Improving the Quality of Economic Data: Lessons from the HRS and AHEAD

F. Thomas JUSTER and James P. SMITH

Missing data are an increasingly important problem in economic surveys, especially when trying to measure household wealth. However, some relatively simple new survey methods such as follow-up brackets appear to appreciably improve the quality of household economic data. Brackets represent partial responses to asset questions and apparently significantly reduce item nonresponse. Brackets also provide a remedy to deal with nonignorable nonresponse bias, a critical problem with economic survey data.

KEY WORDS: Imputation; Missing data; Nonignorable nonresponse.

1. INTRODUCTION

In recent years, our understanding of what determines levels, rates of accumulation, and portfolios of wealth has greatly increased (see Hurd 1990 for an excellent summary). Encouraged by newly available data, analysts have focused theoretical models on fundamental hypotheses about why people save (Deaton 1992; Poterba 1994). Although the issues examined are extremely diverse, these models are linked by a common need: reasonably reliable wealth and savings data to test the basic implications of the models. Unfortunately, the quality of the survey data in many current wealth modules fails to meet that need.

Data quality is an issue of longstanding concern among researchers interested in wealth accumulation (Curtin, Juster, and Morgan 1988; 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 article 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 in this form to wealth modules, these extensions originated in the Panel Study of Income Dynamics (PSID) and were used extensively in the recently fielded Health and Retirement Survey (HRS) and the Asset and Health Dynamics Among the Oldest Old Survey (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 also be able to use brackets successfully by tailoring brackets to specific assets or using them judiciously. Use of follow-up brackets appears to provide a partial remedy to deal with nonignorable nonresponse 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, the failure to use brackets understated population estimates of nonhousing 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.

This article is organized as follows. Section 2 briefly describes why and how follow-up brackets are used, relying mainly on the HRS and AHEAD surveys. Section 3 documents the extent of nonresponse to asset questions and the brackets' ability to modify the consequences of initial nonresponse. Section 4 summarizes the results of our imputations for respondents with missing asset data. Section 5 contains a parallel analysis for the complete HRS and AHEAD samples. Section 6 summarizes our findings and points to directions for future research.

2. BACKGROUND

Assets are notoriously poorly reported on surveys. Nonresponse is pervasive, and other evidence (Curtin et al. 1989) suggested 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 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. 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

F. Thomas Juster is Research Scientist, Survey Research Center, University of Michigan, Ann Arbor, MI 48104. James P. Smith is Senior Economist, RAND, Santa Monica, CA 90407. This research was supported in part by grants 5PO1-AG08291 and 1RO1-AG12394 awarded by the National Institute on Aging, Department of Health and Human Services, to RAND. Excellent programming assistance was provided by Iva Maclennan.

Journal of the American Statistical Association
December 1997, Vol. 92, No. 440, Applications and Case Studies

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 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, 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.

This research relies on data from two important new surveys fielded by the Institute for Social Research at the University of Michigan. The HRS is a national sample of about 7,600 households (12,654 individuals) with at least one person in the household born between 1931 and 1941 (51–61 years old at the interview date). At baseline, an in-home face-to-face interview of some 90 minutes was conducted starting in spring 1992 and extending into early 1993. The sample was obtained by screening for age eligibility from an area probability sample of some 70,000 household address listings. Given its focus on the preretirement years, the principal objective of HRS is to monitor economic transitions in work, income, and wealth, as well as changes in many dimensions of health status.

The companion survey to the HRS—AHEAD—includes 6,052 households (8,204 individuals) from the birth cohorts of 1923 or before, thus with at least one person age 70 or older in 1993. All AHEAD-sampled households with age-eligible respondents under age 80 were obtained by screening from the HRS area probability sample. To guard against underrepresentation of the extremely disabled in an area sample, AHEAD added a supplemental sample of respondents age 80 and above from the Health Care Financing Administration (HCFA) Medicare enrollment file. The baseline AHEAD interview was conducted in 1993 using computer-assisted telephone techniques for respondents age 70–79 and computer-assisted in-person interviews for persons age 80 and older. Given its older age span, AHEAD's objectives shift toward the relationship between changes in physical and cognitive health in old age, the maintenance of independent living arrangements, and asset decumulation. In both surveys, blacks, Hispanics, and residents of Florida were oversampled at a rate of two to one. Baseline response rates were 82% in HRS and 81% in AHEAD, and each survey plans follow-ups every 2 years.

