	-3,956	-2,499	-2,232

a. Based on CPS respondents using monthly reporting intervals.
b. Based on CPS respondents using monthly reporting intervals and after Medicare deduction.

---

[18] From "Enhancing the quality of data on income: recent innovations from the HRS." Michael Hurd, F. Thomas Juster and James P. Smith. In the Journal of Human Resources Summer 2003; 38(3): 758-72.

Security benefits paid during the most recent calendar year. Thus, it seems better to estimate Social Security benefits by asking about last month's Social Security check, multiplying it by twelve for respondents who began to receive Social Security payments prior to the beginning of the most recent calendar year (and multiplying it by the appropriate number of months for households who began to receive payments sometime during the prior calendar year).

Since at least for sub-populations of recipients the truth is known, Social Security may also represent the ideal income source to gauge respondents' ability to report their income accurately. By age 70 when there are no earnings tests or Social Security disability income, Social Security income is fixed legislatively by a formula that depends on the history of past earnings and on family composition. If there are no changes in family composition due to divorce, separation, or death, Social Security income is only revised across calendar years by a universal Cost of Living Adjustment (COLA) first given in the January check each year. To eliminate such demographic reasons for changes in Social Security income, we restricted our AHEAD sample to households where both respondents were at least 70 years old in the first wave and where no marital status changes or deaths occurred between the first and second wave. We also required both respondents to have received some Social Security income in each wave so that there is no ambiguity that we are dealing with program beneficiaries. Finally, cases were deleted when Social Security income was imputed in either wave of the panel.

Given these sample restrictions, Social Security income in our remaining sample should only change due to a COLA. To compare reports of Social Security income across successive waves, we adjusted the wave one report by any COLA that would have taken place given the month and year of interviews. Between waves, most (86.5 percent) AHEAD respondents had two COLA adjustments, but 8.4 percent had only one while 5.1 percent had three. If all respondent reports were completely accurate, these adjusted wave one and actual wave two reports of Social Security income would be identical. Differences between them therefore reflect reporting error.

The first column in Table 15 displays percentile distributions of arithmetic differences in wave one Social Security income (adjusted for subsequent COLA's) and wave two Social Security income. While respondents report monthly incomes, for comparison with other surveys we list differences on an annual basis for the year 1995. The specific year chosen does not affect the

results. The median difference in Social Security income is small-the COLA adjusted wave 1 report is $57 higher per year than the wave 2 report of Social Security income. Half of respondents give reports that are no more than $200 apart, 80 percent give reports within roughly $800 of each other, and 90 percent lie no more than $1,500 (or 23 percent) apart. Reporting errors appear to be symmetric so that each wave is equally likely to be higher than the other.

Are these AHEAD income reporting errors large or small? The answer depends on the context in which the data are used. For cross-sectional analyses since mean Social Security incomes were about $9600 in 1995, Table 15 indicates that AHEAD reporting errors are nine percent or more for one in five respondents. But for analysis relying on the panel nature of the data (within person changes in Social Security income), the problem is far more serious. To illustrate, all within person variation in Social Security income in our sample in Table 15 represents measurement error by construction.

Another way to answer this question is to compare AHEAD income reports to those obtained from other prominent surveys that rely on different methodologies to obtain data on income. The Current Population Surveys (CPS) provide such a comparison. During the 1990s, CPS made several revisions in the way it asks income questions, including Social Security income. Before 1994, CPS respondents were asked to report Social Security income for the last calendar year. Starting in 1994, respondents first selected the periodicity (monthly, quarterly, or annual) in which they wanted to report and then gave a dollar amount for this periodicity. There is a clear preference for a monthly interval for Social Security income. For example, in 1996, 77 percent of CPS respondents selected monthly as the easiest way of reporting Social Security income while 23 percent selected yearly. No matter which periodicity was chosen, the income still referred to the last calendar year. For example, if the respondent chose monthly, they were asked to give their monthly income during an average month last year. CPS staff would then convert all incomes to an annual basis which is the way income is available on public use tapes.

We matched respondents across two successive March panels for 1992 and 1993 (when CPS asked for annual Social Security income) and 1996 and 1997 (when the new CPS reporting system had been in place for a while). Individuals were matched based on their sex, race, age, education, and line number. Matches had to be exact on sex, race, and line number and no more

than two years apart in age and at most one year of schooling apart. We then imposed the same sample deletions used in the AHEAD sample. That is, we retained only cases in which each respondent (and spouse) were at least 70 years old in the first March survey, no deaths or marital changes occurred between March interviews, Social Security incomes were not imputed in either interview, and there was a positive report of Social Security income in both March interviews.

The second and third columns in Table 15 list percentile differences in Social Security income from the second March CPS interview minus the COLA adjusted Social Security income from the previous March CPS. Once again, the median difference was small, less than 50 dollars a year. However, differences in CPS reports of Social Security income are considerably larger than those in AHEAD. For example, the 90th and 10th percentiles in the CPS were about plus and minus $1,900 compared to approximately $800 in AHEAD. Alternatively, roughly one fifth of CPS respondents had measurement errors in their Social Security incomes of 20 percent or more. In general, reporting errors appear to be about twice as large in CPS as in AHEAD. Moreover, the size of these CPS reporting errors seem to be about the same when the new reporting methodology of March of 1996 and 1997 is used as when the old CPS annual income methodology was used in March of 1992 and 1993. Apparently, these revised CPS methods did not lead to any overall improvement in the quality of income reports for Social Security income.

Why then are the quality of AHEAD reports on Social Security income apparently superior to those obtained in CPS? Several factors could produce these differences. For example, CPS does not necessarily interview the most knowledgeable financial respondent, a problem that may be compounded by interviewing someone else other than the older person or his/her spouse. However, when we restricted our analysis to single person households (where there were no options about whom to interview), we found that reporting errors were still about twice as large in CPS as in AHEAD. A more likely explanation for the quality difference is that CPS respondents do not report in the form in which they received their most recent monthly check, which excludes the deduction of the Medicare Part B premium.

To see this, the penultimate column in Table 15 lists differences in CPS Social Security income among those reporting in a monthly interval in both 1996 and 1997. CPS errors in Social Security incomes are much smaller when consistent monthly units reporting is employed. In fact,

47

more than 60 percent of the difference between CPS and AHEAD reporting errors is explained by the use of a monthly interval. The final column in Table 15 indicates some additional quality improvement is obtained by limiting CPS respondents to those reporting in a monthly interval and after Medicare premium deductions in both 1996 and 1997. Much of the remaining difference with AHEAD is likely a consequence of the fact that, even using monthly intervals, CPS is asking respondents to perform the more difficult computational task of calculating what they received in an average month last year while AHEAD is simply asking them to remember the last check. Requiring those respondents who said they found it easier to report in a yearly interval to report monthly instead is likely to result in improved reports as the preference for yearly reporting has little conviction behind it. Even among respondents who reported in a yearly interval in 1996, two-thirds of them reported in a monthly interval one year later.

Conclusion

Although under-reporting of income is often thought to be a problem for those at the bottom of the economic strata, the results presented in this paper indicate that at least for some sources of income it is more of a problem for those at the top of the heap. These income sources include income from financial assets, rental income from property, and income from business or farms. These income sources are understated by a factor of two in conventional household surveys. Fortunately, this appears to be a problem with a solution at hand—integration of asset and income modules in surveys. Such an integration was introduced into the third wave of the Health and Retirement Survey and the second wave of AHEAD. The net result was an almost doubling of these income components as well as a much more consistent reporting by households of their capital income and their assets. Can the benefits of this innovation carry over to other surveys? The merged income/asset module will work best for surveys like PSID, NLS and SIPP which are designed to collect information about asset holdings and about income flows and which have about the same number of asset categories as HRS. But the merged module may work less well in studies like the SCF, which has very detailed asset holdings (roughly 100 categories in all) so that a merger of the income and asset modules may be impractical.

An interesting case involves surveys like CPS that do not currently obtain data on asset holdings in part because data on assets are thought to be sensitive (thereby encouraging refusals) and also to take too much survey time to administer. To deal with these concerns, one possibility is to

experiment randomly with modified versions of the merged income/asset module design that may be less sensitive and less time consuming than the full HRS treatment. One idea would be to ask about the presence or absence of asset holdings, but not about amounts. If assets were present, one would next ask whether there is any income associated with those assets and if so, about the periodicity and amount of income flows. Simply asking about the presence of assets is unlikely to be as sensitive or time consuming as asking about amounts, but may produce some of the data quality benefits of associating income flows with assets. Another possibility is to ask about asset values but only within very broad intervals. Such knowledge may be sufficient to remind respondents of the likely income amounts they receive from these assets. Similarly, asking respondents to answer using a time interval consistent with how income is received significantly improves the quality of reports about income. This is certainly the case with Social Security, where the same amount is received many times in a regular periodicity. The same rationale may hold for many major sources of income. Pension payments are much like Social Security payments, except that some fraction of pension payments will involve tax withholding, and many pensions are not adjusted for Cost of Living changes. But question sequences that ask about tax withholding and about Cost of Living changes should handle this problem quite well. A similar situation is likely to be the case for Veterans' Benefit payments which have the same features as Social Security or Pension payments-once they start, they continue until the death of the recipient, and may continue beyond that depending on demographic circumstances.

III. Underestimates of Income From Assets-Part II

The last decade has seen substantial progress in improving the quality of micro-data on both income and wealth. Some of these developments are documented in recent papers by Juster and Smith (1997), Juster, Lupton, Smith and Stafford (2004), and Hurd, Juster and Smith (2003). These papers explore a number of quality enhancements: the use of unfolding brackets for income or wealth components that convert "don't know" or "refusal" responses into quantitative imputations that contain measurement error but little or no bias; the use of improved estimates of changes over time in wealth and active saving to generate measures of capital gains or losses; the use of a merged questionnaire sequence that integrates survey questions about asset holdings and income flows from these assets to reduce the bias in estimates of income from capital; and finally, matching of the periodicity specified in income questions to the actual periodicity of income receipts as a way to enhance the quality of reports for certain income categories.

These enhancements of survey data on income and wealth, while substantially improving the quality of the cross section data, do not come without a cost. A major problem associated with any change made to the methodology used in a panel survey is that they tend to produce time series inconsistencies. By definition, quality improvements reduce the bias and/or measurement error of the cross section point estimate but, by doing so, introduce a bias in the estimate of change over time.

One way to avoid producing such a time series inconsistency is to freeze the survey technology, thus eliminating any quality enhancement. As a long run strategy, this is clearly a bad idea – robust empirical findings cannot be obtained from poor data. A preferred alternative would be to develop methods of recovering time series consistency in the face of data enhancements. In this section of the paper, we explore methods of recovering time series consistency in the measurement of income from capital in the Health and Retirement Study (HRS) and the AHEAD Study.

This innovation was implemented in the HRS beginning in Wave 3 (1996) and continues to be the methodology used in all following waves including Wave 4 (1998), Wave 5 (2000), Wave 6 (2002), and Wave 7 (2004). Hurd, Juster and Smith (2003) examine the effect of this data collection enhancement and find that the income from capital almost doubles between Waves 2 and 3, suggesting the reduction of a serious bias resulting from the stock/flow separation of asset amounts and income. And as noted above, other surveys, such as the Current Population Survey, also suggest a serious underestimate of income from assets using the conventional survey design that has income from assets reported in one module and the assets reported in a separate module.

Although clearly indicating a substantial improvement in the measurement of asset income, the mean doubling between Waves 2 and 3 of the HRS is problematic for researchers wishing to utilize the panel aspect of the survey. The results of any time series study of HRS asset income will be dominated by this technology change in data collection. To correct the problem we propose a strategy that utilizes the distribution of the rates of return to assets obtained in the unbiased data. Random imputation of asset income rates of return in Waves 1 and 2, using Wave 3 as the donor distribution, are used in conjunction with the asset values of Waves 1 and 2 to generate an estimate of asset income.

Two crucial assumptions are required if this strategy is to be successful. First, it is assumed that although there is a time series inconsistency in the estimate of asset income, the estimates of asset values are not contaminated by this bias. We provide evidence that the measurement of asset values is indeed consistent over time and that the major source of bias in the rate of return to assets stems from the measurement of asset income. Second, the donor distribution must be an adequate representation of the true distribution in the time period where the imputations are being made. To determine how robust our strategy is to this assumption, we provide imputed estimates based on donor distributions coming from HRS Waves 3, 4 and 5. The stability of the imputed estimates across donor distributions is noteworthy.

In the next section, we examine the HRS data on household financial wealth and income flowing from that wealth. We discuss the possibility of various sources of measurement error in the time series across Waves 1 through 5 and provide the mean rate of return to financial assets in these years. We outline three imputation procedures and discuss their relative advantages and disadvantages. These procedures are applied to the HRS asset income data and the results are reported below. The robustness of each procedure is examined as are the various imputation strategies.

Survey Structure Induced Bias in HRS income from Financial Assets

Financial wealth in the HRS is defined as the sum of four components: checking, saving and money market accounts; CD's, government savings bonds and Treasury bills; publicly traded corporate equities and equity mutual funds; and corporate bonds. Each of these potentially yields some amount of asset income. Data from the 1992 and 1994 HRS Waves are based on the conventional survey format while the 1996, 1998 and 2000 Waves are based on the revised format that integrates questions about asset holdings with questions about income from assets. In the conventional format, respondents are asked whether they own any of the four financial assets, and how much they own if they report owning any. In a later section of the questionnaire, respondents are asked about income from a variety of sources (wages or salary, workers compensation, veterans' benefits, business income, rent, Social Security, pensions, interest or dividends, etc.).

In the revised question sequence, households are asked whether they have each of the four

financial asset components noted above. If the respondent claims to own a particular asset, they are asked about the value of their holdings, and if greater than zero, whether they received any dividend or interest income from that asset. If they claim to have received asset income, they are asked how much and how often. Similar question sequences are asked for each of the four types of financial assets.

Gross differences in the reporting of income from assets across the five HRS waves are enormous. These are shown in Table 16. In 1992 and 1994, using conventional methodology, only about a third of the HRS sample reported income from financial assets while almost two-thirds reported zero income from assets. These proportions were approximately reversed in 1996, 1998 and 2000 using the experimental methodology, with almost two-thirds reporting income from assets. Interestingly enough, the proportion of the sample reporting ownership of financial assets is essentially the same, with some upward drift, on all five waves: the fraction owning financial assets is a bit over 80% in 1992, and goes up slightly in each later year as one would expect during a period of vigorous economic expansion. A similar pattern shows up in the AHEAD data in Table 17, where the 1993 wave uses the standard methodology and the 1995, 1998 and 2000 waves use the experimental methodology.

Another way to look at the linkage between assets and income from assets is to examine the proportion of the sample reporting zero income from assets within different asset percentiles across survey years. This comparison is provided in Table 18. In the lowest asset category (zero to the 25th percentile), the proportion of the sample reporting zero income from assets is over 90% in all five survey years, although it is a bit higher in 1992 and 1994 than in 1996, 1998, or 2000. The differences by year become substantial when we look at higher asset percentiles. For example, in the 90th percentile and above, the 1992 and 1994 proportions of households reporting zero income from assets are, respectively, 22% and 36%, extraordinarily high numbers for households in the upper 10% of the financial asset distribution.

Table 16. HRS Financial Assets and Income from these Assets, 1992-2000 (% of Households)

Owns Financial Assets	1992 Income from Assets			1994 Income from Assets			1996 Income from Assets			1998 Income from Assets			2000 Income from Assets		
	Yes	No	Total	Yes	No	Total	Yes	No	Total	Yes	No	Total	Yes	No	Total
Yes	35.7	45.4	81.1	29.2	53.2	82.4	62.3	21.0	83.3	63.0	21.3	84.3	60.1	25.1	85.2
No	0.2	18.8	19.0	0.2	17.4	17.6	0.1	16.7	16.7	0.0	15.8	15.8	0.1	14.8	14.8
Total	35.9	63.2	100.0	29.4	70.6	100.0	62.3	37.7	100.0	63.0	37.1	100.0	60.1	39.9	100.0

Table 17. AHEAD Financial Assets and Income from these Assets, 1993-2002 (% of Households)

Owns Financial Assets	1993 Income from Assets			1995 Income from Assets			1998 Income from Assets			2000 Income from Assets			2002 Income from Assets		
	Yes	No	Total	Yes	No	Total	Yes	No	Total	Yes	No	Total	Yes	No	Total
Yes	36.3	39.2	75.6	65.7	16.9	82.6	62.9	19.1	82.0	58.2	23.5	81.8	57.1	26.4	83.5
No	3.6	20.8	24.4	0.1	17.3	17.4	0.1	17.9	18.0	0	18.2	18.2	0.1	16.4	16.5
Total	39.9	60.1	100.0	65.8	34.2	100.0	63.0	37.0	100.0	58.2	41.7	100.0	57.2	42.8	100.0

Table 18. Percent Reporting Zero Income from Assets, by Asset Percentiles

Year	Percentile Group of Financial Assets					All Households
	[0-24%]	[25-49%]	[50-74%]	[75-89%]	[90+%]	
1992	98.0	79.1	51.5	32.0	26.2	63.2
1994	97.7	81.9	61.1	47.6	36.3	70.6
1996	93.7	42.5	11.1	5.2	4.5	37.7
1998	92.8	43.7	10.8	5.8	1.3	37.1
2000	92.9	45.7	16.3	6.7	2.3	39.9

Integrating the survey questions on asset income into the asset and liabilities module reduces the proportion of households reporting zero income from assets to about 3% in that percentile group. Substantial differences in the fraction of households reporting zero income from assets also show up in other percentile groups. In the 25th -50th percentile, the fraction of households reporting zero income from assets goes from about 80% using the conventional survey format to between 40 and 50% using the revised format. The fraction reporting zero in the $50^{th} – 75^{th}$ percentile goes from over 50% in the conventional mode to about 12% in the revised mode and in the $75^{th} – 90^{th}$ percentile the fraction reporting zero goes from about a third in 1992 and 1994 to around 5% in 1996, 1998 and 2000.

Tables 19A, 19B, and 19C contain a more detailed picture of the change in income from

financial assets and in asset holdings over the five survey years and over the percentile distribution of financial asset holdings. The pattern of the data in these tables is very consistent. In Table 19A, which has mean income from financial asset holdings by percentiles of financial asset holdings, the full sample means in 1992 and 1994 are roughly 50% of the means in 1996, 1998 or 2000. This across year mean difference stems largely from differences among households whose financial asset holdings are in the 75th percentile or higher. For example, in the 90th-100th percentile, mean asset income is about $8,000 in 1992 and 1994, but about $18,000, $22,000 and $23,000 in 1996, 1998 and 2000, respectively – between a two-fold and a three-fold increase. In contrast, in the 50th -75th percentile, the 1996, 1998 and 2000 data look to be about the same size as in the 1992 data, all of which are higher than the 1994 mean.

Table 19B contains mean financial asset holdings across asset holding percentile groups. No pattern difference is evident between the 1992-1994 data and the 1996-1998-2000 data. By year, the mean grows substantially, as one would expect during a period of economic prosperity with substantial capital gains. In the 50th-74th percentile, the mean grows from roughly $15,000 to slightly over $22,500 – a 50% increase over the eight-year period. In the 90th+ percentile group, the mean grows from around $300,000 in 1992 to about $650,000 in 2000 – roughly a two-fold increase. Thus the pattern that one would expect in the absence of any survey innovation is exactly what one finds in Table 19B. Mean financial asset holdings grow steadily and substantially over the 1992 to 2000 period with no indication that the growth rate is affected by the transition from conventional survey methods to the revised method. Generally speaking, the growth rates in assets over the entire period tend to average about 9% per year with growth being larger in the higher percentiles than in the lower ones.

Table 19A: Mean Income from Financial Assets by Percentiles of Financial Asset Holdings (1996 Dollars)

Year	(0-25%)	(25-50%)	(50-75%)	(75-90%)	(90-100%)	All Households
			Financial Asset Percentile			
1992	25	360	1,081	2,882	8,776	1,876
1994	16	311	706	1,883	7,683	1,481
1996	11	143	1,070	4,680	18,451	3,190
1998	6	163	1,057	4,643	22,545	3,740
2000	31	284	1,015	4,889	23,307	4,024

Table 19B: Mean Financial Asset Holdings in Dollars by Percentiles of Financial Asset Holdings (1996 Dollars)

			Financial Asset Percentile				
Year	(0-25%)	(25-50%)	(50-75%)	(75-90%)	(90-100%)	All Households
1992	13	1,960	14,723	62,493	318,749	51,197
1994	34	2,793	19,047	71,070	369,886	60,887
1996	41	2,479	19,335	80,113	454,030	70,656
1998	30	2,190	18,909	85,009	589,991	88,957
2000	54	2,674	22,550	100,480	649,099	100,539

Table 19C: Mean Average Rate of Return by Percentiles of Financial Asset Holdings (Percent)

			Financial Asset Percentile				
Year	(0-25%)	(25-50%)	(50-75%)	(75-90%)	(90-100%)	All Households
1992	1.3	3.3	4.7	4.1	3.3	3.7
1994	0.6	2.3	2.7	2.4	2.7	2.3
1996	3.9	5.0	5.2	5.8	4.6	5.0
1998	3.0	5.0	5.0	4.6	4.5	4.6
2000	2.8	4.5	4.3	4.6	4.1	4.2

Note: Table 19C presents the mean of individual average rates of return, defined as the ratio of financial asset income to financial assets. This requires all households with no financial wealth to be dropped from the sample. In addition, ratios above one are trimmed in the calculation. This drops roughly one percent of the sample in each year with most coming from the first quartile (about 3% dropped in the first quartile).

The effect of the revised survey format conditioned on asset holdings is presented in Table 19C which provides the mean of the average rate of return to financial assets, defined as the ratio of financial asset income to financial assets. Note that this is a mean of individual rates of return rather than the ratio of the means. The mean average rate of return over all households increases by roughly 50% from the conventional format to the revised format. This pattern can be seen across the asset groups as well. For households with financial assets above the 90th percentile, the mean of the average rate of return jumps from 3.3% and 2.7% in 1992 and 1994, respectively, to 4.6% in 1996 after which it stays relatively constant.

The data displayed in Tables 19A, 19B and 19C make it clear that time-series analysis of the effect of income change on various types of behavior would be greatly aided if the income component that reflected the return on financial assets could be adjusted to ensure consistency. The problem is that all datasets using the conventional survey design (asking about a long set of

income components, including dividends and interest income) will seriously underestimate income flows from financial assets and hence overstate the change across the conventional and revised survey years.

There are at least two potentially important ways in which biased measurement error is introduced into reported financial income in 1992 and 1994 – error in reporting having any asset income at all, and error in reporting the value of asset income conditional on having any at all. As indicated in Table 16, a striking features of the quality enhancement in measuring income from capital is that the merged question sequence converts the proportion of respondents who report zero asset income from 71% in Wave 2 to 38% in Wave 3. Even more striking is that the merged module converts the proportion of households with assets above the 90th percentile who reported zero interest or dividend income, from 36% in Wave 2 to 4.5% in Wave 3 (See table 18). Thus, one possibility is that the bias in reported financial income is generated solely by households who actually have but report no asset income. This would imply no bias among households who reported asset income and thus require imputation for only those households who report owning assets but no asset income. If we limit comparisons to households reporting some asset income in each year we might find the same degree of time series consistency that we find in the level of asset holdings from Table 19B. If that were true, we could focus on devices for imputing values to households that reported owning financial assets with no asset income in Waves 1 and 2, based on the relationships observed in wave 3.

To examine this hypothesis, Table 20 reproduces Table 18 for households who report positive income from financial assets. The mean average rate of return in 1992 and 1994 for these households seems much more in line with that from the later waves. However, this masks some remaining time series inconsistencies across the financial asset distribution. The average rate of return for households with financial assets above the 90th percentile, households with by far the most asset income on average (See Table 19A), remains roughly 40% lower in 1992 and 1994 than in 1996, 1998 or 2000. Thus, while the elimination of households who report no asset income alleviates some of the time series inconsistency, it fails to do so for the most relevant households, i.e. households with significant asset income. This is strong evidence against the hypothesis that the only survey induced bias is among households reporting no asset income. The

existence of survey structure induced bias appears to be present both in households reporting positive asset income as well as in those reporting zero asset income.

Imputation Strategy

The average rates of return reported in Table 21 are not only evidence of measurement error in asset income from Waves 1 and 2 of the HRS, but also suggest a possible solution to correcting the problem. As noted above, there is a high degree of consistency in financial wealth across all waves in the HRS. The time series consistency is a product of the fact that the survey instrument did not change over the years. Furthermore, the use of a follow-up sequence of unfolding bracket questions for respondents reporting 'don't know' or 'refuse' in the collection of asset and liability data, combined with random imputation within brackets, greatly minimizes any bias in the measurement of financial wealth. The result is that, while the time series consistency of financial asset income is clearly suspect, the reliability of measured financial wealth is strong.[19] It is thus possible to use the rates of return computed for the 1996 data to assign a rate of return to households in 1992 and 1994. These rates of return can then be combined with the financial wealth data for those households to impute an unbiased measure of financial asset income.

Imputation Results

The imputation procedures used in this part of the paper rely heavily on the distribution of the rate of return in 1996. The central assumption is that the rate of return distributions for Wave 1 and 2 of the HRS are biased downward while the Wave 3 distribution, although not free of measurement error, has no such bias. The rate of return distributions for Waves 1, 2 and 3 are provided in Table 21A, 21B and 21C.

The survey induced bias is clear by comparing the Wave 3 distribution with that of Wave1 and 2. The median rate of return for all households in 1996 is 2.4%. This value is zero for households in 1992 and 1994. Moving up the rate of return distribution, the bias remains. The average rate of return in 1996 is 5.7% at the 75th percentile while being only 3.3% and 1.5% in 1992 and 1994, respectively. Not surprisingly, the average rates of return at a given percentile are smaller for lower values of financial assets, which is largely a result of the fact that the number of

---

[19] Note that unfolding brackets were implemented in the collection of asset income in all waves except Wave 1. This makes the reliability of asset income in Wave 1 even more suspect.

households with zero asset income increases. Households with small amounts of financial assets are more likely to have a portfolio that yields little to no asset income. For households in the lowest asset group, the median rate of return is zero in all years of the survey. Nevertheless, the

Table 20: Mean Average Rate of Return by Percentiles of Financial Asset Holdings, Only Households With Positive Asset Income (Percent)

Year	(0-25%)	(25-50%)	(50-75%)	(75-90%)	(90-100%)	All Households
			Financial Asset Percentile			
1992	14.5	9.0	6.0	4.8	3.2	8.2
1994	11.3	6.0	4.4	5.1	3.3	6.4
1996	9.4	5.9	6.0	5.8	4.6	6.6
1998	9.6	5.9	5.2	4.8	4.6	6.3
2000	8.6	5.3	4.9	4.6	4.3	5.8

Note: Table 20 presents the mean of individual average rates of return, defined as the ratio of financial asset income to financial assets. This requires all households with no financial wealth to be dropped from the sample.

Table 21A: Distribution of Rate of Return to Financial Assets (Percent), HRS Wave I (1992)

Percentile	(0-25%)	(25-50%)	(50-75%)	(75-90%)	(90-100%)	All Households
			Financial Asset Percentile			
5th	0.0	0.0	0.0	0.0	0.0	0.0
10th	0.0	0.0	0.0	0.0	0.0	0.0
25th	0.0	0.0	0.0	0.0	0.0	0.0
50th	0.0	0.0	0.8	1.6	1.7	0.0
75th	0.0	1.7	4.7	5.0	4.5	3.3
90th	2.6	10.8	13.9	10.0	7.4	10.0
95th	44.4	35.7	25.0	16.3	10.0	23.1

Table 21B: Distribution of Rate of Return to Financial Assets (Percent), HRS Wave II (1994)

			Financial Asset Percentile			
Percentile	(0-25%)	(25-50%)	(50-75%)	(75-90%)	(90-100%)	All Households
5th	0.0	0.0	0.0	0.0	0.0	0.0
10th	0.0	0.0	0.0	0.0	0.0	0.0
25th	0.0	0.0	0.0	0.0	0.0	0.0
50th	0.0	0.0	0.0	0.4	0.8	0.0
75th	0.0	0.8	2.4	2.5	2.9	1.5
90th	0.0	8.0	7.1	5.7	5.7	5.8
95th	10.0	16.3	12.5	8.8	8.9	12.0

Table 21C: Distribution of Rate of Return to Financial Assets (Percent), HRS Wave III (1996)

			Financial Asset Percentile			
Percentile	(0-25%)	(25-50%)	(50-75%)	(75-90%)	(90-100%)	All Households
5th	0.0	0.0	0.0	0.0	0.1	0.0
10th	0.0	0.0	0.0	0.6	0.4	0.0
25th	0.0	0.1	1.0	2.0	1.8	0.0
50th	0.0	1.8	3.1	4.0	3.6	2.4
75th	3.0	5.3	6.2	6.2	6.0	5.7
90th	20.0	12.9	12.0	11.9	9.8	12.1
95th	60.0	21.8	17.0	16.9	14.1	21.5

Table 22: Mean Income from Financial Assets by Imputation Method (1996 dollars)

				Financial Asset Percentile			
Data Year	Imputation Method	(0-25%)	(25-50%)	(50-75%)	(75-90%)	(90-100%)	All Households
1992	None	25	360	1,081	2,882	8,776	1,876
	(A)	11	565	1,446	4,011	11,675	2,543
	(B)	2	272	734	3,745	15,306	2,633
	(C)	19	202	958	4,443	18,901	2,886
1994	None	16	311	706	1,883	7,683	1,481
	(A)	11	524	1,293	3,622	12,960	2,600
	(B)	4	252	996	3,976	17,010	2,984
	(C)	6	240	993	3,693	17,256	2,961
1996	None	11	143	1,070	4,680	18,451	3,190
1998	None	6	163	1,057	4,643	22,545	3,740
2000	None	31	284	1,015	4,889	23,307	4,024

pattern of the bias is consistent. The median rate of return for households with financial assets in the 50th to 75th percentile is 3.1% in 1996. In 1992 and 1994, this value remains at zero.

Conclusion

In this part of the paper we note the substantial effects of asking survey respondents about asset income in a merged asset/income module in which the income question sequences follow directly after the asset sequences rather than being asked in a separate income module. The inability of surveys to ascertain accurate asset income data is certainly a product of this phenomenon. We go on to note that the improvements made by correcting this survey flaw do not come without a cost - the substantial seam problem that exists between the years in which the survey technology is improved. In an attempt to improve cross-year consistency in the financial asset income series of the Health and Retirement Study, we propose a number of imputation strategies that take advantage of the fact that cross-year consistency is maintained in the levels of financial assets.

Using various schemes to impute an average rate of return to households in 1992 and 1994, we are able to establish a time series of financial asset income with consistency similar to that of financial wealth. The strategy that yields the best results is one that combines a household's own portfolio allocation information from later waves of the data with random imputation of rates of return within various financial asset groups where the donor distributions come from the 1996 survey year. These results are notably robust to replacing the 1996 donor distribution with that of either the 1998 or 2000 survey years. A version of this imputation procedure that also accounts for gross outliers in the average rate of return yields a time series of financial asset income that is consistent with macroeconomic trends.

Future work will include correcting the income from privately owned business or farms, and real estate. Income from these two assets shares the same time series inconsistency as the financial asset income examined in this paper since it was also asked in a separate model from the value of the assets. The bias in business, farm and real estate asset returns is more difficult to correct since the rates of return are far more idiosyncratic than they are for financial wealth. Nonetheless, once these issues are adequately resolved, a superior measure of total household income will be made available.

IV. Underestimation of Assets in AHEAD 1993

In the design of the income and asset sections of the AHEAD 1993 survey, a number of experiments were tried both in terms of question sequences and question wording. Not all of the enhancements worked, and some had clearly negative consequences. For example, AHEAD 1993 asked about income before asking about assets, which appears to have resulted in a substantial understatement of the level of financial asset holdings. The apparent reason is the inclusion of the introductory phrase, "Aside from anything that you have already told me about, do you (or your h/w/p) have any holdings of common stock, money market funds, CDs, corporate bonds…" Many respondents apparently took the phrase "Aside from anything you have already told me about" as not referring solely to questions about assets, where the phrase is highly important and needs to be used to avoid duplication, but to apply to the questions in the previous section that asked about income from financial assets. Thus a substantial number of people reported income from financial assets and then said that "aside from…" they did not own any such asset. The result is a severe underestimate of the 1993 levels of financial asset holdings (see Rohwedder, Haider and Hurd 2004).

In addition, there is a moderate underestimate of asset holdings in the form of IRA and Keogh accounts in the AHEAD 1993 survey, probably due to the fact that questions were asked about holdings of multiple IRAs or Keoghs in waves subsequent to 1993, but only a single question was asked in 1993. As a result, reports of holdings of IRAs and Keoghs tend to be biased downward in the 1993 AHEAD survey.

Finally, the value of owned businesses or farms appears to be severely underestimated in AHEAD 1993. This is probably due in part to the omission of farm assets from the question

(which asked about business assets only, not about business or farm assets). However, this is unlikely to be the only explanation. The underestimate appears severe compared to business or farm asset holdings in other waves, and we would have expected that many respondents who owned farms would have reported their farm as a business asset.

One of the consequences associated with ownership underreporting of individual asset components is a clear time-series inconsistency in total assets, which is illustrated in Figure 1 below and documented in more detail in Tables 23 and 24. If we assume that the general tendency of asset holdings is a slow decline as people reach increasingly older age, the asset shortfall in 1993 looks to be of the order of $80,000 to $90,000 per household-a very large number.