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), 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 (e.g., an amount between $5,000 and $7,499 would be represented by the letter E).

Unfolding brackets taking the form of simple questions that follow immediately after a "don't know" or "refuse" response are another device to determine the interval in which the respondent's assets lie. Because the 1984 and 1989 PSID wealth modules had used similar unfolding brackets, HRS and AHEAD were not the first studies to use follow-up brackets for wealth questions. Although PSID respondents have been interviewed every year since 1968, they may still be as sensitive to privacy concerns as other respondents.

Although SCF has historically used range cards for nonresponse, it has recently experimented with follow-up brackets. In the 1983, 1989, 1992, and 1995 SCFs, nonresponders to asset value questions were given the option of selecting category limits listed in range cards (with a more detailed list of categories in 1992 and 1995). The 1995 survey also allowed respondents to select their own limits and used the unfolding brackets techniques used in HRS and AHEAD. The Survey of Income and Program Participation (SIPP) also has made limited use of follow-up brackets to impute property income, but has not used them in its wealth module.

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. 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 features. First, unfolding brackets (is the amount more than x?) 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 to categorical data. In HRS wave 1, the strategy used in the 1984 and 1989 PSID wealth module 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 (on a voluntary basis) in the financial asset module. In later waves where telephones were the primary medium (AHEAD 1 and 2, HRS 2 and 3), range cards were not used, and all assets used unfolding brackets.

The HRS and AHEAD basic design is shown in Figure 1. First, ownership status is obtained with allowable responses of "yes," "no," "don't know," or "refuse." Next,

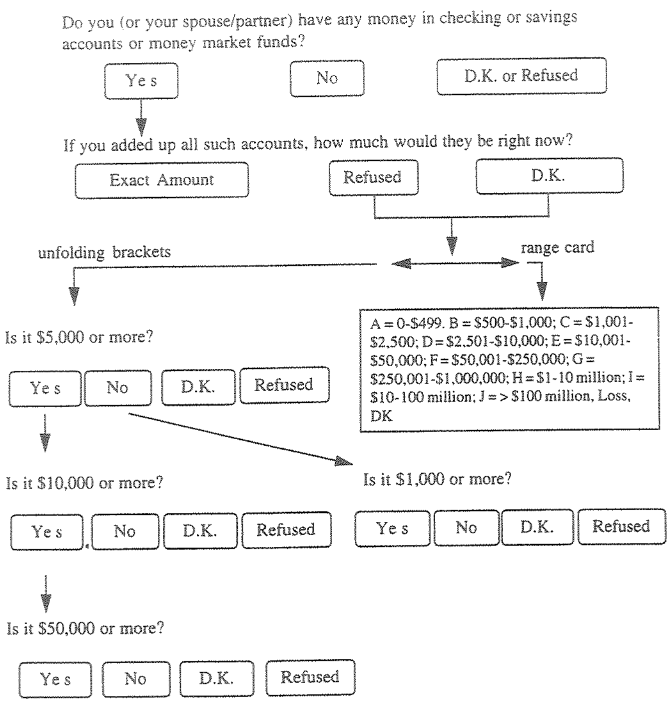

Figure 1. Illustrative Example of HRS Question Sequence on Checking Accounts, Savings Accounts, or Money Market Funds.

respondents reporting asset ownership are asked about total value; possible responses are a dollar amount, "don't know," or "refuse." Respondents in the last two categories were asked a set of "is it more than x" questions that placed their asset values within categorical limits (unfolding brackets). Anytime in this sequence of bracket questions, respondents could give a "refuse" or a "don't know" response, thus ending the sequence. For housing values, HRS wave 1 nonrespondents were shown a range card. Range cards were occasionally used instead of unfolding brackets for assets other than housing in HRS wave 1, possibly because some respondents kept the range card after it was first used and routinely used it. In addition, respondents and interviewers may have felt that range cards shortened the interview.