Figure 1. Mean Values of the Total Non-Housing Assets in AHEAD

($1,000)

Weighted Results, in 2002 dollars
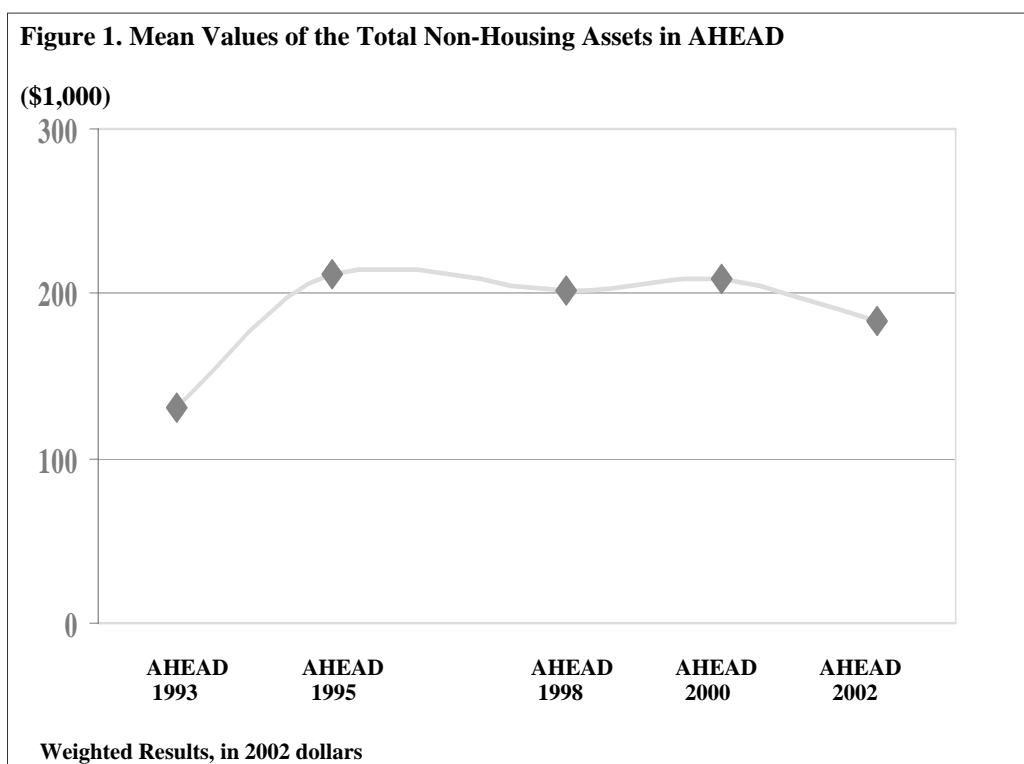
In this section of the paper, we first describe the survey design problems in AHEAD 1993 that we speculate to be the primary cause of the problems with the asset data. We then describe various asset correction procedures. This section concludes with a comparison of the AHEAD 1993 asset data before and after corrections, as well as the use of this corrected asset data to impute financial asset income in AHEAD 1993.

Survey Design Problems in AHEAD 1993

The strategy used for collecting asset data in AHEAD 1993 differed from that used in any other AHEAD wave in two respects. First, unlike as in AHEAD 1995, 1998 and 2000 where questions regarding an asset and the income from the asset were closely aligned in an integrated questionnaire module, asset information was collected separately from income information in AHEAD 1993.[20]

Second, when asset questions were asked in AHEAD 1993, the wording and sequencing of some questions was problematic. Combined with the fact that the asset section followed, rather than preceded, the income section, this sequence appears to have caused a significant number of

---

[20] This problem also exists in HRS 1992 and 1994.

households who owned financial assets to report not owning any such assets. To see how this might have occurred, consider the following asset ownership question for stocks and mutual funds:

> "(Aside from anything you have already told me about,) Do you [or your (husband/Wife/partner)] have any shares of stock in publicly held corporations, or mutual funds?"

While the wording beginning with "Do you" was exactly the same as in later waves, the phrase at the beginning of the question "Aside from anything you have already told me about" was not. This could have encouraged many respondents who owned stocks or mutual funds to give a negative response to this question, on the grounds that information about stock or mutual funds had already been covered in the income section, when respondents were asked whether they were receiving income from various sources-IRA distributions, stocks, bonds, savings accounts, CDs, rental properties, etc. In other words, a respondent who owned stock or mutual funds might think that he/she no longer needed to report the asset because he/she had already (implicitly) reported on it in the income section.

As noted above, two other survey design problems also lead to downward-biased asset measurements in the AHEAD 1993 survey. First, AHEAD 1993 contains only a single question about IRA holdings, while in later waves, the sequence begins by asking how many IRAs the respondent and spouse has and if more than one, about the amount in the largest IRA, then the amount in the second largest , etc. Second, in AHEAD 1993, the question about businesses/farms owned by respondents omits the word "farm," hence potentially missing some assets for respondents who owned a farm.

The impact of these problems on the quality of the asset data may be seen in Tables 23 and 24, where the percentages of asset-owning households and the mean values of the individual asset components are compared across waves. The most conspicuous differences between the AHEAD 1993 asset data and the data in any other wave were significant ownership underreporting in businesses/farms, IRAs, stocks and mutual funds, corporate bonds, checking/savings, money market accounts, and CDs, T-bills, and government bonds. The underreporting appears to apply both to the incidence of ownership and the mean values of assets for owners. The financial assets, in particular, were strongly downward biased. Aggregating the individual asset

components, for example, yields a financial asset total of roughly $137 thousand in the 1995 wave, compared to a total of roughly $71 thousand in the 1993 wave. And the 1993 ownership rates for the financial asset components are fully 10 percentage points lower (out of a total of 30 percentage points) for two of the financial assets, substantially lower for the other two.

Correction Procedure

We used several procedures to try to identify and correct for these asset shortfalls, including corrections for the incidence of ownership and corrections for the mean value of assets, given ownership. We look first at corrections for the incidence of ownership, then at corrections for mean value, given ownership.

Table 23. AHEAD Asset Ownership

% of HH Owning

Asset Type	AHEAD 1993	AHEAD 1995	AHEAD 1998	AHEAD 2000	AHEAD 2002
Real Estate	18	15	12	13	11
Business/Farm	4	7	6	6	6
IRA	17	20	20	22	22
Stocks/Mutual Fund	20	30	31	33	31
Corporate Bonds	6	9	8	9	9
Ck/Sv/MM Acct	77	84	84	83	87
CD's/T-bills/Gov sv bd	22	32	32	34	32
Vehicle	72	69	68	69	68
Other Assets	10	9	10	10	10
Debt	14	13	12	11	10
Total Non-Housing Assets	91	92	92	93	93
Net Worth	94	96	96	97	96

Table 24. Mean Values of Assets in 2002 Dollars

Asset Type	AHEAD 1993	AHEAD 1995	AHEAD 1998	AHEAD 2000	AHEAD 2002
Real Estate	$25,254	$29,185	$29,804	$31,225	$24,497
Business/Farm	$8,769	$17,883	$17,638	$18,687	$18,821
IRA	$10,582	$15,091	$16,405	$16,875	$15,349
Stocks/Mutual Fund	$30,154	$74,014	$68,171	$73,109	$56,084
Corporate Bonds	$6,403	$13,057	$10,384	$8,264	$11,626
Ck/Sv/MM Acct	$22,953	$28,632	$24,132	$23,661	$26,368
CD's/T-bills/Gov sv bd	$11,376	$21,385	$21,741	$23,101	$20,254
Vehicle	$8,837	$8,013	$7,563	$7,223	$6,961
Other Assets	$3,425	$5,090	$7,270	$6,440	$4,374
Debt	$1,152	$815	$611	$689	$1,029
Total Non-Housing Assets	$127,501	$211,534	$202,495	$208,894	$183,126
Net Worth	$219,860	$311,135	$302,215	$312,375	$292,640

Note: Weighted results in 2002 dollars. The top panel shows the percentage of asset-owning households, the second panel the mean value of the asset holdings. "Total Non-Housing Assets" is equal to the sum of the first nine asset components minus "Debt". "Net Worth" is equal to "Total Non-Housing Assets" + "Housing Equity".

Ownership Corrections

The best way to think about ownership corrections is to recognize that multi-wave ownership patterns have a stability that single-wave patterns may not. Thus three-wave ownership patterns that include or exclude the 1993 wave provide a way to identify the extent of underreporting in the 1993 wave.

Tables 25 and 26 provide the relevant multi-wave comparisons of ownership. Table 25 compares ownership rates for two sets of three-wave data—the 1995, 1998 and 2000 waves, and the 1998, 2000 and 2002 waves. Comparisons are shown for six asset categories- business and farm, IRAs and Keoghs, Stocks and Mutual Funds, Corporate Bonds, Checking, Saving and Money Market Accounts, and CDs, T-Bills, and Government Savings Bonds. Column one shows the proportion of the sample that reports ownership in at least one of the three-waves (1995, 1998 and 2000), column two shows the proportion of the sample reporting ownership in a different set of three-

waves (1998, 2000 and 2002), while column three shows the difference in the two sets of three-wave comparisons.

Table 26 shows the same type of comparison except that the 1993 wave of data is included as one of the three data waves in the analysis. A comparison of the data in Tables 25 and 26 indicates that three-wave asset holding patterns are very consistent if 1993 is one of the waves (holding patterns are clearly lower) while the three-wave pattern is not systematically different if the 1993 wave is excluded (holding patterns are basically random).

Table 25. Percent of Sample Owning Selected Assets, Specified AHEAD Waves

ASSET	1995-98-00	1998-00-02	Δ
Business/Farm	9.49	9.13	+.36
IRA-Keogh	25.55	25.87	-.32
Stocks/Mutual Funds	41.48	41.89	-.41
Corporate Bonds	16.22	15.78	+.44
CK/Sv/MM Accounts	94.37	94.49	-.12
CDs/T-bills/Gov sv bd	49.53	48.26	+1.27

Table 26. Percent of Sample Owning Selected Assets, Specified AHEAD Waves

ASSET	1993-95-98	1995-98-00	Δ
Business/Farm	8.96	9.49	-.53
IRA-Keogh	22.44	25.55	-3.11
Stocks/Mutual Funds	37.52	41.48	-3.96
Corporate Bonds	14.40	16.22	-1.82
CK/Sv/MM Acct	92.78	94.37	-1.59
CDs/T-bills/Gov sv bd	46.37	49.53	-3.16

Imputed Income in the AHEAD sample

The imputation procedure that we employ assumes that the rate of return to financial assets in AHEAD 1995 is unbiased. We then use three somewhat different imputation strategies (labeled 1, 2, and 3) to apply these rates of return to AHEAD 1993 asset measurements. In strategy 1, anyone who reports positive financial income and financial assets in 1993 is left as is, and anyone who reports financial assets but no financial income receives a random rate of return

drawn from the distribution of rates of return in the 1993 data (classified by financial asset percentile group, defined as those in percentiles 0-25, 25-50, 50-75, 75-90 and 90-100). In strategy 2, anyone with financial assets in 1993 receives a random rate of return imputation.

Strategy 3 is the same as 2, except that those individuals who own financial assets in both 1993 and 1995, and for whom financial assets are in the same asset percentile group in both years, get their 1995 rate of return applied to their 1993 assets. Anyone who has financial assets, but no financial income in 1995, receives a rate of return of zero in 1993. In all three strategies, anyone with no financial assets in 1993 gets a rate of return of zero.

The results are shown in Table 27. Average AHEAD income from financial assets is $3,959 in 1993 as compared to approximately $6000 in the next three waves. Strategies 1, 2 and 3 result in estimates of $5179, $4969 and $4720, respectively, in the 1993 wave.

Table 27.Corrected Financial Income in AHEAD by Financial Asset Percentile Group

Year	Imputation Strategy	Financial Asset Percentile					
		0-25	25-50	50-75	75-90	90-100	total
1993	none	1,218	1,062	3,117	6,116	16,146	3,959
	1	126	1,686	5,003	8,560	23,614	5,179
	2	62	1,062	3,817	8,853	25,768	4,969
	3	57	1,010	3,873	8,585	22,575	4,720
1995	none	55	776	3,821	9,643	32,210	6,138
1998	none	55	611	3,224	10,812	35,627	6,453
2000	none	85	579	2,692	9,997	33,304	5,960
2002	none	43	661	2,297	6,753	24,778	4,449

Note: Weighted results in 2002 dollars.

Correction of Mean Values for Owners

Correction of the data on ownership status to eliminate the ownership shortfall, however useful it may be, is not guaranteed to eliminate the total ownership bias, since correcting for the ownership shortfall will not eliminate any bias that takes the form of an underestimate of value, given ownership. And the design of the 1993 AHEAD survey does appear to leave room for this sort of bias. For example, the question sequence about financial assets and income from financial

assets may well leave room for underestimating value if some assets in a particular category yield income but others do not - e.g., some stocks pay dividends but others do not. Given the 1993 survey design, some respondents might judge that they have already told the IWER about the value of assets for assets that yield income, but had not said anything about assets that do not – e.g., the value of stocks that pay no dividends. Some respondents might easily report asset ownership but a lower asset amount than respondents who did not make a distinction between assets that yield income and assets that do not.

A careful examination of the data for the AHEAD sample suggests that the mean value of ownership, for households who report that they own an asset, is relatively low for comparisons involving the 1993 AHEAD sample. Table 28 summarizes differences in mean values for AHEAD owners in each asset category in successive pairs of years - 1995 and 1993, 1998 and 1995, 2000 and 1998, 2002 and 2000, and 2004 and 2002. For the most part, the mean differences are unrelated to the year of interview, e.g., the mean holdings for owners in 1998 minus the means for owners in 1995 is a plus for 6 asset categories and a minus for 3 categories; the difference between 2000 and 1998 is a plus for 4 asset categories and a minus for 5 categories, etc. It appears that this relationship is about the same for each pair of years except for 1995-1993, where the mean differences are all plus. Thus the data tells us that the 1993 means for AHEAD asset holders of each type is less than the mean for asset holders in other waves.

Table 28: Differences in Mean Value of Holdings for Asset Owners, 1993-2004 (000 dollars)

	1995-1993	1998-1995	2000-1998	2002-2000	2004-2002	Σ +, -
Real Estate	+32.4	+21.7	-22.1	-53.6	+60.2	3,2
Business/Farm	+21.5	+78.4	+31.0	+182.4	-3.7	4,1
IRA	+19.6	+34.6	+3.4	-12.9	-13.4	3,2
Stock, Mutual Fund	+151.2	-26.1	+6.1	+21.5	+10.4	4,1
Corporate Bonds	+22.0	+9.1	-8.4	+8.1	+12.2	4,1
CK, SAV, Money Market Accounts	+10.6	+1.6	-0.5	+2.2	-11.1	3,2
CD's, T-Bills, Government Savings Bonds	+20.5	+14.6	-13.0	-9.2	+3.7	3,2
Transportation	+2.4	-1.4	-0.5	+1.2	+0.3	3,2
Other Assets	+31.8	-3.6	+13.1	-19.9	+8.6	3,2
Σ +	9	6	4	5	6	
Σ -	0	3	5	4	3	

The remaining part of the imputation strategy is to impute new asset values for the six categories of assets whose values appear under-reported. Asset values are only imputed for owners. To do this we employ a prediction method based on ordinary-least-squares. First, we estimate a model of prediction of 1995 asset values (a separate model for each relevant asset category) using information available about the respondents in 1998. Although much information is available about respondents (e.g. demographic data, health status), we find that no significant advantage is gained by using any covariates beyond 1998 asset values. So for each of stocks, bonds, accounts, CDs, IRAs, and businesses we regress the 1995 asset value on those asset values in 1998, plus real estate and first and second home values[21]. We then use the coefficient values (reported in table 26) to predict 1993 asset values for the under-reported assets using 1995 asset values as inputs. We also do the same procedure, using the OLS relationship between 1998 and 2000 instead of that between 1995 and 1998. These are not economically sophisticated strategies, as they assume that whatever patterns governed asset-value transitions between the later survey waves are adequately expressible using a linear combination of asset values, and they assume that those patterns are similar to the true pattern governing asset-value transitions between 1993 and 1995. The primary justification for using this method is that it produces population-wide distributions of assets that appear to be closer to the true distributions than what are currently available for 1993 AHEAD.

Table 29 shows the results of these procedures. It shows the original population means in 1993 and 1995, as well as the new means using the 1995-1998 OLS coefficients (Imputation A) and the 1998-2000 OLS coefficients (Imputation B). In addition to means, it shows standard deviations, and means, medians and the 95th percentile level, the latter three conditional on ownership. Based on the fact that in later waves of AHEAD, mean assets show a slow, steady decline, we might expect the true values of 1993 assets to be slightly above 1995. Neither of the imputation methods consistently produce this result, but both sets do produce mean asset values that are significantly closer to 1995 than the current 1993 values. Aside from the mean values, the other distributional qualities of the imputations are reasonable as compared to 1995. Standard deviations for both sets of imputations are roughly in line with 1995 standard deviations, as are means conditional on ownership.

---

[21] We also attempted a log-log specification for the regressions, but this produced significantly lower R^2 values, and, more importantly, imputation values that were significantly lower than the original 1993 data.

However, it should be noted that there is a significant difference between the distributions of both sets of imputations for all the asset categories and the distribution of the 1995 values. That is, the imputations consistently over-predict the median and under-predict the 95th percentile, conditional on ownership, as compared to 1995. This means that in both sets of imputations, there are too many people with moderately high values for each asset and too few with very high values. So for a given respondent we cannot have much confidence that we have predicted the asset values well. Only in cross-sectional means do the data appear reasonable.