To be more concrete, Figure 1 displays the HRS question flow for assets held in checking and savings accounts. Imputations were needed for the relatively small number of respondents who replied either do not know or refuse to the question about asset ownership. No further questions about this asset were asked of this group. Respondents who did not report an exact amount were asked up to three questions, starting with "Is this amount greater than $5,000?", that ultimately yielded five bracket categories: $0–$999, $1,000–$4,999, $5,000–$9,999, $10,000–$49,999, and $50,000 or more. At any point in this sequence of questions, a respondent could report don't know or refuse, producing additional bracket categories (e.g., $10,000 or more). Alternatively, the range card option was used with intervals of $0–$499, $500–$1,000, $1,001–$2,500, $2,501–$10,000, $10,001–$50,000, $50,001–$250,000, $250,001–$1,000,000, 1–10 million, 10 million–100 million, and more than 100 million. The range card intervals are the same for all assets, whereas the unfolding bracket intervals varied by asset type.

Our brief summary of the history of attempts to reduce the missing data problem in household wealth surveys suggests that a variety of early and contemporary experiments have been tried in an attempt to improve data quality. In the following sections, we examine the HRS and AHEAD experience using this unfolding bracket technique.

3. MISSING VALUES AND DATA QUALITY

This section documents the ability of follow-up brackets to limit the effects of initial nonresponse. Table 1 lists the prevalence of item nonresponse in the HRS and AHEAD asset modules; exact data nonresponse is shown in column 2. Housing yields the lowest nonresponse rates, with less than 5% of HRS respondents not providing an exact home value and almost twice as many having trouble with the mortgage. Missing values are considerably larger among 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 nonresponses on the value of their businesses or stocks. In most cases, a larger fraction of AHEAD households than HRS households did not give an exact value to their assets. Among asset owners, 32% of AHEAD (28% of HRS) households did not report the exact amount in their checking and savings accounts. In general, item nonresponse 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 nonresponse rates. Severe cognitive problems were more likely to result in the use of a proxy respondent.

Nonresponse 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 SIPP; the comparable figure for the 1983 SCF was 25%. Roughly one-third of respondents in both of these surveys did not respond with an exact amount about 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 nonrespondents to asset questions are hard-nut cases, reluctant for privacy reasons to reveal their asset values. In this common view of nonresponse as dogmatic refusal, the cost of countering the initial nonresponse with more probing is thought to be high and the yield in new information low. But our experience in HRS and AHEAD suggests that convincing nonrespondents to provide bracketed responses is often easy. To illustrate, Table 1 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. The proportion of all missing data converted to range card or unfolding bracket responses is shown in the last column.

Table 1. Response Rates (Percent of Total)

Variable	No asset (1)	Owners only: Exact data report (2)	Exact data missing (3)	Range card (4)	Unfold brackets (5)	No information (6)	(4) + (5)/ 3 (7)
HRS							
House*	28	96	4	1	n/a	3	.25
First mortgage	55	92	8	1	n/a	7	.12
Other real estate	75	74	26	6	15	5	.81
Vehicles	0	86	14	3	9	2	.83
Business equity	82	68	32	5	20	7	.76
IRA and Keoghs	58	73	27	5	14	8	.78
Stocks	70	67	33	6	19	9	.73
Checking-savings	18	72	28	5	14	8	.78
CDs, treasury bills	73	70	30	6	14	10	.68
Bonds	92	69	31	6	12	13	.43
Other savings	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
First mortgage	89	86	14	n/a	13	2	.87
Other real estate	80	74	26	n/a	21	4	.84
Vehicles	0	83	17	n/a	15	2	.88
Business equity	95	59	41	n/a	36	5	.88
IRA and Keoghs	83	74	26	n/a	19	7	.73
Stocks	79	55	45	n/a	37	8	.82
Checking-savings	24	68	32	n/a	25	7	.78
CDs, treasury bills	77	62	38	n/a	28	10	.74
Bonds	92	59	41	n/a	31	10	.76
Other savings	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).

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 technique. For example, the range categories converted a 33% item nonresponse 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 nonresponse (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 nonresponse is not the same as eliminating item nonresponse 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 1 shows that brackets were even more successful in decreasing item nonresponse in AHEAD. For example, brackets converted a 45% full-item nonresponse in stock value to only 8% of cases with no information on value. On average, brackets reduced nonresponse for asset items by more than 80%, a conversion rate that exceeds even HRS. In general, full item nonresponse (no information on value) in both surveys ends up in the 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 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. 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 may learn from the bracket questions that a rough approximation to asset value is of sufficient accuracy and use 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 nonresponse. 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 translates into nonresponse. It turns out that most of these unsure respondents can be persuaded to place their asset values within range limits, information that turns out to be very valuable indeed.

Table 2 provides some insight into this issue by listing the distribution of HRS cases originally recorded as "DK" or "REF" on asset questions. Respondents who went com-

Table 2. Bracket Response Distributions, HRS Data (% of Total)

Asset type	Don't know response: Complete bracket	Partial bracket	DK Ref	Refusal response: Complete bracket	Partial bracket	DK Ref
Real estate	84	6	10	40	11	49
Vehicles	89	3	8	32	3	65
Business equity	83	4	14	41	7	53
IRA and Keoghs	82	7	10	42	12	46
Stock	82	6	12	34	12	54
Checking and savings	85	7	8	45	11	46
CD's, treasury bills	80	10	11	35	8	58
Bonds	71	5	24	19	7	74
Other	83	5	13	28	5	67

pletely 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 between 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 typically 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.

4. IMPUTATION OF MISSING VALUES-METHODS

Follow-up bracket questions persuaded many initial nonrespondents 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, exact-answer cases 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.

Although many imputation procedures are available, not all are equally appropriate for assets, where the distribution of possible values is inherently constrained. For example, most asset values cannot be less than 0, and the frequent use of range values imposes varying upper and lower bounds. To maintain the distribution in the population, imputation should also have variance–preserving properties.

In this article we use hot-deck algorithms to assign missing values. Like any method, hot-deck imputation has advantages and disadvantages (see Little-Rubin 1987). Its advantage is the absence of parametric assumptions about unobservables, relying implicitly on the functional form among the donors. A principal disadvantage is that it reduces to a fully saturated ANOVA model so that donor samples can quickly become thin if many covariates are used.

An important distinction in missing-data models is between ignorable and nonignorable nonresponse (Little and Rubin 1987). When nonresponse is nonignorable, there is response bias, because observationally equivalent respondents and nonrespondents will have different distributions of missing values. Although significant advances have been made recently on ignorable models for missing data (Fay 1996; Rubin 1996), progress on nonignorable nonresponse has been stymied, because inferences are sensitive to untestable assumptions about what generates differences between respondents and nonrespondents (Lillard, Smith, and Welch 1986). The difficulty with nonignorable response models is that there rarely is an external way of testing model assumptions. An important exception is the work of Greenlees, Reece, and Zieschang (1982), which used the CPS-IRS-SSA exact match file to model nonresponse in wages. While having some external source for missing values is unusual, follow-up brackets can be viewed as a self-reported external source for missing values. Instead of providing an exact number for the missing value (as with IRS-SSA values for missing CPS wages), follow-up brackets provide a range within which the missing value must lie. For this reason, follow-up brackets provide mileage in dealing with nonignorable nonresponse, an especially important issue for wealth.

There exists nontrivial estimation error in any imputation procedure, and hot deck is certainly no exception. The estimation error associated with any single set of imputations can make it difficult to distinguish with any confidence between differences that emerge because of bias and those due to random imputation estimation error. Multiple imputation methods (Rao 1996; Rao and Shao 1992; Rubin 1987), which acknowledge this underlying uncertainty by providing multiple estimates, can also be used productively with follow-up brackets. We minimized the effect of imputation estimation error by computing all hot-deck imputations across 25 independent trials. The means across these 25 trials are the imputations presented in this article. Because estimation error is proportional to the square root of N, estimation error associated with imputation should be reduced five-fold. The differences reported herein are simply far too large to be due to imputation estimation error.

As suggested by Figure 1, there are three distinct groups of respondents with missing values for whom imputations must be obtained. The first relatively small group involves those who did not know or refused to say whether they had the asset at all. For this group, hot-deck imputations are made using the complete sample—owners and nonowners of the assets—as potential donors. Given the small number of respondents in this group, how it is treated is not very important, and we ignore it in this article.