Table 29. Imputation Results

Means	1993			1995
	Original Data	Imputation A	Imputation B	
Stocks/Mutual Funds	23,120	52,128	45,646	56,357
Corporate Bonds	5,092	9,321	9,983	9,486
CK/SAV/MM Accounts	16,660	19,855	18,274	22,268
CDs, T-bills, Gov't Sav Bonds	8,641	16,228	14,851	17,441
Business/Farm	6,001	14,113	16,962	12,264
IRAs	7,366	9,176	10,030	10,734

Standard Deviations	1993			1995
	Original Data	Imputation A	Imputation B	
Stocks/Mutual Funds	180,691	398,513	501,696	634,320
Corporate Bonds	46,152	52,038	78,450	73,773
CK/SAV/MM Accounts	44,578	33,040	18,886	62,916
CDs, T-bills, Gov't Sav Bonds	37,361	42,988	46,712	60,344
Business/Farm	52,904	130,802	90,621	114,982
IRAs	37,529	49,531	40,511	65,986

Means*	1993			1995
	Original Data	Imputation A	Imputation B	
Stocks/Mutual Funds	128,616	227,681	199,367	207,397
Corporate Bonds	97,009	123,295	132,043	113,354
CK/SAV/MM Accounts	22,922	26,035	23,962	27,812
CDs, T-bills, Gov't Sav Bonds	42,385	59,945	54,860	59,842
Business/Farm	191,322	231,705	278,469	226,299
IRAs	51,945	60,223	65,833	65,561

Medians*	1993			1995
	Original Data	Imputation A	Imputation B	
Stocks/Mutual Funds	40,000	125,005	56,173	50,000
Corporate Bonds	31,500	85,698	66,917	40,000
CK/SAV/MM Accounts	7,000	16,222	21,562	9,400
CDs, T-bills, Gov't Sav Bonds	20,000	42,145	39,725	30,000
Business/Farm	100,000	147,648	217,225	100,000
IRAs	23,000	32,124	46,598	26,800

95th Percentile*	1993			1995
	Original Data	Imputation A	Imputation B	
Stocks/Mutual Funds	500,000	648,215	687,188	800,000
Corporate Bonds	630,000	359,044	433,360	450,000
CK/SAV/MM Accounts	100,000	72,886	44,768	118,000
CDs, T-bills, Gov't Sav Bonds	150,000	154,403	133,180	200,000
Business/Farm	650,000	575,796	612,172	650,000
IRAs	250,000	190,354	173,581	247,000

*Conditional on ownership

This last point is corroborated by tables 30, which shows root mean square differences (RMSD) and mean absolute differences (MAD) between the two imputation strategies among all owners of the particular asset. While the MAD have relatively modest magnitude, the RMSD are very large—particularly for stocks, bonds and businesses. These two facts together suggest that for many respondents the two imputation strategies produce similar results, but that for a few, the differences are very large. That would cause the RMSD to be much larger than the MAD.

Table 30: Differences in Imputations

Asset	RMSD*	MAD**
Stocks/Mutual Funds	276,275	23,591
Corporate Bonds	139,654	4193
CK/SAV/MM Accounts	23,843	7547
CDs/T-bills/ Gov't Sav.Bonds	64,376	5699
Business/Farm	415,907	10,231
IRAs	37,853	2869

Table 31 shows the regression coefficients and R^2 values for the regressions that we used to produce each set of imputations. Two points are worth making. First, R^2 values are generally low—varying between 9% and 59%, but mainly around 30%. This means that these covariates do not explain the bulk of the variation in asset values among asset owners. As already pointed out, adding more covariates does not help the situation. Second, the regression coefficients used for imputation A (taking the coefficients from 1995-1998) and imputation B (taking the coefficients from 1998-2000) do not show a great deal of stability across imputation strategies. This is not true of every coefficient—some are similar for each strategy—but it is certainly true of many. This fact helps explain the differences shown in table 31. All of this should lead to reservations about using the 1993 AHEAD data for economic analysis. At the least, the analyst should be well aware of how the imputations were done and their limitations.

Table 31: Regression Coefficients and and R-Squared Values

Dep. Variable	Imp. Strategy	Independent Variable Coefficients									Constant	R-squared
		Stock	Bond	Account	CD	Business	IRA	Real Estate	Home	2nd Home		
Stock	A	0.64	-0.14	0.34	-0.19	-0.19	-0.08	-0.09	0.10	1.55	89,803	18%
Stock	B	0.84	0.73	0.07	0.44	0.03	-0.14	0.01	-0.02	-0.01	12,650	23%
Bond	A	0.02	0.42	0.12	0.16	-0.02	-0.004	-0.02	0.44	0.002	21,144	29%
Bond	B	0.07	0.77	-0.10	0.49	-0.05	-0.06	0.08	0.14	0.05	10,978	52%
Account	A	0.00	0.19	0.31	0.23	0.05	0.06	-0.06	0.001	0.13	12,704	22%
Account	B	0.01	0.01	0.11	0.10	0.01	0.03	0.03	0.002	0.03	19,532	9%
CD	A	0.00	0.05	0.21	0.56	-0.17	0.03	0.21	0.02	-0.05	24,920	39%
CD	B	0.01	0.72	0.11	0.27	-0.08	-0.08	0.16	0.01	0.06	29,342	35%
Business	A	0.16	-1.66	0.66	-0.03	0.06	2.34	0.25	-0.01	-0.76	94,735	34%
Business	B	0.06	0.28	-0.27	-0.01	0.45	-0.16	0.23	0.57	0.56	98,781	59%
IRA	A	0.02	-0.02	0.07	-0.02	0.05	0.68	-0.12	0.003	0.55	15,050	36%
IRA	B	0.01	0.05	-0.04	0.03	-0.02	0.54	-0.01	0.001	0.29	33,719	25%

V. Correcting Second Home Equity in HRS/AHEAD

Overview

Second home equity is an important component of both housing equity and net worth for the old population. It has been covered, implicitly or explicitly, across all waves of HRS and AHEAD surveys. But due to a skip-pattern error, not all households with second homes were asked detailed questions about current market value, amount of mortgage, etc… The negative impact of the inconsistent treatment of second home on the estimation of housing equity and net worth is substantial. When the second home information is not collected for all the households who own second homes (as in AHEAD 1995 and HRS 1996), the second home equity measure based on the partial data is likely to suffer from selection bias, rendering vulnerable both measures of total housing equity and total net worth. This paper reports on an imputation method to correct for this bias that we demonstrate and find effective.

The Issues

Second home equity is an important component of both housing equity and net worth for the old population. It has been covered, implicitly or explicitly, across all waves of HRS and AHEAD surveys. The treatment of second home equity, however, has not been consistent. Questions regarding second home should be asked for all households who have second homes at the time of the interview. Parallel to the questions about the primary (or main) home, the question sequence about second home should be independent of the sequence about real estate investment, making second home equity distinguishable from real estate equity. HRS92, HRS94, and HRS98 and after are the only survey waves that have exactly followed these rules.

In AHEAD93, second home was explicitly treated as part of real estate investment. When asked about their real estate assets, a household was directed to include "any real estate (other than its main home), such as land, a second home, rental real estate, a partnership, or money owed to you on a land contract or mortgage" (see Question K2, AHEAD93 Codebook). In both AHEAD95 and HRS96, second home was correctly treated as independent of primary home and real estate investment. But due to a skip-pattern error, not all households with second homes were asked detailed questions about current market value, amount of mortgage, etc. Specifically, any respondents who had not lived in their second homes for at least two months of the year would

not have been asked about their second home equity (see Questions CS31, CS35, and F40 for the relevant question flow in the AHEAD95 and HRS96 Codebooks). Since most people do not live in their second homes for two months or more of the year, this problem has skipped most second-home owners effectively past the detailed questions about value, mortgage, etc, and mis-classified most second home owners as not owning second homes.

The negative impact of the inconsistent treatment of second home on the estimation of housing equity and net worth is substantial. Based on results from HRS 1992, 1994, 1996, 1998, 2000, and 2002, second home equity in the aggregate accounts for more than 10% of total housing equity, more than 3% of total net worth for the HRS cohort (Table 32) and about 8% of total housing equity and about 3% of the total net worth for the AHEAD cohort (Table 33). When second home equity is combined with real estate investment (as in AHEAD 1993), there is no direct way to get an accurate measure of second home—and thus, total housing—equity. On the other hand, when the second home information is not collected for all the households who own second homes (as in AHEAD 1995 and HRS 1996), the second home equity measure based on the partial data is likely to suffer from selection bias, rendering vulnerable both measures of total housing equity and total net worth.

Table 32. The Role of Second Home Equity in Total Housing Equity and Total Net Worth: the HRS Cohort, HRS 1992-2002

	HRS 1992	HRS 1994	HRS 1996	HRS 1998	HRS 2000	HRS 2002
% 2nd Home Ownership	13.9	13.0	4.0	13.2	13.1	13.3
Second Home Equity	15,140	11,515	4,735	12,528	14,372	17,855
Primary Home Equity	82,566	88,207	90,512	99,428	110,248	114,780
Total Housing Equity	97,707	99,723	95,248	111,956	124,621	132,635
Total Net Worth	274,366	310,365	319,485	378,375	418,389	410,450
% of 2nd Home Equity in Total Housing Equity	15.5	11.5	4.9	11.2	11.5	13.5
% of 2nd Home Equity in Total Net Worth	5.5	3.7	1.4	3.3	3.4	4.3

Note: Total net worth is a combination of the total housing equity and non-housing assets, which exclude "trusts not reported earlier". All the housing equity and net worth variables are weighted means in 2002 dollars. The HRS 1996 results contain obvious errors in the percentage of second home ownership and second home equity.

Table 33. The Role of Second Home Equity in Total Housing Equity and Total Net Worth: the AHEAD Cohort, AHEAD 1995-2002

	AHEAD 1993	AHEAD 1995	AHEAD 1998	AHEAD 2000	AHEAD 2002
% 2nd Home Ownership	NA	5.0	7.7	10.8	10.9
Second Home Equity	NA	4,845	8,591	7,956	11,032
Primary Home Equity	87,159	90,351	91,128	95,499	95,129
Total Housing Equity	-	95,197	99,719	103,455	106,161
Total Net Worth	217,933	307,000	302,214	312,356	304,137
% of 2nd Home Equity in Total Housing Equity	NA	5.1	8.6	7.7	10.4
% of 2nd Home Equity in Total Net Worth	-	1.6	2.8	2.5	3.6

Note: Total net worth is a combination of the total housing equity and non-housing assets, which exclude "trusts not reported earlier". All the housing equity and net worth variables are weighted means in 2002 dollars. The AHEAD 1995 results contain obvious errors in second home equity. The AHEAD 1993 total net worth is apparently also flawed, an issue to be addressed elsewhere.

In this part of the paper, we intend to correct—at least partially—the second-home data. The plan is to use the information in HRS98 as the gold standard, imputing second home equity for HRS96 and AHEAD95 from data in later waves on ownership and data of purchase. Section II explores the cross-wave relationship in second home ownership between HRS98 and HRS96,

and between HRS98 and AHEAD95. Based on these connections, Section III proposes a simple method for correcting the second-home errors in HRS96 and AHEAD95, and Section IV reports some preliminary results after the data corrections. The paper concludes with an extension of our simple correction method to the second home problem for AHEAD93, where the problem is that second-home equity is combined with real estate investment.

Building Cross-Wave Connections

Two facts in HRS98 about housing and assets make it feasible to correct second home equity for HRS96 and AHEAD95. First, information on the year of purchase for second home is available in HRS98. This allows us to be able to predict second home ownership in a previous wave. If a second-home-owning household reported in HRS98 that it had purchased its second home in 1994, for example, it should also report having a second home in HRS96 or AHEAD95. Conversely, if the purchase year was 1997, the household would usually have no second home in HRS96 or AHEAD95.

HRS98 also has information on housing transactions. In Section N (Widowhood and Divorce), each household was asked if it had bought or sold any home (main or second) since the last interview. While this information may not help us identify all previously mis-classified second home owners, it would help us separate a previously mis-classified second home owner from an owner who had bought his/her second home after his/her previous interview. We shall elaborate this point as we proceed.

There are four possible answers to a question whether a household owned a second home in HRS98 and/or in a previous wave, say, HRS96: it owned a second home in both waves, it owned a second home in neither waves, and it owned a second home in only one of the waves. Figure 2 depicts the four potential scenarios generated from the question.

Cell A represents all households who owned second homes in both waves. Theoretically, it includes second-home-owning households who made no housing transactions since the HRS96 interview, and households who sold and bought second homes after the HRS96 interview. The information on second home equity is available in both waves for these households, and this information will be the backbone in our exercise of second home equity imputation.

Figure 2: Second Home Ownership in HRS98 and HRS96: A Theoretical Model

		HRS 1998	
		Yes	**No**
HRS 1996	**Yes**	**A**	**B** (New Sale)
	No	**C** (New Purchase)	**D**

Cell D represents all households who did not own second homes in HRS98 and HRS96. For these households, second home transactions are possible, but unlikely. (They could, in principle, have both bought and sold a second home in 1997.)

Cells B and C include most of the households who made housing transactions after the HRS96 interview. A household without a second home in HRS96 would be in Cell C if it purchased one after the interview, while a household with a second home in 1996 would be in Cell B if it sold the home after the interview. Information on second home equity is available only in HRS98 for the households in Cell C, and in HRS96 for the households in Cell B.

One consequence of the inconsistent treatment of second home equity described earlier is that a great number of the households who are supposed to be in Cell A are mis-classified into Cell C, thereby reducing the percentage of households who had second homes in both waves (Cell A). As evidenced in Tables 34 and 35, the percentages of second-home owning households in HRS96 and AHEAD95 were, respectively, 3.4 and 4.1, both substantially lower than their counterparts in HRS98 (12.4 for the HRS96 households, and 6.9 for the AHEAD95 households).

Table 34. Second Home Ownership in HRS98 and HRS96: Empirical Results before Correction

		HRS 1998			
		Yes	No	Total	(%)
HRS 1996	Yes	163	47	210	(3.4)
	No	612	5,343	5,955	(96.6)
	Total	775	5,390	6,165	(100)
	(%)	(12.4)	(87.4)	(100)	

Note: Only the households interviewed in both HRS96 and HRS98 are listed in this table.

Table 35. Second Home Ownership in HRS98 and AHEAD95: Empirical Results before Correction

		HRS 1998			
		Yes	No	Total	(%)
AHEAD 1995	Yes	101	68	169	(4.1)
	No	192	3,892	4,084	(96.0)
	Total	293	3,960	4,253	(100)
	(%)	(6.9)	(93.1)	(100)	

Note: Only the households interviewed in both AHEAD95 and HRS98 are listed in this table.

The Correction Method

Our method of correcting second home equity consists of two stages. In the first stage, we identify the mis-classified households in Cell C, assigning them back to Cell A. This may be done based on the following two sequential rules ---

(a) If a HRS (or AHEAD) household in Cell C reported in HRS98 that it had purchased its second home before 1996 (or 1995), this household will be treated as mis-classified, and assigned to Cell A; and

(b) If the first rule fails to assign the household to Cell A, but records in HRS98 show that the household did not sell any home after the HRS96 (or AHEAD95) interview, the household will also be assigned to Cell A.[22]

For simplicity, a household who can be identified as mis-classified through these rules will be called as *identifiable mis-classified household*. Our second home equity corrections are limited to such households. Any households in Cell C who cannot be identified as mis-classified will remain in that cell, and we will not correct second home equity in HRS96 or AHEAD95.

The second stage of our correction method involves an estimation of the second-home equity for the identifiable mis-classified households in HRS96 or AHEAD95, based on the information available for the households in Cell A that are identified in Tables 34 and 35. Obviously, there are various ways to do this. The method presented below seems to be one of the simplest.

Let second-home equity as reported in HRS98 and HRS96 (or AHEAD95) be, respectively, X and Y. The relationship between the two is assumed to be (1),

$$Y = X\updownarrow + \odot \qquad\qquad (1),$$

where $\updownarrow$ is a factor related to the rate of appreciation of second-home equity,[23] and $\odot$ is a random error term.

Equation (1) may be estimated by least squares. Based on this equation, we then generate predicted value for each household in HRS96 (or AHEAD95) for which a reported X is available in HRS98. The final estimate of the second home equity for each identifiable mis-classified household may then be determined by a hotdeck imputation procedure that is based on the predicted value of Y.

---

[22] We understand that these rules cannot identify all mis-classified households in Cell C (for example, a household who had made multiple housing transactions after the HRS96 (or AHEAD95) interview), nor can they prevent some households in Cell C from being mis-identified (for example, a household who reported in HRS98 inaccurate information on the purchase year of its second home or its housing transaction history). But we believe that these rules should be able to correctly identify most of the mis-classified households.

[23] To be exact, if the rate of appreciation of the second-home equity is *r*, then $\updownarrow = 1/(1+r)$.

Preliminary Results

The effect of the corrections on the HRS/AHEAD second home data may be best seen in Tables 36 and 37. Before the corrections, only 3.4% of HRS96 and 4.0% of AHEAD95 households have reported to have second homes, and have non-missing information on their second-home equity. After the corrections, the number increases to 13.1% for the HRS96 households, and 8.4% for the AHEAD95 households.

Table 36. Second Home Ownership in HRS98 and HRS96: The Effect of Corrections

		HRS 1998			
HRS 1996		Yes	No	Total	(%)
Pre-Correction	Yes	163	47	210	(3.4)
	No	612	5,343	5,955	(96.6)
Post-Correction	Yes	763	47	810	(13.1)
	No	12	5,343	5,355	(86.9)
	Total	775	5,390	6,165	(100)
	(%)	(12.6)	(87.4)		(100)

Note: Only the households interviewed in both HRS96 and HRS98 are listed in this table.