Brackets are critical, however, when assigning missing values to the other two groups. Our second group includes those who selected a bracket range for their asset value. The direct benefit from brackets is that they place an otherwise unknown asset value within a prespecified range. In making imputations for bracket respondents, we used as potential donors those exact-answer respondents whose asset values fell within the respondent's bracket range.

Our final group are those who provided neither an exact value nor a bracketed value for their assets. Brackets have the indirect benefit of providing a more relevant pool of donors for these cases. To make imputations for these complete nonresponse cases, only the pool of bracket responses is used as a source of potential donors, in contrast to standard imputation methods that traditionally use the entire sample. We would argue that respondents who use brackets are a more representative donor pool for complete nonresponse cases, because they share a common initial reluctance to answer wealth questions.

The following subsections summarize our imputations for nonhousing wealth, separately for bracket response cases and complete nonresponse cases. Section 5 combines these subsamples to derive the implications for the complete HRS and AHEAD samples. Whether brackets are used or not, we use an identical list of personal attributes for each hot-deck imputation in the article. The mose extensive list of covariates used included rate, ethnicity (Hispanic), education (0–11, 12–15, 16 or more years), married, law, M.D. or doctorate degree, and dummy variables for quintiles of household income and quintiles of housing equity (for nonhousing assets). As with most hot-deck procedures, the coarseness of the matching variables depends on the ability to find matches (Rubin 1987). The actual covariates used to impute specific assets vary depending on range width and thinness of the sample of potential donors.

4.1 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 conventional procedure—all exact-answer respondents serve as potential donors. In both cases the full list of personal attributes described earlier is used in the imputation algorithm. Table 3 shows means and medians (averaged over 25 iterations) for each nonhousing asset. The

Table 3. The Effect of Ignoring Brackets for Imputation of Missing Nonhousing Values Among Respondents Providing Bracketed Responses

	HRS brackets		AHEAD brackets	
Asset	Used	Ignored	Used	Ignored
Mean Values				
Real estate	221,676	129,098	146,149	107,472
Vehicles	18,079	12,539	6,606	6,141
Business	348,600	165,986	219,580	99,872
IRA, Keoghs	56,415	44,357	55,110	52,608
Stocks	74,736	56,982	104,694	74,866
Checking, savings	23,409	16,014	21,648	20,750
CDs, treasury	47,665	27,253	34,823	39,852
Bonds	67,846	47,447	90,208	54,275
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
Vehicles	10,000	7,800	4,272	2,024
Business	98,000	24,260	110,400	15,960
IRA, Keoghs	30,000	20,080	24,500	23,224
Stocks	22,928	17,017	39,340	27,760
Checking, savings	6,672	5,000	7,780	6,320
CDs, treasury	10,000	9,760	13,440	19,400
Bonds	24,220	14,340	41,540	31,060
Other assets	20,000	13,980	8,050	9,950
Other debts	−2,918	−2,544	−2,000	−1,554

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 3 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 they 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 two 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 do, especially at very high asset values. Thus it is possible that the unfolding bracket categories may still understate respondents' asset values.

Our reliance on hot-deck methods is not an endorsement, because these methods are not inherently superior to the alternatives. We use them here to facilitate comparisons with Census imputation of missing wealth data, which relies almost exclusively on hot-deck methods. How sensitive are our main conclusions to our reliance on hot decks? The relatively few covariates in the hot-deck model may imply that correlations may be preserved only among a small subset of variables. In particular, our conclusions on bias may be sensitive to not including the values of other assets of respondents. For example, an excellent predictor of respondents' stock holdings may be the value of their bonds. The list of possible household assets is far too large to fit into a hot-deck procedure, so more explicit model-based approaches must be used.

To explore this issue, Table 4 lists two additional sets of HRS imputations alongside our hot-deck estimates, which are repeated in the second column. The third column contains mean imputations obtained from an ordinary least squares (OLS) regression model with the same list of covariates in the hot-deck model. The fourth column is also derived from an OLS regression on asset values, but now the covariates are expanded to include the presence and dollar value of each other asset held by the respondent. In both new models, predicted values are augmented by a random selection from the residual variance (using the estimated mean squared error from the regressions). To reduce the effect of imputation error, 25 different independent draws from this residual distribution are made. The numbers presented for the new models in Table 4 are also average values across the 25 iterations.