Since the percentages of second home ownership in HRS96 and AHEAD95 are significantly increased after corrections, we expect that the mean values of the second home equities in the two waves will be increased significantly as well.

In Tables 38 and 39 we have replicated Tables 31 and 32 with the corrected information on the second home equity for HRS 1996 and AHEAD 1995. The time-series patterns of second home ownership and equity are more consistent now for both the HRS and AHEAD cohorts. In HRS 1996, 13.1% of the households owned second homes, compared to 13.9% in HRS 1992, 13.0% in HRS 1994, 13.2% in HRS 1998, 13.1% in HRS 2000, and 13.3% in HRS 2002. In AHEAD 1995, 8.2% of the households owned second homes, compared to 7.7% in HRS 1998, 10.8% in HRS 2000, and 10.9% in 2002.

With the corrections, the mean value of the second home equity in HRS 1996 has increased by about 165%, from $4,735 to $12,590, while the mean value of the second home equity in

AHEAD 1995 has increased by about 90%, from $4,845 to $9,262. The second home equity now comprises 12.2% of total housing equity and 3.8% of total net worth in HRS 1996, and 9.3% of total housing equity and 3.0% of total net worth in AHEAD 1995. All these numbers are quite comparable to the counterpart data in other HRS or AHEAD waves, suggesting that our corrections have indeed improved the data quality.

Table 37. Second Home Ownership in HRS98 and AHEAD95: The Effect of Corrections

		HRS 1998			
AHEAD 1995		Yes	No	Total	(%)
Pre-Correction	Yes	101	68	169	(4.0)
	No	192	3,892	4,084	(96.0)
Post-Correction	Yes	290	68	358	(8.4)
	No	3	3,892	3,895	(91.6)
	Total	293	3,960	4,253	(100)
	(%)	(6.9)	(93.1)	(100)	

Note: Only the households interviewed in both AHEAD95 and HRS98 are listed in this table.

It is interesting to note that the households who were mis-classified in both AHEAD 1995 and HRS 1996 appeared to have lower values on their second home equities. For those mis-classified households in AHEAD 1995, the mean value of their second home equities was $100,852, compared to $126,656 for those reported to own second homes. In HRS 1996, the two numbers were $81,525 and $109,137, respectively. Since a mis-classified household was one who lived in second home for less than two months of the year, the results suggest a positive correlation between the duration of second home stay and the quality of the second home.

Table 38. **The Role of Second Home Equity in Total Housing Equity and Total Net Worth after the Second Home Equity Problem in HRS 1996 Corrected: the HRS Cohort, HRS 1992-2002**

	HRS 1992	HRS 1994	HRS 1996	HRS 1998	HRS 2000	HRS 2002
% 2nd Home Ownership	13.9	13.0	13.1	13.2	13.1	13.3
Second Home Equity	15,140	11,515	12,590	12,528	14,372	17,855
Primary Home Equity	82,566	88,207	90,512	99,428	110,248	114,780
Total Housing Equity	97,707	99,723	103,102	111,956	124,621	132,635
Total Net Worth	274,366	310,365	327,340	378,375	418,389	410,450
% of 2nd Home Equity in Total Housing Equity	15.5	11.5	12.2	11.2	11.5	13.5
% of 2nd Home Equity in Total Net Worth	5.5	3.7	3.8	3.3	3.4	4.3

Note: Total net worth is a combination of the total housing equity and non-housing assets, which exclude "trusts not reported earlier". All the housing equity and net worth variables are weighted means in 2002 dollars. Both the second home ownership and equity now show more consistent time-series patterns.

Table 39. **The Role of Second Home Equity in Total Housing Equity and Total Net Worth after the Second Home Equity Problem in AHEAD 1995 Corrected: the AHEAD Cohort, AHEAD 1995-2002**

	AHEAD 1993	AHEAD 1995	AHEAD 1998	AHEAD 2000	AHEAD 2002
% 2nd Home Ownership	7.8	8.2	7.7	10.8	10.9
Second Home Equity	7,857	9,262	8,591	7,956	11,032
Primary Home Equity	87,159	90,351	91,128	95,499	95,129
Total Housing Equity	95,016	99,613	99,719	103,455	106,161
Total Net Worth	217,933	311,417	302,214	312,356	304,137
% of 2nd Home Equity in Total Housing Equity	8.3	9.3	8.6	7.7	10.4
% of 2nd Home Equity in Total Net Worth	3.6	3.0	2.8	2.5	3.6

Note: Total net worth is a combination of the total housing equity and non-housing assets, which exclude "trusts not reported earlier". All the housing equity and net worth variables are weighted means in 2002 dollars. Both the second home ownership and equity now show more consistent time-series patterns. The percentage of second home equity in total net worth in AHEAD 1993 seems to be on the high side, but that is due to the very low level of net worth in AHEAD 1993.

The AHEAD 1993 Problem

The AHEAD93 problem (i.e., second home equity combined with real estate investment equity) is different from the problem troubling HRS96 and AHEAD95. But it may be handled in a way similar to what we have done for HRS96 and AHEAD95. Basically, we may continue to use the information about second home in a later wave (e.g., purchase year, housing transaction history in AHEAD 1995) to predict the existence (or the lack of it) of second home in AHEAD93, and then impute second home equity for those identifiable second home owners.

Two points need to be made here. First, in predicting the second home ownership for AHEAD 1993, one needs to use the corrected—not just reported—second home ownership information in AHEAD 1995. Second, to impute second home equity, one has to make an assumption about the rate of appreciation of the housing market.

Table 39, Column 1 summarizes the second home ownership and equity results for AHEAD 1993. According to the table, 7.8% of the households owned second homes in AHEAD 1993, and the mean value of the second home equity was about $7,857.[24] Compared to the results in other AHEAD waves, both numbers seem to be quite reasonable.

---

[24] To impute second home equity for AHEAD 1993, we first estimated a simple linear relationship between the second home equities in AHEAD 1995 and HRS 1998. We then predicted the AHEAD 1993 equity based on the relationship and the observed or imputed equity values in AHEAD 1995 or HRS 1998.

Appendix 1. Structure of Unfolding Bracket Question Sequence

Unfolding brackets as a survey technique aiming to reduce item non-responses have been used widely in HRS in questions related to income, assets, and health care expenditures. A complete unfolding bracket question sequence generally consists of three sets of questions: ownership, amount if owned, and bracket questions, DK or REF. These questions are illustrated in the following example for stocks/mutual funds holdings (Table A1).

Table A1. Unfolding Bracket Question Sequence for Stocks/Mutual Funds Holdings

Panel 1. For Low Entry Point Group

(Ownership Question)
A). (Aside from anything you have already told me about,) Do you (or your [husband/wife/partner]) have any shares of stock or stock mutual funds?

       1. YES   5. NO   8. DK   9. RF
                 |_____|_____|
       ⇓
               SKIP OUT

(Amount Question)
B. If you sold all those and paid off anything you owed on them, about how much would you have?

                            DK   RF
   $---------------------------   ------------------------
    AMOUNT [SKIP OUT]
                           ⇓

(Bracket Questions)
C1. Would it amount to less than $2,500, more than $2,500, or what?

     1. Less than $2,500 [SKIP OUT]   3. About $2,500 [SKIP OUT]
     5. More than $2,500 [Go to C2]   8. DK   9. RF
                           |____SKIP OUT____|

C2. Would it amount to less than $25,000, more than $25,000, or what?

     1. Less than $25,000 [SKIP OUT]   3. About $25,000 [SKIP OUT]
     5. More than $25,000 [Go to C3]   8. DK   9. RF
                           |____SKIP OUT____|

C3. Would it amount to less than $125,000, more than $125,000, or what?

     1. Less than $125,000 [SKIP OUT]   3. About $125,000 [SKIP OUT]
     5. More than $125,000 [Go to C4]   8. DK   9. RF
                           |____SKIP OUT____|

C4. Would it amount to less than $400,000, more than $400,000, or what?

     1. Less than $400,000 [SKIP OUT]   3. About $400,000   [SKIP OUT]

5. More than $400,000 [SKIP OUT]   8. DK   9. RF
                                                                |____SKIP OUT____|

Panel 2. For Middle Entry Point Group

(Ownership Question)
A). (Aside from anything you have already told me about,) Do you (or your [husband/wife/partner]) have any shares of stock or stock mutual funds?

          1. YES   5. NO   8. DK   9. RF
                              |_____|_____|
           ⇓   SKIP OUT

(Amount Question)
B. If you sold all those and paid off anything you owed on them, about how much would you have?

                                        DK   RF
     $------------------   -----------------------
     AMOUNT [SKIP OUT]
                                              ⇓

(Bracket Questions)
C2. Would it amount to less than $25,000, more than $25,000, or what?

       1. Less than $25,000 [Go to C1]   3. About $25,000 [SKIP OUT]
       5. More than $25,000 [Go to C3]   8. DK   9. RF
                                                         |____SKIP OUT____|

C1. Would it amount to less than $2,500, more than $2,500, or what?

       1. Less than $2,500 [SKIP OUT]   3. About $2,500 [SKIP OUT]
       5. More than $2,500   [SKIP OUT]   8. DK   9. RF
                                                         |____SKIP OUT____|

C3. Would it amount to less than $125,000, more than $125,000, or what?

       1. Less than $125,000 [SKIP OUT]   3. About $125,000 [SKIP OUT]
       5. More than $125,000 [Go to C4]   8. DK   9. RF
                                                         |____SKIP OUT____|

C4. Would it amount to less than $400,000, more than $400,000, or what?

       1. Less than $400,000 [SKIP OUT]   3. About $400,000   [SKIP OUT]
       5. More than $400,000 [SKIP OUT]   8. DK   9. RF
                                                         |____SKIP OUT____|

Panel 3. For High Entry Point Group

(Ownership Question)
A). (Aside from anything you have already told me about,) Do you (or your [husband/wife/partner]) have any shares of stock or stock mutual funds?

1. YES   5. NO   8. DK   9. RF

|_____|_____|

⇓

SKIP OUT

(Amount Question)

B. If you sold all those and paid off anything you owed on them, about how much would you have?

                             DK   RF

$------------------   ------------------------

AMOUNT [SKIP OUT]

⇓

(Bracket Questions)

C3. Would it amount to less than $125,000, more than $125,000, or what?

     1. Less than $125,000 [Go to C2]   3. About $125,000 [SKIP OUT]
     5. More than $125,000 [Go to C4]   8. DK   9. RF

                               |____SKIP OUT____|

C2. Would it amount to less than $25,000, more than $25,000, or what?

     1. Less than $25,000 [Go to C1]   3. About $25,000 [SKIP OUT]
     5. More than $25,000 [SKIP OUT]   8. DK   9. RF

                               |____SKIP OUT____|

C1. Would it amount to less than $2,500, more than $2,500, or what?

     1. Less than $2,500 [SKIP OUT]   3. About $2,500 [SKIP OUT]
     5. More than $2,500   [SKIP OUT]   8. DK   9. RF

                               |____SKIP OUT____|

C4. Would it amount to less than $400,000, more than $400,000, or what?

     1. Less than $400,000 [SKIP OUT]   3. About $400,000   [SKIP OUT]
     5. More than $400,000 [SKIP OUT]   8. DK   9. RF

                               |____SKIP OUT____|

What distinguishes an unfolding bracket sequence from traditional survey questions is its inclusion of range questions (less than x, about x, more than x), which become necessary for those who give a positive answer to the ownership question but fail to provide a specific amount to the open-ended question.

An unfolding bracket sequence can be characterized by the number of breakpoints and the values of these breakpoints, which in HRS are pre-determined based on an algorithm described in Hill (1999). In the above example, the unfolding bracket sequence has four breakpoints, valued at $2,500, $25,000, $125, 000, and $400,000, respectively. Different bracket sequences may have a

different number of breakpoints, and/or different breakpoint values. The bracket sequence for transportation vehicles in HRS 1998, for example, has only three breakpoints, with breakpoint values at $5,000, $50,000, and $100,000.

Another feature of an HRS unfolding bracket sequence is that the sequence stops whenever a DK or RF answer is given to a bracket question. Thus, many stock-owning households do not provide any definitive information at all regarding the value of their stock holdings even when prompted with bracket questions.[25] The sequence also stops when bracket questions are at an upper or lower limit (e.g., less than the lowest breakpoint, or more than the highest breakpoint), when an "about" answer is given, or when a lower and upper bracket limit is identified.

Entry points become an issue in the unfolding bracket sequence because the distribution of respondents among bracket categories may depend on the entry bracket. Prior to interview and since HRS 1998, each HRS household is assigned to one of three randomly selected groups, which may conveniently be called a low-entry-point group, a middle-entry-point group, and a high-entry-point group, respectively.

Households assigned to the low-entry-point group will get bracket questions starting with lowest breakpoint—in this case, $2,500. They will first be asked question C1, then C2, and C3, and conclude the bracket sequence with C4 if needed (See Table A1, Panel 1). In contrast, households assigned to the middle-entry-point group will get bracket questions starting with the second lowest breakpoint—in this case, $25,000. They will first be asked question C2, and then—depending on their answers to the question—C1, or C3 and C4 (See Table A1, Panel 2). The bracket question sequence for households assigned to the high-entry-point group is determined in a similar way: starting with C3, and then C4, or C2 and C1—depending on their answers to a previous question (See Table A1, Panel 3).

"Unfolding range" is a term used for characterizing respondent answers to the entire sequence of bracket questions. In combination with entry-point, unfolding range uniquely determines a respondent's response pattern to various bracket questions. For a stock-owning household who belongs to the low-entry-point group and who fails to reveal the exact value of its stock holdings, for example, its answers to the bracket questions may be as follows: C1 = 5, C2 = 5, and C3 = 1.

---

[25] That is, they give a DK/RF answer to the entry bracket question, and the bracket sequence stops right there.

We know that this household's holdings are greater than $25,000 and less than $125,000, and denote its unfolding range as "> $25,000 and < $125,000." Similarly, the unfolding range for a household in the high-entry-point group with bracket answers C3 = 1, C2=1, and C1= 5 is "> $2,500 and < $25,000."

Table A2 lists all the possible unfolding ranges for stock-owning households in the four breakpoint unfolding sequence illustrated above. Readers may find it useful for understanding the data in Appendixes 2 and 3 of this paper.

Appendices 2 and 3: A Summary

Appendix 2 provides the details of the imputation results for each of the 10 net worth components in the HRS study (nine assets and debt). The top half of the table shows the raw counts of numbers of household respondents by unfolding bracket range, while the bottom half of the table shows the distribution of household respondents by unfolding bracket categories.

Table A2. Possible Unfolding Ranges for Stocks/Mutual Funds Holdings

Bracket Pattern	Entry-Point Group	Why Bracket Sequence Stops	Unfolding Range
C1=1	Low-entry-point	1	< $2,500
C1=3	Low-entry-point	2	About $2,500
C1=8 or 9	Low-entry-point	3	DK/RF if <> $2,500
C1=5, C2=1	Low-entry-point	4	> $2,500 & < $25,000
C1=5, C2=3	Low-entry-point	2	About $25,000
C1=5, C2= 8 or 9	Low-entry-point	3	DK/RF if <> $25,000
C1=5, C2=5, C3=1	Low-entry-point	4	> $25,000 & < $125,000
C1=5, C2=5, C3=3	Low-entry-point	2	About $125,000
C1=5, C2=5, C3=8 or 9	Low-entry-point	3	DK/RF if <> $125,000
C1=5, C2=5, C3=5, C4=1	Low-entry-point	4	> $125,000 & < $400,000
C1=5, C2=5, C3=5, C4=3	Low-entry-point	2	About $400,000
C1=5, C2=5, C3=5, C4=8 or 9	Low-entry-point	3	DK/RF if <> $400,000
C1=5, C2=5, C3=5, C5=5	Low-entry-point	1	> $400,000
C2=1, C1=1	Middle-entry-point	1	< $2,500
C2=1, C1=3	Middle-entry-point	2	About $2,500
C2=1, C1=8 or 9	Middle-entry-point	3	DK/RF if <> $2,500
C2=1, C1=5	Middle-entry-point	4	> $2,500 & < $25,000
C2=3	Middle-entry-point	2	About $25,000
C2=8 or 9	Middle-entry-point	3	DK/RF if <> $25,000
C2=5, C3=1	Middle-entry-point	4	> $25,000 & < $125,000
C2=5, C3=3	Middle-entry-point	2	About $125,000
C2=5, C3=8 or 9	Middle-entry-point	3	DK/RF if <> $125,000
C2=5, C3=5, C4=1	Middle-entry-point	4	> $125,000 & < $400,000
C2=5, C3=5, C4=3	Middle-entry-point	2	About $400,000
C2=5, C3=5, C4= 8 or 9	Middle-entry-point	3	DK/RF if <> $400,000
C2=5, C3=5, C4=5	Middle-entry-point	1	> $400,000
C3=1, C2=1, C1=1	High-entry-point	1	< $2,500
C3=1, C2=1, C1=3	High-entry-point	2	About $2,500
C3=1, C2=1, C1=8 or 9	High-entry-point	3	DK/RF if <> $2,500
C3=1, C2=1, C1=5	High-entry-point	4	> $2,500 & < $25,000
C3=1, C2=3	High-entry-point	2	About $25,000
C3=1, C2=8 or 9	High-entry-point	3	DK/RF if <> $25,000
C3=1, C2=5	High-entry-point	4	> $25,000 & < $125,000
C3=3	High-entry-point	2	About $125,000
C3=8 or 9	High-entry-point	3	DK/RF if <> $125,000
C3=5, C4=1	High-entry-point	4	> $125,000 & < $400,000
C3=5, C4=3	High-entry-point	2	About $400,000
C3=5, C4=8 or 9	High-entry-point	3	DK/RF if <> $400,000
C3=5, C4=5	High-entry-point	1	> $400,000

Note: 1) The unfolding sequence has four breakpoints valued at $2,500, $25,000, $125,000, and $400,000. 2)Bracket pattern is arranged in the order of bracket questions answered. 2) Households from the low-entry-point group always start with question C1. Households from the middle-entry-point group always start with question C2. Households from the high-entry-point group always start with question C3. 3) Reason why bracket sequence stops: 1 = bracket question is at an upper or lower limit; 2 = an "about" answer is given; 3 = a "DK" or "RF" answer is given; 4 = a lower and upper bracket limit is identified (e.g., the amount is less than $50,000 and greater than $5000).