Table 4. Mean HRS Asset Imputations Under Three Alternative Models

Asset	Hot-decks (brackets-used)	OLS-simple	OLS-extended
Real estate	221,676	130,958	148,129
Vehicles	18,079	12,755	12,948
Business	348,600	133,206	160,825
IRA, Keoghs	56,415	43,544	45,627
Stocks	74,736	52,621	68,700
Checking, savings	23,409	15,086	16,127
CD's, treasury	47,665	24,075	29,589
Bonds	67,846	42,767	72,525
Other assets	78,711	47,011	50,646
Other debts	−7,118	−9,797	−9,964
Average value	75,647	42,560	49,949

Comparing the third columns of Tables 3 and 4, mean imputed values for bracketed cases are quite similar. This similarity implies that our imputations may not be overly sensitive to the type of imputation model (hot deck or regression) when the same covariates are used. With one notable exception, expanding the covariate list to include other assets does not have a great impact on the extent of nonignorable nonresponse bias, largely because exact dollar and bracketed respondents do not differ a great deal in their asset portfolios. The exception relates to stocks and bonds where mutual knowledge about their coexistence and values significantly raises estimates of missing values. The principal advantage of enlarging the list of covariates to include other assets is that it improves predictions of within-group allocations of asset values. Averaged across all assets, R^2 in the regressions average about .22 with the expanded covariate list, compared to .10 with the more limited list. Better within-group predictions are a good enough reason to include other assets in any imputation algorithm, but they apparently are not a substitute for brackets in dealing with nonignorable nonresponse bias. (See Kennickell 1979 for an application to SCF.)

The problems entailed in accurate imputation when brackets are unavailable are succinctly summarized by calculating the percentage of cases in which imputation without brackets assigns values outside respondents' self-reported range. Table 5 indicates that only 35% of HRS missing values are correctly assigned to respondents' self-reported brackets when brackets are ignored. Ignoring

Table 5. Fraction of HRS Misassigned Cases When Brackets Are Ignored

	Below brackets	*Within brackets*	*Above brackets*
All Cases	27.3	35.4	37.1
0–$5,000		16.7	83.3
Above $500,000	88.7	11.3	

brackets has particularly severe consequences at the extremes of the distribution. Among those HRS respondents who said that an asset was worth more then $500,000, standard imputation using personal attributes predicted a value below that threshold in 89% of the cases. Similarly, when respondents indicated that an asset was less than $5,000, a value larger than that threshold was assigned in 83% of the cases.

4.2 Bracket Respondents: Refusals Versus Don't Knows

We argued earlier that refusal (REF) and don't know (DK) responses may have different motives and consequently different distributions of asset values, other things equal. If refusals stem largely from the size of assets, then the imputed assets of REF cases will be greater than those of DK. Table 6 provides some evidence on this distinction by listing mean imputed assets separately for REF and DK respondents who gave unfolding bracket values. To conserve space, details for selected specific assets are provided for the HRS sample only.

REF responses show much higher asset values than DK responses. These sharp differences suggest that the distinction between REF and DK should be recorded in public use tapes to help researchers in making their own imputations. Whether a respondent was a REF or DK case was used as an attribute for all imputations in this article.

4.3 Final Nonresponse 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 nonresponse cases. Table 7 lists imputed mean values for all "final nonresponse 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 responses. We believe that the latter is more representative of the final nonresponse 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 nonresponses, who are even more reluctant than bracket respondents to reveal their assets.

Table 7 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 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.

5. 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 those who did not possess the asset. Table 8 summarizes the effect of using brackets on total asset values for the complete HRS and AHEAD samples. In the full HRS sample, mean nonhousing 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 and savings accounts. The size of the discrepancy in the full sample varies with the type of asset. Although there is little difference in housing equity, use of brackets increases total business and real estate equity by 37% and total financial assets by 17%.

Because they vary systematically with age, these discrepancies may affect our views on such basic questions as the adequacy of savings for future retirement. Table 8 shows that mean nonhousing wealth is 9% larger when brackets are used in AHEAD. Although this is a nontrivial effect, it is much smaller than the 19% reported for HRS. On the basis of our estimates from these two samples, wealth imputations without brackets may understate by roughly 10% the asset holdings of those in their 50s relative to those age 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 tests of the size of the bias across age groups should be conducted.

To this point, our discussion has concentrated solely on nonhousing assets. Nonresponse on housing was less severe in both surveys, so the ultimate impact of missing data on total household wealth is much smaller. In contrast to other assets, mean imputed home equity is little different in AHEAD and is actually slightly smaller in HRS when range

Table 6. Means of Nonhousing Assets for Unfolding Bracket Responses

	All unfolding brackets	*Refusal unfolding brackets*	*Don't Know unfolding brackets*
HRS assets			
Other real estate	184,458	254,047	176,033
Equity in business	361,009	448,285	351,145
IRA, Keoghs	49,360	66,013	43,781
Stocks, mutual funds	68,098	133,428	58,735
Checking, savings	19,502	23,908	17,619
CD's, treasury	32,389	45,021	28,414
Bonds	70,823	54,509	74,754
Other assets	52,071	78,788	48,743

Table 7. Imputation of Mean Nonhousing Values for "Final Nonresponses"

	HRS donor pool		AHEAD pool	
	Exact data responses	Bracket responses	Exact data responses	Bracket responses
Other real estate	109,449	226,308	91,108	165,454
Vehicles	11,209	20,684	7,774	5,697
Equity in business	280,105	413,221	43,429	251,780
IRA, Keoghs	37,554	61,272	46,554	47,555
Stocks, mutual funds	63,258	98,055	75,891	153,968
Checking, savings	16,823	24,585	20,880	23,571
CD's, Treasury	24,805	46,259	38,399	33,204
Bonds	45,681	51,747	50,322	92,842
Other assets	51,683	106,653	36,883	20,479
Other debts	−6,665	−7,170	−6,211	−4,635
Average value	44,185	72,118	40,297	57,156

card respondents are used as the donors. The HRS result is a direct consequence of the distribution of range card responses compared to the distribution of respondents with exact data. HRS nonresponse cases who gave a range card answer were only one-third as likely as exact data cases to own a house worth more than $250,000. Across many stratifications of economic status, including education and income, the less well-off are less likely to report their house value. For example, 7% of HRS households in the bottom family income quintile do not report house values, compared to 3% in the top quintile.

Why is the missing-data pattern for housing so different from all other missing asset values? Particularly among respondents who have lived in their current home for many years, unwillingness to report housing values may reflect uncertainty rather than sensitivity about value, especially among less-educated and lower-income respondents. In addition, housing prices exhibit significant regional variation, introducing disparities in the norm of what constitutes an expensive home. Respondent sensitivity about value may exist only relative to this norm, blurring the simple relation of reporting to actual value.

6. REPORT ON SOME EXTENSIONS AND CONCLUSIONS

Although unfolding brackets can improve the quality of financial data, research on their optimal design and implementation is just at the beginning stages. Three issues should be placed particularly high on the data quality research agenda (Sudman, Bradburn, and Schwarz 1966). These issues are complex and in need of additional research; their potential importance is briefly sketched here.

The first question concerns how to "optimally" select bracket categories. Practical survey constraints will always limit brackets to a small number of distinct categories. However, the breakpoints selected can still be chosen to maximize their predictive power in imputation. For example, if most of the missing cases had values over $1 million, then the chosen brackets should be concentrated above that number. Because their empirical distributions are so diverse, the "best" brackets will vary across individual assets. The general idea behind selecting "optimal" breakpoints is that the chosen thresholds should maximize explained variance in a one-way ANOVA (Hill, Heeringa, and Howell 1994). The breakpoints actually used in HRS and AHEAD baselines were essentially the same as those in the 1984 and 1989 PSID wealth modules. Based on the optimal breakpoint strategy, these breakpoints were revised in subsequent waves of HRS and AHEAD. For example, the HRS wave-1 brackets for investment real estate were $1–$4,999; $5,000–$49,999; $50,000–$149,999, and $150,000 or more. For wave 2, the analysis yielded optimized brackets of $1–$2,499; $2,500–$124,999; $125,000–$499,999; $500,000–$999,999 and $1 million or more. Essentially, the optimal brackets involved finer partitions for very small and very large real estate values where many observations were concentrated.