The table is divided into 10 panels (one panel for each net worth component); three major components for each panel (respondents who reported don't know (DK) when asked asset amount, those who reported refuse (RF) when asked about asset amount, and the sum of the don't know and refuse respondents); and three randomly selected entry points which vary by type of asset. Thus, for example, Panel 1 deals with Real Estate holdings, shows the numbers of households who replied don't know or refused for each of the three entry points (2.5 K, 125 K, and 500K) and for each of the 13 unfolding bracket ranges. The bottom half of the table shows the percent distribution of households in each of six bracket categories.

The largest and most consistent difference in the data shown in Appendix 2 is in the distributions of DK cases and RF cases. For DK cases, only about ten percent of households could not provide a bracket amount when asked the entry bracket question, and another roughly 8-9 percent could not provide a response when asked the bracket question somewhere other than in the entry bracket sequence. (These numbers are taken from Panel 1 of Appendix 2, which asks about Real Estate assets.) In contrast, fully sixty percent of REF households refused to provide a bracket response when asked the entry bracket question, and another roughly 15% would not provide a bracket response when asked the bracket question somewhere other than in the entry bracket sequence. Comparable numbers are found for other assets, with the general pattern being that the refusal rate for the entry bracket questions is even higher—more like 70% than sixty.

Appendix 3 contains mean amounts for each of the bracket categories, with the amount being obtained by random draws for each household in each of the bracket categories. That is, if 50 households wound up in the bracket category "< $50K > $5K", the program would search the data for continuous amount cases in that category, then assign a randomly selected case for each of the 50 households. In cases where the respondent said that their asset was worth "about 10K", when asked whether it was worth "< $10 K, > $10K, or what?", the program would assign $10 K to each of the households in the "about $10 K" category.

Appendix 2, Panel 1: Real Estate Holdings

Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether DK or REF in Response to questions about Asset holdings (unweighted N)

HRS 1998: Real Estate Holdings

	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
Unfolding range	$2,500	$125,000	$500,000	$2,500	$125,000	$500,000	$2,500	$125,000	$500,000
< $2,500	9	7	9	1	0	1	10	7	10
About $2,500	10	3	2	0	0	0	10	3	2
DK/RF if <> $2,500	**16**	5	7	**12**	0	2	**28**	5	9
> $2,500 & < $125,000	74	67	59	3	1	3	77	68	62
About $125,000	13	6	5	2	0	1	15	6	6
DK/RF if <> $125,000	9	**24**	6	3	**23**	3	12	**47**	9
> $125,000 & < $500,000	29	16	33	0	2	0	29	18	33
About $500,000	1	6	2	0	0	0	1	6	2
DK/RF if <> $500,000	4	3	**17**	1	2	**14**	5	5	**31**
> $500,000 & < $1,000,000	8	11	8	0	1	1	8	12	9
About $1,000,000	3	2	4	0	0	0	3	2	4
DK/RF if <> $1,000,000	0	2	1	0	2	0	0	4	1
> $1,000,000	2	4	4	1	0	1	3	4	5
Total Case Number	178	156	157	23	31	26	201	187	183
Percentage Distribution by Bracket Category									
DK/RF to entry-bracket question	9.0%	15.4%	10.8%	52.2%	74.2%	53.8%	13.9%	25.1%	16.9%
DK/RF to non-entry-bracket question	7.3%	6.4%	8.9%	17.4%	12.9%	19.2%	8.5%	7.5%	10.4%
< $2,500	5.1%	4.5%	5.7%	4.3%	0.0%	3.8%	5.0%	3.7%	5.5%
>= $2,500 & < $125,000	47.2%	44.9%	38.9%	13.0%	3.2%	11.5%	43.3%	38.0%	35.0%
>= $125,000 & < $500,000	23.6%	14.1%	24.2%	8.7%	6.5%	3.8%	21.9%	12.8%	21.3%
>= $500,000	7.9%	14.7%	11.5%	4.3%	3.2%	7.7%	7.5%	12.8%	10.9%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Note: 1) Entry bracket categories are highlighted. 2) "DK/RF to entry-bracket question" is calculated as ratio of the highlighted cell to the total case number. 3) "DK/RF to non-entry-bracket question" is calculated as ratio of all the non-highlighted cells ending in DK/RF to the total case number.

Appendix 2, Panel 2: Business/Farm Asset Holdings
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether DK or REF
in Response to questions about Asset holdings (unweighted N)

HRS 1998: Business/Farm Asset Holding

Unfolding range	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
	$5,000	$10,000	$100,000	$5,000	$10,000	$100,000	$5,000	$10,000	$100,000
< $5,000	13	6	6	0	0	0	13	6	6
About $5,000	4	2	3	1	0	0	5	2	3
DK/RF if <> $5,000	**25**	0	0	**21**	0	0	**46**	0	0
> $5,000 & < $10,000	3	5	2	0	0	0	3	5	2
About $10,000	4	5	4	1	0	0	5	5	4
DK/RF if <> $10,000	3	**22**	1	1	**14**	0	4	**36**	1
> $10,000 & < $100,000	30	25	28	3	2	2	33	27	30
About $100,000	9	9	2	0	1	0	9	10	2
DK/RF if <> $100,000	5	7	**30**	3	1	**22**	8	8	**52**
> $100,000 & < $1,000,000	22	39	33	2	3	0	24	42	33
About $1,000,000	1	3	3	0	0	0	1	3	3
DK/RF if <> $1,000,000	1	1	0	1	2	1	2	3	1
> $1,000,000	3	3	10	0	0	1	3	3	11
Total Case Number	123	127	122	33	23	26	156	150	148

Percentage Distribution by Bracket Category

	$5,000	$10,000	$100,000	$5,000	$10,000	$100,000	$5,000	$10,000	$100,000
DK/RF to entry-bracket question	20.3%	17.3%	24.6%	63.6%	60.9%	84.6%	29.5%	24.0%	35.1%
DK/RF to non-entry-bracket question	7.3%	6.3%	0.8%	15.2%	13.0%	3.8%	9.0%	7.3%	1.4%
< $5,000	10.6%	4.7%	4.9%	0.0%	0.0%	0.0%	8.3%	4.0%	4.1%
>= $5,000 & < $10,000	5.7%	5.5%	4.1%	3.0%	0.0%	0.0%	5.1%	4.7%	3.4%
>= $10,000 & < $100,000	27.6%	23.6%	26.2%	12.1%	8.7%	7.7%	24.4%	21.3%	23.0%
>= $100,000	28.5%	42.5%	39.3%	6.1%	17.4%	3.8%	23.7%	38.7%	33.1%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Note: 1) Entry bracket categories are highlighted. 2) "DK/RF to entry-bracket question" is calculated as ratio of the highlighted cell to the total case number.
3) "DK/RF to non-entry-bracket question" is calculated as ratio of all the non-highlighted cells ending in DK/RF to the total case number.

Appendix 2, Panel 3: IRA Holdings
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether DK or REF
in Response to questions about Asset holdings (unweighted N)

HRS 1998: IRA Holdings

Unfolding range	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
	$10,000	$25,000	$100,000	$10,000	$25,000	$100,000	$10,000	$25,000	$100,000
< $10,000	48	40	29	12	6	5	60	46	34
About $10,000	15	13	7	2	0	3	17	13	10
DK/RF if <> $10,000	**41**	4	2	**85**	3	2	**126**	7	4
> $10,000 & < $25,000	36	44	22	4	6	8	40	50	30
About $25,000	12	13	23	1	3	3	13	16	26
DK/RF if <> $25,000	9	**45**	13	10	**97**	13	19	**142**	26
> $25,000 & < $100,000	42	68	62	12	17	23	54	85	85
About $100,000	3	8	4	0	0	0	3	8	4
DK/RF if <> $100,000	6	4	**53**	7	4	**85**	13	8	**138**
> $100,000 & < $400,000	21	19	21	5	5	8	26	24	29
About $400,000	2	0	2	0	1	0	2	1	2
DK/RF if <> $400,000	1	2	2	1	3	2	2	5	4
> $400,000	6	6	2	2	0	3	8	6	5
Total Case Number	242	266	242	141	145	155	383	411	397
Percentage Distribution by Bracket Category									
DK/RF to entry-bracket question	16.9%	16.9%	21.9%	60.3%	66.9%	54.8%	32.9%	34.5%	34.8%
DK/RF to non-entry-bracket question	6.6%	3.8%	7.0%	12.8%	6.9%	11.0%	8.9%	4.9%	8.6%
< $10,000	19.8%	15.0%	12.0%	8.5%	4.1%	3.2%	15.7%	11.2%	8.6%
>= $10,000 & < $25,000	21.1%	21.4%	12.0%	4.3%	4.1%	7.1%	14.9%	15.3%	10.1%
>= $25,000 & < $100,000	22.3%	30.5%	35.1%	9.2%	13.8%	16.8%	17.5%	24.6%	28.0%
>= $100,000	13.2%	12.4%	12.0%	5.0%	4.1%	7.1%	10.2%	9.5%	10.1%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Note: 1) Entry bracket categories are highlighted. 2) "DK/RF to entry-bracket question" is calculated as ratio of the highlighted cell to the total case number.
3) "DK/RF to non-entry-bracket question" is calculated as ratio of all the non-highlighted cells ending in DK/RF to the total case number.

Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether DK or REF
in Response to questions about Asset holdings (unweighted N)

HRS 1998: Stock Holdings

	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
Unfolding range	$2,500	$25,000	$125,000	$2,500	$25,000	$125,000	$2,500	$25,000	$125,000
< $2,500	46	25	37	2	0	0	48	25	37
About $2,500	12	11	9	0	0	0	12	11	9
DK/RF if <> $2,500	**58**	11	2	**69**	2	0	**127**	13	2
> $2,500 & < $25,000	71	91	68	7	10	6	78	101	74
About $25,000	18	19	16	1	4	1	19	23	17
DK/RF if <> $25,000	12	**68**	26	11	**88**	2	23	**156**	28
> $25,000 & < $125,000	61	71	96	7	14	9	68	85	105
About $125,000	11	12	13	2	0	5	13	12	18
DK/RF if <> $125,000	10	12	**78**	3	9	**86**	13	21	**164**
> $125,000 & < $400,000	27	49	48	1	5	10	28	54	58
About $400,000	0	4	8	0	2	0	0	6	8
DK/RF if <> $400,000	7	2	5	0	2	7	7	4	12
> $400,000	17	26	24	3	6	5	20	32	29
Total Case Number	350	401	430	106	142	131	456	543	561
Percentage Distribution by Bracket Category									
DK/RF to entry-bracket question	16.6%	17.0%	18.1%	65.1%	62.0%	65.6%	27.9%	28.7%	29.2%
DK/RF to non-entry-bracket question	8.3%	6.2%	7.7%	13.2%	9.2%	6.9%	9.4%	7.0%	7.5%
< $2,500	13.1%	6.2%	8.6%	1.9%	0.0%	0.0%	10.5%	4.6%	6.6%
>= $2,500 & < $25,000	23.7%	25.4%	17.9%	6.6%	7.0%	4.6%	19.7%	20.6%	14.8%
>= $25,000 & < $125,000	22.6%	22.4%	26.0%	7.5%	12.7%	7.6%	19.1%	19.9%	21.7%
>= $125,000	15.7%	22.7%	21.6%	5.7%	9.2%	15.3%	13.4%	19.2%	20.1%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Note: 1) Entry bracket categories are highlighted. 2) "DK/RF to entry-bracket question" is calculated as ratio of the highlighted cell to the total case number.
3) "DK/RF to non-entry-bracket question" is calculated as ratio of all the non-highlighted cells ending in DK/RF to the total case number.

Appendix 2, Panel 5: Checking, Savings, and MM Account Holdings
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether DK or REF
in Response to questions about Asset holdings (unweighted N)

HRS 1998: Checking, Savings and MM Account Holdings

Unfolding range	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
	$5,000	$50,000	$150,000	$5,000	$50,000	$150,000	$5,000	$50,000	$150,000
< $5,000	254	244	214	50	45	33	304	289	247
About $5,000	41	48	51	10	10	10	51	58	61
DK/RF if <> $5,000	**94**	32	12	**264**	21	13	**358**	53	25
> $5,000 & < $50,000	185	187	159	37	45	35	222	232	194
About $50,000	10	19	27	3	4	6	13	23	33
DK/RF if <> $50,000	16	**73**	27	17	**247**	20	33	**320**	47
> $50,000 & < $150,000	33	45	58	5	13	7	38	58	65
About $150,000	4	4	6	0	1	2	4	5	8
DK/RF if <> $150,000	4	4	**68**	2	5	**246**	6	9	**314**
> $150,000 & < $300,000	6	8	13	3	3	4	9	11	17
About $300,000	2	2	2	0	0	0	2	2	2
DK/RF if <> $300,000	2	3	2	0	1	4	2	4	6
> $300,000	0	8	8	1	2	3	1	10	11
Total Case Number	651	677	647	392	397	383	1043	1074	1030
Percentage Distribution by Bracket Category									
DK/RF to entry-bracket question	14.4%	10.8%	10.5%	67.3%	62.2%	64.2%	34.3%	29.8%	30.5%
DK/RF to non-entry-bracket question	3.4%	5.8%	6.3%	4.8%	6.8%	9.7%	3.9%	6.1%	7.6%
< $5,000	39.0%	36.0%	33.1%	12.8%	11.3%	8.6%	29.1%	26.9%	24.0%
>= $5,000 & < $50,000	34.7%	34.7%	32.5%	12.0%	13.9%	11.7%	26.2%	27.0%	24.8%
>= $50,000 & < $150,000	6.6%	9.5%	13.1%	2.0%	4.3%	3.4%	4.9%	7.5%	9.5%
>= $150,000	1.8%	3.2%	4.5%	1.0%	1.5%	2.3%	1.5%	2.6%	3.7%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Note: 1) Entry bracket categories are highlighted. 2) "DK/RF to entry-bracket question" is calculated as ratio of the highlighted cell to the total case number.
3) "DK/RF to non-entry-bracket question" is calculated as ratio of all the non-highlighted cells ending in DK/RF to the total case number.

Appendix 2, Panel 6: Bond Holdings
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether DK or REF
in Response to questions about Asset holdings (unweighted N)

HRS 1998: Bond Holdings

Unfolding range	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
	$2,500	$10,000	$100,000	$2,500	$10,000	$100,000	$2,500	$10,000	$100,000
< $2,500	8	6	4	0	0	0	8	6	4
About $2,500	2	4	1	0	0	1	2	4	2
DK/RF if <> $2,500	**14**	2	1	**29**	1	0	**43**	3	1
> $2,500 & < $10,000	17	12	7	0	0	0	17	12	7
About $10,000	4	5	3	0	1	0	4	6	3
DK/RF if <> $10,000	2	**16**	6	2	**33**	0	4	**49**	6
> $10,000 & < $100,000	19	29	26	6	3	6	25	32	32
About $100,000	2	3	2	0	0	0	2	3	2
DK/RF if <> $100,000	5	3	**21**	0	3	**24**	5	6	**45**
> $100,000 & < $400,000	8	5	11	0	0	0	8	5	11
About $400,000	1	0	0	0	0	0	1	0	0
DK/RF if <> $400,000				0	0	2	0	0	2
> $400,000	5	1	4	0	1	0	5	2	4
Total Case Number	87	86	86	37	42	33	124	128	119
Percentage Distribution by Bracket Category									
DK/RF to entry-bracket question	16.1%	18.6%	24.4%	78.4%	78.6%	72.7%	34.7%	38.3%	37.8%
DK/RF to non-entry-bracket question	8.0%	5.8%	8.1%	5.4%	9.5%	6.1%	7.3%	7.0%	7.6%
< $2,500	9.2%	7.0%	4.7%	0.0%	0.0%	0.0%	6.5%	4.7%	3.4%
>= $2,500 & < $10,000	21.8%	18.6%	9.3%	0.0%	0.0%	3.0%	15.3%	12.5%	7.6%
>= $10,000 & < $100,000	26.4%	39.5%	33.7%	16.2%	9.5%	18.2%	23.4%	29.7%	29.4%
>= $100,000	18.4%	10.5%	19.8%	0.0%	2.4%	0.0%	12.9%	7.8%	14.3%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Note: 1) Entry bracket categories are highlighted. 2) "DK/RF to entry-bracket question" is calculated as ratio of the highlighted cell to the total case number.
3) "DK/RF to non-entry-bracket question" is calculated as ratio of all the non-highlighted cells ending in DK/RF to the total case number.