Even if the best set of bracket thresholds are chosen, the issue of whether there exists an anchoring effect associated with the choice of an initial threshold in the sequence remains. Anchoring occurs when the content of the question itself conveys information about what the probable "correct" answer is. For example, if respondents are asked about the size of their checking accounts, responses may be influenced by whether the first question is at the $100 level, the

Table 8. Value of Assets in Full HRS and AHEAD Samples

	HRS		AHEAD	
	Imputations using brackets	Imputations ignoring brackets	Imputations using brackets	Imputations ignoring brackets
Housing	75,459	75,864	66,882	66,705
Nonhousing	162,253	136,904	100,583	91,694
All assets	237,712	212,768	167,465	158,399

\$1,000, or the \$10,000 level, even if the final set of bracket categories are the same. Because respondents may assume that question designers know more than they do, the entry point may tell respondents something about what the "correct" answer is. A sequence that starts with \$100 will convey the impression that small numbers are more likely to be correct than large numbers, whereas a sequence starting with \$10,000 may give the opposite impression.

To address this question, a group of respondents in the second wave of AHEAD were asked to place their savings account values into bracket thresholds. While the final set of thresholds were the same, the initial threshold value varied randomly across respondents. The cumulative distribution of savings account values varied systematically with alternative initial entry points. For example, the cumulative fraction of cases less than \$10,000 was 49% when the initial entry point was \$1,000 compared to 37% when it was \$20,000. Anchoring effects produced less bias in mean values when the initial entry point was in the middle rather than at either end of the distribution. Because most HRS and AHEAD bracket sequences start toward the middle of the distribution, the bias in mean values in these surveys may be moderate.

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 about equal to x?" 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 about equal to x) provides a somewhat different distribution of responses, with about 5–10% of respondents reporting that "about equal to x" is the correct answer.

7. CONCLUSIONS

This article has investigated some survey techniques used in the HRS and AHEAD surveys. These techniques—follow-up bracket responses—reduce the implications of initial nonresponse to wealth questions and narrow uncertainty about precise asset values. Because initial levels of item nonresponse in HRS and AHEAD are similar to those obtained in other household surveys, follow-up brackets may also lower the pervasiveness of complete item nonresponse in other surveys.

The potential value to other household surveys of follow-up brackets goes beyond simply reducing nonresponse. 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 nonresponse. 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.

[Received May 1995. Revised May 1997.]

REFERENCES

Curtin, R., Juster, F. T., and Morgan, J. (1989), "Survey Estimates of Wealth: An Assessment of Quality," in *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, 490–499.

Ferber, R. (1959), "Collecting Financial Data by Consumer Panel Techniques," report, University of Illinois-Urbana, 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, 251–261.

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

Hurd, M. D. (1990), "Research on the Elderly: Economic Status, Retirement, Consumption and Saving," *Journal of Economic Literature*, 27, 565–637.

Juster, F. T. (1977), "An Experiment in Data Collection: The Use of Financial Records," in *Proceedings of the Social Science Section, American Statistical Association*, pp. 504–510.

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, 33–62.

Kennickell, A. B. (1991), "Imputation of the 1989 Survey of Consumer Finances," in *Proceedings of the Section of Survey Research Methods*, American Statistical Association, pp. 1–9.

——— (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, June 1961.

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 Economy*, 94, 489–506.

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

Madow, W. 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, 538–573.

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

——— (1994b), *International Comparisons of Household Saving*, Chicago: University of Chicago Press.

Rao, J. N. K., and Shao, J. (1992), "Jackknife Variance Estimation With Survey Data Under Hot Deck Imputation," *Biometrika*, 79, 811–822.

——— (1996), "On Variance Estimation With Imputed Survey Data," *Journal of the American Statistical Association*, 91, 499–506.

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

——— (1996), "Multiple Imputation After 18+ Years," *Journal of the American Statistical Association*, 91, 473–490.

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