Appendix 2, Panel 7: CDs, T-bills and Gov't Savings Bond Holdings
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether DK or REF
in Response to questions about Asset holdings (unweighted N)

HRS 1998: CDs, T-bills and Gov't Savings Bond Holdings

	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
Unfolding range	$2,500	$25,000	$125,000	$2,500	$25,000	$125,000	$2,500	$25,000	$125,000
< $2,500	43	18	17	5	6	1	48	24	18
About $2,500	7	11	9	1	1	1	8	12	10
DK/RF if <> $2,500	**37**	3	2	**100**	2	1	**137**	5	3
> $2,500 & < $25,000	51	55	41	6	10	7	57	65	48
About $25,000	4	11	21	0	2	1	4	13	22
DK/RF if <> $25,000	5	**44**	9	5	**116**	3	10	**160**	12
> $25,000 & < $125,000	43	42	50	6	4	9	49	46	59
About $125,000	7	3	2	0	2	1	7	5	3
DK/RF if <> $125,000	5	5	**45**	2	5	**112**	7	10	**157**
> $125,000 & < $250,000	7	6	6	0	0	2	7	6	8
About $250,000	2	2	0	0	0	1	2	2	1
DK/RF if <> $250,000	0	0	3	1	1	0	1	1	3
> $250,000	6	9	4	2	0	2	8	9	6
Total Case Number	217	209	209	128	149	141	345	358	350
Percentage Distribution by Bracket Category									
DK/RF to entry-bracket question	17.1%	21.1%	21.5%	78.1%	77.9%	79.4%	39.7%	44.7%	44.9%
DK/RF to non-entry-bracket question	4.6%	3.8%	6.7%	6.3%	5.4%	2.8%	5.2%	4.5%	5.1%
< $2,500	19.8%	8.6%	8.1%	3.9%	4.0%	0.7%	13.9%	6.7%	5.1%
>= $2,500 & < $25,000	26.7%	31.6%	23.9%	5.5%	7.4%	5.7%	18.8%	21.5%	16.6%
>= $25,000 & < $125,000	21.7%	25.4%	34.0%	4.7%	4.0%	7.1%	15.4%	16.5%	23.1%
>= $125,000	10.1%	9.6%	5.7%	1.6%	1.3%	4.3%	7.0%	6.1%	5.1%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Note: 1) Entry bracket categories are highlighted. 2) "DK/RF to entry-bracket question" is calculated as ratio of the highlighted cell to the total case number.
3) "DK/RF to non-entry-bracket question" is calculated as ratio of all the non-highlighted cells ending in DK/RF to the total case number.

Appendix 2, Panel 8: Vehicle Holdings
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether DK or REF
in Response to questions about Asset holdings (unweighted N)

HRS 1998: Vehicles

	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
Unfolding range	$5,000	$25,000	$200,000	$5,000	$25,000	$200,000	$5,000	$25,000	$200,000
< $5,000	257	193	220	6	1	10	263	194	230
About $5,000	76	84	87	3	3	3	79	87	90
DK/RF if <> $5,000	**80**	22	23	**49**	1	0	**129**	23	23
> $5,000 & < $25,000	268	272	230	4	8	6	272	280	236
About $25,000	19	24	32	0	1	2	19	25	34
DK/RF if <> $25,000	5	**52**	8	2	**44**	3	7	**96**	11
> $25,000 & < $200,000	50	66	63	0	2	4	50	68	67
About $200,000	1	1	3	0	1	1	1	2	4
DK/RF if <> $200,000	0	0	**32**	0	0	**39**	0	0	**71**
> $200,000	2	3	12	0	0	0	2	3	12
Total Case Number	758	717	710	64	61	68	822	778	778
Percentage Distribution by Bracket Category									
DK/RF to entry-bracket question	10.6%	7.3%	4.5%	76.6%	72.1%	57.4%	15.7%	12.3%	9.1%
DK/RF to non-entry-bracket question	0.7%	3.1%	4.4%	3.1%	1.6%	4.4%	0.9%	3.0%	4.4%
< $5,000	33.9%	26.9%	31.0%	9.4%	1.6%	14.7%	32.0%	24.9%	29.6%
>= $5,000 & < $25,000	45.4%	49.7%	44.6%	10.9%	18.0%	13.2%	42.7%	47.2%	41.9%
>= $25,000 & < $200,000	9.1%	12.6%	13.4%	0.0%	4.9%	8.8%	8.4%	12.0%	13.0%
>= $200,000	0.4%	0.6%	2.1%	0.0%	1.6%	1.5%	0.4%	0.6%	2.1%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Note: 1) Entry bracket categories are highlighted. 2) "DK/RF to entry-bracket question" is calculated as ratio of the highlighted cell to the total case number.
3) "DK/RF to non-entry-bracket question" is calculated as ratio of all the non-highlighted cells ending in DK/RF to the total case number.

Appendix 2, Panel 9: Other Asset Holdings
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether DK or REF
in Response to questions about Asset holdings (unweighted N)

HRS 1998: Other Assets

	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
Unfolding range	$5,000	$50,000	$100,000	$5,000	$50,000	$100,000	$5,000	$50,000	$100,000
< $5,000	24	14	18	3	1	2	27	15	20
About $5,000	11	3	9	0	0	0	11	3	9
DK/RF if <> $5,000	**18**	2	1	**26**	0	0	**44**	2	1
> $5,000 & < $50,000	39	44	40	3	3	3	42	47	43
About $50,000	6	8	7	0	3	0	6	11	7
DK/RF if <> $50,000	1	**10**	8	3	**24**	0	4	**34**	8
> $50,000 & < $100,000	3	8	9	1	2	1	4	10	10
About $100,000	2	5	4	0	0	0	2	5	4
DK/RF if <> $100,000	2	0	**13**	0	0	**23**	2	0	**36**
> $100,000	13	10	12	0	0	2	13	10	14
Total Case Number	119	104	121	36	33	31	155	137	152
Percentage Distribution by Bracket Category									
DK/RF to entry-bracket question	15.1%	9.6%	10.7%	72.2%	72.7%	74.2%	28.4%	24.8%	23.7%
DK/RF to non-entry-bracket question	2.5%	1.9%	7.4%	8.3%	0.0%	0.0%	3.9%	1.5%	5.9%
< $5,000	20.2%	13.5%	14.9%	8.3%	3.0%	6.5%	17.4%	10.9%	13.2%
>= $5,000 & < $50,000	42.0%	45.2%	40.5%	8.3%	9.1%	9.7%	34.2%	36.5%	34.2%
>= $50,000 & < $100,000	7.6%	15.4%	13.2%	2.8%	15.2%	3.2%	6.5%	15.3%	11.2%
>= $100,000	12.6%	14.4%	13.2%	0.0%	0.0%	6.5%	9.7%	10.9%	11.8%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Note: 1) Entry bracket categories are highlighted. 2) "DK/RF to entry-bracket question" is calculated as ratio of the highlighted cell to the total case number.
3) "DK/RF to non-entry-bracket question" is calculated as ratio of all the non-highlighted cells ending in DK/RF to the total case number.

Appendix 2, Panel 10: Debts
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether DK or REF
in Response to questions about Asset holdings (unweighted N)

HRS 1998: Debts

	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
Unfolding range	$500	$5,000	$50,000	$500	$5,000	$50,000	$500	$5,000	$50,000
< $500	23	20	4	1	0	1	24	20	5
About $500	7	2	6	0	1	0	7	3	6
DK/RF if <> $500	**18**	2	0	**16**	1	0	**34**	3	0
> $500 & < $5,000	47	42	38	11	5	5	58	47	43
About $5,000	15	10	9	1	1	2	16	11	11
DK/RF if <> $5,000	5	**24**	5	3	**19**	1	8	**43**	6
> $5,000 & < $50,000	41	55	41	5	4	12	46	59	53
About $50,000	2	1	3	0	0	0	2	1	3
DK/RF if <> $50,000	1	1	**10**	0	0	14	1	1	**24**
> $50,000	4	5	5	2	1	3	6	6	8
Total Case Number	163	162	121	39	32	38	202	194	159
Percentage Distribution by Bracket Category									
DK/RF to entry-bracket question	11.0%	14.8%	8.3%	41.0%	59.4%	36.8%	16.8%	22.2%	15.1%
DK/RF to non-entry-bracket question	3.7%	1.9%	4.1%	7.7%	3.1%	2.6%	4.5%	2.1%	3.8%
< $500	14.1%	12.3%	3.3%	2.6%	0.0%	2.6%	11.9%	10.3%	3.1%
>= $500 & < $5,000	33.1%	27.2%	36.4%	28.2%	18.8%	13.2%	32.2%	25.8%	30.8%
>= $5,000 & < $50,000	34.4%	40.1%	41.3%	15.4%	15.6%	36.8%	30.7%	36.1%	40.3%
>= $50,000	3.7%	3.7%	6.6%	5.1%	3.1%	7.9%	4.0%	3.6%	6.9%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Note: 1) Entry bracket categories are highlighted. 2) "DK/RF to entry-bracket question" is calculated as ratio of the highlighted cell to the total case number.
3) "DK/RF to non-entry-bracket question" is calculated as ratio of all the non-highlighted cells ending in DK/RF to the total case number.

Appendix 3, Panel 1: Value of Real Estate Holdings
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether
DK or REF in Response to questions about Asset holdings (unweighted dollars)

HRS 1998: Value of Real Estate Holdings

	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
Unfolding range	$2,500	$125,000	$500,000	$2,500	$125,000	$500,000	$2,500	$125,000	$500,000
< $2,500	1,178	771	856	1,000	0	1,000	1,160	771	870
About $2,500	2,500	2,500	2,500	0	0	0	2,500	2,500	2,500
DK/RF if <> $2,500	**197,188**	67,000	30,857	**185,500**	0	55,000	**192,179**	67,000	36,222
> $2,500 & < $125,000	43,182	40,872	43,746	28,833	40,000	38,333	42,623	40,859	43,484
About $125,000	125,000	125,000	125,000	125,000	0	125,000	125,000	125,000	125,000
DK/RF if <> $125,000	175,667	**104,196**	72,000	275,000	**183,304**	49,667	200,500	**142,909**	64,556
> $125,000 & < $500,000	196,552	232,500	245,361	0	250,000	0	196,552	234,444	245,361
About $500,000	500,000	500,000	500,000	0	0	0	500,000	500,000	500,000
DK/RF if <> $500,000	492,975	300,000	**126,882**	100,000	420,000	**318,786**	414,380	348,000	**213,548**
> $500,000 & < $1,000,000	626,250	709,091	662,500	0	750,000	800,000	626,250	712,500	677,778
About $1,000,000	1,000,000	1,000,000	1,000,000	0	0	0	1,000,000	1,000,000	1,000,000
DK/RF if <> $1,000,000	0.	725,000	750,000	0	1,875,000	0	0	1,300,000	750,000
> $1,000,000	2,065,500	1,562,500	3,062,500	2,000,000	0	3,000,000	2,043,667	1,562,500	3,050,000
All Categories	168,006	201,647	238,348	238,630	325,667	337,039	176,087	222,209	252,369

Appendix 3, Panel 2: Value of Business/Farm Asset Holding
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether
DK or REF in Response to questions about Asset holdings (unweighted dollars)

HRS 1998: Value of Business/Farm Asset Holding

Unfolding range	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
	$5,000	$10,000	$100,000	$5,000	$10,000	$100,000	$5,000	$10,000	$100,000
< $5,000	346	0	0	0	0	0	346	0	0
About $5,000	5,000	5,000	5,000	5,000	0	0	5,000	5,000	5,000
DK/RF if <> $5,000	**186,600**	0	0	**285,648**	0	0	**231,818**	0	0
> $5,000 & < $10,000	6,667	8,600	8,500	0	0	0	6,667	8,600	8,500
About $10,000	10,000	10,000	10,000	10,000	0	0	10,000	10,000	10,000
DK/RF if <> $10,000	98,333	**344,701**	10,000	50,000	**157,893**	0	86,250	**272,053**	10,000
> $10,000 & < $100,000	47,667	42,880	48,250	43,333	65,000	18,500	47,273	44,519	46,267
About $100,000	100,000	100,000	100,000	0	100,000	0	100,000	100,000	100,000
DK/RF if <> $100,000	1,256,000	520,000	**219,300**	668,667	1,000,000	**181,386**	1,035,750	580,000	**203,260**
> $100,000 & < $1,000,000	344,091	291,487	288,333	450,000	233,333	0	352,917	287,333	288,333
About $1,000,000	1,000,000	1,000,000	1,000,000	0	0	0	1,000,000	1,000,000	1,000,000
DK/RF if <> $1,000,000	500,000	700,000	0	300,000	500,000	900,000	400,000	566,667	900,000
> $1,000,000	2,816,667	4,066,667	1,920,000	0	0	2,500,000	2,816,667	4,066,667	1,972,727
All Categories	253,451	319,421	327,271	284,836	223,500	285,673	260,090	304,713	319,963

Appendix 3, Panel 3: Value of IRA Holdings
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether
DK or REF in Response to questions about Asset holdings (unweighted dollars)

HRS 1998: Value of IRA Holdings

Unfolding range	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
	$10,000	$25,000	$100,000	$10,000	$25,000	$100,000	$10,000	$25,000	$100,000
< $10,000	4,545	4,874	4,312	5,383	5,117	3,260	4,713	4,905	4,158
About $10,000	10,000	10,000	10,000	10,000	0	10,000	10,000	10,000	10,000
DK/RF if <> $10,000	**54,090**	8,900	4,000	**46,951**	10,667	9,500	**49,274**	9,657	6,750
> $10,000 & < $25,000	16,972	16,954	15,818	16,850	17,000	17,500	16,960	16,960	16,267
About $25,000	25,000	25,000	25,000	25,000	25,000	25,000	25,000	25,000	25,000
DK/RF if <> $25,000	61,111	**48,566**	35,769	115,530	**82,525**	23,446	89,753	**71,763**	29,608
> $25,000 & < $100,000	52,633	48,969	50,789	46,333	50,941	53,783	51,233	49,364	51,599
About $100,000	100,000	100,000	100,000	0	0	0	100,000	100,000	100,000
DK/RF if <> $100,000	127,000	85,500	**55,275**	94,571	53,750	**54,653**	109,539	69,625	**54,892**
> $100,000 & < $400,000	194,191	183,947	230,762	151,066	118,200	240,625	185,897	170,250	233,483
About $400,000	400,000	0	400,000	0	400,000	0	400,000	400,000	400,000
DK/RF if <> $400,000	300,000	115,000	230,000	120,000	801,667	175,000	210,000	527,000	202,500
> $400,000	955,667	557,500	1,250,000	587,500	0	708,333	863,625	557,500	925,000
All Categories	75,337	56,988	68,907	60,933	87,735	70,114	70,034	67,835	69,378

Appendix 3, Panel 4: Value of Stock Holdings
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether
DK or REF in Response to questions about Asset holdings (unweighted dollars)

HRS 1998: Value of Stock Holdings

| | DK to open-ended question | | | Refusal to open-ended question | | | DK or REF | | |
| | Random entry point: | | | Random entry point: | | | Random entry point: | | |
Unfolding range	$2,500	$25,000	$125,000	$2,500	$25,000	$125,000	$2,500	$25,000	$125,000
< $2,500	992	1,152	918	1,100	0	0	997	1,152	918
About $2,500	2,500	2,500	2,500	0	0	0	2,500	2,500	2,500
DK/RF if <> $2,500	**275,373**	8,509	625	**131,395**	6,000	0	**197,149**	8,123	625
> $2,500 & < $25,000	10,310	11,002	11,319	7,500	11,630	10,333	10,058	11,064	11,239
About $25,000	25,000	25,000	25,000	25,000	25,000	25,000	25,000	25,000	25,000
DK/RF if <> $25,000	52,842	**89,910**	35,385	103,455	**116,733**	17,500	77,048	**105,041**	34,107
> $25,000 & < $125,000	61,485	65,400	62,167	49,143	63,500	66,111	60,215	65,087	62,505
About $125,000	125,000	125,000	125,000	125,000	0	125,000	125,000	125,000	125,000
DK/RF if <> $125,000	211,100	155,000	**93,567**	323,343	204,444	**135,315**	237,002	176,191	**115,459**
> $125,000 & < $400,000	203,333	210,699	216,667	146,000	200,000	196,200	201,286	209,709	213,138
About $400,000	0	400,000	400,000	0	400,000	0	0	400,000	400,000
DK/RF if <> $400,000	368,571	250,000	320,000	0	325,000	388,571	368,571	287,500	360,000
> $400,000	1,225,294	2,112,000	1,504,332	2,177,322	700,000	570,000	1,368,098	1,847,250	1,343,240
All Categories	154,286	207,181	158,936	174,773	139,999	156,573	159,048	189,612	158,384

Appendix 3, Panel 5: Value of Checking, Savings and MM Account Holdings
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether
DK or REF in Response to questions about Asset holdings (unweighted dollars)

HRS 1998: Value of Checking, Savings and MM Account Holdings

	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
Unfolding range	$5,000	$50,000	$150,000	$5,000	$50,000	$150,000	$5,000	$50,000	$150,000
< $5,000	1,531	1,532	1,506	1,516	1,359	1,717	1,529	1,505	1,534
About $5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000
DK/RF if <> $5,000	**15,085**	8,477	7,550	**16,418**	10,293	7,835	**16,068**	9,197	7,698
> $5,000 & < $50,000	17,264	17,226	16,901	16,216	20,322	17,031	17,090	17,826	16,924
About $50,000	50,000	50,000	50,000	50,000	50,000	50,000	50,000	50,000	50,000
DK/RF if <> $50,000	33,000	**30,456**	18,268	35,900	**19,818**	16,430	34,494	**22,244**	17,486
> $50,000 & < $150,000	77,545	77,378	81,783	85,000	80,923	97,623	78,526	78,172	83,489
About $150,000	150,000	150,000	150,000	0	150,000	150,000	150,000	150,000	150,000
DK/RF if <> $150,000	122,500	352,500	**15,492**	332,500	79,200	**30,376**	192,500	200,667	**27,152**
> $150,000 & < $300,000	223,500	204,625	209,539	225,000	233,333	216,000	224,000	212,455	211,059
About $300,000	300,000	300,000	300,000	0	0	0	300,000	300,000	300,000
DK/RF if <> $300,000	200,000	211,333	212,500	0	175,000	196,750	200,000	202,250	202,000
> $300,000	0	521,091	531,250	330,000	425,000	500,000	330,000	501,873	522,727
All Categories	18,777	29,263	30,748	20,193	24,332	34,047	19,309	27,440	31,975

Appendix 3, Panel 6: Value of Bond Holdings
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether
DK or REF in Response to questions about Asset holdings (unweighted dollars)

HRS 1998: Value of Bond Holdings

	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
Unfolding range	$2,500	$10,000	$100,000	$2,500	$10,000	$100,000	$2,500	$10,000	$100,000
< $2,500	900	1,267	875	0	0	0	900	1,267	875
About $2,500	2,500	2,500	2,500	0	0	2,500	2,500	2,500	2,500
DK/RF if <> $2,500	**43,643**	1,250	9,000	**161,452**	300,000	0	**123,095**	100,833	9,000
> $2,500 & < $10,000	5,088	5,833	5,429	0	0	0	5,088	5,833	5,429
About $10,000	10,000	10,000	10,000	0	10,000	0	10,000	10,000	10,000
DK/RF if <> $10,000	85,500	**80,563**	37,333	55,000	**98,903**	0	70,250	**92,914**	37,333
> $10,000 & < $100,000	39,932	38,293	33,577	23,917	33,333	30,333	36,088	37,828	32,969
About $100,000	100,000	100,000	100,000	0	0	0	100,000	100,000	100,000
DK/RF if <> $100,000	75,200	298,333	**81,090**	0	35,000	**69,133**	75,200	166,667	**74,713**
> $100,000 & < $400,000	169,375	227,000	222,727	0	0	0	169,375	227,000	222,727
About $400,000	400,000	0	0	0	0	0	400,000	0	0
DK/RF if <> $400,000	0	0	0	0	0	158,000	0	0	158,000
> $400,000	680,000	700,000	625,000	0	500,000	0	680,000	600,000	625,000
All Categories	85,177	64,763	93,406	133,395	101,876	65,445	99,564	76,941	85,652

Appendix 3, Panel 7: Value of CDs, T-bills and Gov't Savings Bond Holdings
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether
DK or REF in Response to questions about Asset holdings (unweighted dollars)

HRS 1998: Value of CDs, T-bills, and Gov't Savings
Bond Holdings

Unfolding range	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
	$2,500	$25,000	$125,000	$2,500	$25,000	$125,000	$2,500	$25,000	$125,000
< $2,500	1,074	1,406	1,341	1,500	1,200	2,400	1,119	1,354	1,400
About $2,500	2,500	2,500	2,500	2,500	2,500	2,500	2,500	2,500	2,500
DK/RF if <> $2,500	**40,508**	13,667	11,025	**46,899**	1,000	10,000	**45,173**	8,600	10,683
> $2,500 & < $25,000	12,131	11,655	10,220	9,667	11,600	11,000	11,872	11,646	10,333
About $25,000	25,000	25,000	25,000	0	25,000	25,000	25,000	25,000	25,000
DK/RF if <> $25,000	71,400	**34,161**	41,617	22,200	**39,445**	63,333	46,800	**37,992**	47,046
> $25,000 & < $125,000	56,814	67,762	61,180	64,167	57,500	62,444	57,714	66,870	61,373
About $125,000	125,000	125,000	125,000	0	125,000	125,000	125,000	125,000	125,000
DK/RF if <> $125,000	50,000	89,200	**45,092**	35,500	44,000	**49,457**	45,857	66,600	**48,206**
> $125,000 & < $250,000	177,143	184,333	187,500	0	0	165,000	177,143	184,333	181,875
About $250,000	250,000	250,000	0	0	0	250,000	250,000	250,000	250,000
DK/RF if <> $250,000	0	0	250,000	200,000	300,000	0	200,000	300,000	250,000
> $250,000	318,333	380,562	387,500	337,500	0	400,000	323,125	380,562	391,667
All Categories	45,420	53,641	48,560	48,437	38,613	56,122	46,539	47,386	51,606

Appendix 3, Panel 8: Value of Vehicles
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether
DK or REF in Response to questions about Asset holdings (unweighted dollars)

HRS 1998: Value of Vehicles

Unfolding range	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
	$5,000	$25,000	$200,000	$5,000	$25,000	$200,000	$5,000	$25,000	$200,000
< $5,000	1,900	1,861	1,935	1,250	500	2,405	1,885	1,854	1,956
About $5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000
DK/RF if <> $5,000	**10,334**	8,823	9,683	**12,220**	0	0.	**11,051**	8,439	9,683
> $5,000 & < $25,000	12,404	13,044	13,262	14,500	11,125	13,867	12,435	12,989	13,277
About $25,000	25,000	25,000	25,000	0	25,000	25,000	25,000	25,000	25,000
DK/RF if <> $25,000	35,800	**10,354**	14,650	13,500	**18,864**	9,533	29,429	**14,254**	13,255
> $25,000 & < $200,000	50,040	47,545	44,532	0	30,000	40,000	50,040	47,029	44,261
About $200,000	200,000	200,000	200,000	0	200,000	200,000	200,000	200,000	200,000
DK/RF if <> $200,000	0	0	**9,248**	0	0	**16,885**	0	0	**13,443**
> $200,000	250,000	416,667	248,333	0	0	0	250,000	416,667	248,333
All Categories	11,709	14,292	16,524	11,036	19,992	17,932	11,656	14,739	16,647

Appendix 3, Panel 9: Value of Other Assets
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether
DK or REF in Response to questions about Asset holdings (unweighted dollars)

HRS 1998: Value of Other Assets

Unfolding range	DK to open-ended question			Refusal to open-ended question			DK or REF		
	Random entry point:			Random entry point:			Random entry point:		
	$5,000	$50,000	$100,000	$5,000	$50,000	$100,000	$5,000	$50,000	$100,000
< $5,000	2,188	1,891	2,169	2,318	3,000	1,313	2,202	1,965	2,084
About $5,000	5,000	5,000	5,000	0	0	0	5,000	5,000	5,000
DK/RF if <> $5,000	**20,111**	8,500	5,000	**44,385**	0	0	**34,455**	8,500	5,000
> $5,000 & < $50,000	14,518	20,750	19,530	20,000	19,000	21,667	14,910	20,638	19,679
About $50,000	50,000	50,000	50,000	0	50,000	0	50,000	50,000	50,000
DK/RF if <> $50,000	150,000	**58,500**	15,188	20,000	**63,875**	0	52,500	**62,294**	15,188
> $50,000 & < $100,000	75,000	69,375	68,778	85,000	68,500	0	77,500	69,200	61,900
About $100,000	100,000	100,000	100,000	0	0	0	100,000	100,000	100,000
DK/RF if <> $100,000	80,000	0	**32,092**	0	0	**45,735**	80,000	0	**40,809**
> $100,000	285,462	191,000	267,667	0	0	147,500	285,462	191,000	250,500
All Categories	48,586	47,322	49,504	37,943	56,970	45,630	46,114	49,646	48,714

Appendix 3, Panel 10: Value of Debts
Distribution of HRS 1998 sample by unfolding bracket Range, Entry Point, and whether
DK or REF in Response to questions about Asset holdings (unweighted dollars)

HRS 1998: Value of Debts

| | DK to open-ended question | | | Refusal to open-ended question | | | DK or REF | | |
| | Random entry point: | | | Random entry point: | | | Random entry point: | | |
Unfolding range	$500	$5,000	$50,000	$500	$5,000	$50,000	$500	$5,000	$50,000
< $500	193	234	182	100	0	400	190	234	225
About $500	500	500	500	0	500	0	500	500	500
DK/RF if <> $500	**8,573**	600	0	**8,695**	4,000	0	**8,631**	1,733	0
> $500 & < $5,000	1,989	1,979	1,929	1,814	1,900	1,660	1,956	1,970	1,898
About $5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000
DK/RF if <> $5,000	3,400	**4,689**	5,500	1,333	**18,271**	20,000	2,625	**10,690**	7,917
> $5,000 & < $50,000	14,988	14,661	13,929	13,200	14,500	11,808	14,793	14,650	13,449
About $50,000	50,000	50,000	50,000	0	0	0	50,000	50,000	50,000
DK/RF if <> $50,000	6,000	20,000	**4,610**	0	0	13,079	6,000	20,000	**9,550**
> $50,000	57,250	103,002	73,600	125,000	100,000	469,000[a]	79,833	102,502	186,571[a]
All Categories	7,959	10,147	10,618	12,415	16,380	35,175[a]	8,819	11,175	16,388[a]

Note: [a] One imputed outlier (6,666,733) was removed from the calculation.

References

Burkhauser, Richard V. and Paul J. Gertler (eds.), 1995. Special Issue of the Journal of Human Resources, "The Health and Retirement Study: Data Quality and Early Results."

Coder, John, and L. Scoon-Rogers. 1995. "Evaluating the Quality of Income Data Collected in the Annual Supplement to the March CPS and the SIPP," SIPP Working Paper Series No. 215. Washington, DC: U.S. Bureau of the Census.

Current Population Reports. "Consumer Income," Series P-60. Washington, DC: U.S. Bureau of the Census.

Current Population Reports. Consumer Income, Series P-60, Washington D.C., U.S. Bureau of the Census.

Curtin, R., Juster, F.T., and Morgan, J. 1989. "Survey Estimates of Wealth: An Assessment of Quality," The Measurement of Saving, Investment, and Wealth, eds. R.E. Lipsey and H.S. Tice, Chicago: University of Chicago Press.

Deaton, A. 1992. "Understanding Consumption," Oxford, U.K.: Clarendon Press.

Fay, R.E. 1996. "Alternative Paradigms for the Analysis of Imputed Survey Data," Journal of the American Statistical Association, 91, p490-499.

Ferber, R. 1959. "Collecting Financial Data by Consumer Panel Techniques," Report, University of Illinois-Urbana, Bureau of Economic and Business Research.

Ferber, Robert. 1966. "The Reliability of Consumer Reports of Financial Assets and Debts," Urbana, IL: University of Illinois Bureau of Economic and Business Research.

Greenlees, W.S., Reece, J.S., and Zieschang, K.D. 1982. "Imputation of Missing Values When the Probability of Response Depends on the Variable Being Imputed," Journal of the American Statistical Association, 77, p251-261.

Hill, D., Heeringa, H.S., and Howell, D.A. 1994. "Mixed Level Survey Measurement; Optimal Instrument Design and Efficient Analysis of Data," Unpublished paper.

Hill, Daniel H. 1999. "Unfolding Bracket Method in the Measurement of Expenditures and Wealth," In Robert J. Willis and James P. Smith (eds.), Wealth, Health and Work: Innovations in Measurement in the Social Sciences. Ann Arbor: University of Michigan Press.

Hurd, Michael D. 1990. "Research on the Elderly: Economic Status, Retirement, Consumption and Saving," Journal of Economic Literature, 27, p565-637.

Hurd, Michel D. 1999. "Anchoring and Acquiescence Bias in Measuring Assets in Household
Surveys," Journal of Risk and Uncertainty, 19, p111-136.

Hurd, Michael D., Juster, F.Thomas, and Smith, James P. 2003. "Enhancing the Quality of Data
on Income: Recent Innovations from the HRS," Journal of Human Resources, 38 (3),
p758-772.

Juster, F.Thomas, Lupton, Joseph P., Smith, James P., and Stafford, Frank. "The Decline in
Household Saving and the Wealth Effect," The Review of Economics and Statistics,
November 2005, 87 (4), p20-27.

Juster, F.T. 1977. "An Experiment in Data Collection: The Use of Financial Records," in
Proceedings of the Social Science Section, American Statistical Association, p504-510.

Juster, F.T., and J. P. Smith. 1997. "Improving the Quality of Economic Data: Lessons from
HRS and AHEAD," Journal of the American Statistical Association, 92, p1268-78.

Juster, F.T., and Kuester, K.A. 1991. "Differences in the Measurement of Wealth, Wealth
Inequality and Wealth Composition Obtained From Alternative U.S. Wealth Surveys,"
Review of Income and Wealth, 36, p33-62.

Kahneman, Daniel., and Amos Tversky. 1986. "Rational Choice and the Framing of Decisions,"
Journal of Business, 59: 251-278.

Kennickell, A.B. 1991. "Imputation of the 1989 Survey of Consumer Finances," in Proceedings
of the Section of Survey Research Methods, American Statistical Association, p1-9.

Kennickell, A.B. 1997. "Using Range Techniques With CAPI in the 1995 Survey of Consumer
Finances," Federal Reserve Board, Occasional Staff Studies-2.

Lansing, J., Ginsberg, G.B., and Braaten, K. 1961. "An Investigation of Response Error," Bureau
of Economic and Business Research, University of Illinois-Urbana.

Lillard, L., Smith, J.P., and Welch, F. 1986. "What Do We Know About Wages; The Importance
of Nonreporting and Census Imputation," Journal of Political Ecomony, 94, p489-506.

Little, R.J.A., and Rubin, D.B. 1987. Statistical Analysis With Missing Data, New York: Wiley.

Madow, M.G., Oklin, I., and Rubin, D.B. 1983. Incomplete Data in Sample Surveys. Vol.2:
Theories and Bibliographies, New York: Academic Press.

Meng, X.L. 1994. "Multiple-Imputation Inferences With Uncongenial Sources of Input,"
Statistical Science, 9, p538-573.

Moore, Jeffrey, Linda Stinson, and Edward J. Welniak. 1997. "Income Measurement Error in Surveys: A Review," Washington, DC: U.S. Bureau of the Census. S.M. 97/5.

Poterba, J.M. 1994a. "Personal Saving Behavior and Retirement Income Modeling: A Research Assessment," paper prepared for a meeting of the Committee on National Statistics Panel on Retirement Income Modeling Conference, National Academy of Sciences.

Poterba, J.M. 1994b. International Comparisons of Household Saving, Chicago: University of Chicago Press.

Radner, D. 1982. "Distribution of Family Income: Improved Estimates," Social SecurityBulletin. 45, p13-21.

Rao, J.N.K., and Shao, J. 1992. "Jackknife Variance Estimation With Survey Data Under Hot Deck Imputation," Biometrika, 79, p811-822.

Rao, J.N.K., and Shao, J. 1996. "On Variance Estimation With Imputed Survey Data," Journal of the American Statistical Association, 91, p499-506.

Rubin, D.B. 1987. "Multiple Imputation for Non-Response in Surveys," New York: Wiley.

Rubin, D.B. 1996. "Multiple Imputation After 18+ Years," Journal of the American Statistical Association, 91, p473-490.

Schieber, Sylvester J. 1995. "Why Do Pension Benefits Seem So Small," Benefits Quarterly, 11(4), p57-70.

Soest, Arthur van, and Michael Hurd. 2003. "Unfolding Brackets, Anchoring and Acquiescence Bias in Panel Models for Household Asset Holdings," Unpublished manuscript, Rand.

Sudman, S. Bradburn, N.M., and Schwarz, N. 1966. Thinking About Answers, San Francisco: Jossey-Bass.

Sudman, Sydney, and Norman Bradburn. 1974. "Response Errors in Surveys," Chicago. Aldine.

Tourangeau, Roger, Lance Rips, and Kenneth Rasinki. 2000. "The Psychology of Survey Response", Cambridge University Press, Cambridge U.K.

Tversky, Amos, and Daniel Kahneman. 1974. "Judgment Under Uncertainty: Heuristics and Bias," Science, 185, p1124-1131.

Tversky, Amos, and Daniel Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice," Science, 211, p453-458.

Woods, John R. 1996. "Pension Benefits Among the Aged: Conflicting Measures, Unequal Distributions," Social Security Bulletin, 59(3), p3-30.