

HRS

HEALTH AND RETIREMENT STUDY
A Longitudinal Study of Health, Retirement, and Aging
Sponsored by the National Institute on Aging

Unfolding brackets for reducing item nonresponse in economic surveys

Documentation Report

*Head, Sampling Section,

**Associate Research Scientist & Project Manager,

***Programmer Analyst

Steven G. Heeringa*

Daniel H. Hill**

David A. Howell***

Survey Research Center
Institute for Social Research
University of Michigan
Ann Arbor, Michigan

1995

Funding

The Health and Retirement Study is funded by a grant from the National Institute on Aging (U01 AG009740) with supplemental support from the Social Security Administration. HRS is conducted by the University of Michigan.

This research was sponsored, in part, by NIA Grant P01AG10179 as well.

Suggested Citation

Heeringa, S. G., Hill, D. H., & Howell, D. A. (1995). Unfolding brackets for reducing item nonresponse in economic surveys. University of Michigan.
<https://hrs.isr.umich.edu/publications/biblio/10985>

Acknowledgements

The authors would like to thank Tom Juster, Michael Hurd, Jim Smith and Bob Willis for their many helpful comments on earlier related work. Any errors are the responsibility of the authors.

Abstract

This paper describes and analyzes a new survey methodology for reducing item non-response on financial measures. This "unfolding bracket" method is systematic and applicable in both face-to-face and telephone surveys. The proportion of missing observations for financial variables in national surveys is often in the 20-25% range and in some cases is as high as a third. With the unfolding bracket method the proportion of completely missing data can be cut by two-thirds. Furthermore, with appropriately chosen bracket breakpoints, the amount of the variance in the underlying measure recovered is quite high. We propose and demonstrate one method for choosing the breakpoints which employs the Downhill Simplex algorithm to maximize their explanatory value. Additionally, use of a Box-Cox transform of the actual data in conjunction with this algorithm, can result in breakpoints which are effective in explaining most of the underlying variance in both actual values and their log transforms. Since each of these metrics is appropriate for some uses this compromise is quite useful in meeting the needs of a wide variety of potential users. Finally, we investigate the effects of bracketing on the empirical validity of survey data. While we do find lower empirical validity for data from individuals exposed to brackets early in the survey instrument, this appears to be the result of self-selection rather than a direct effect of exposure to the methodology.

Key Words: Economic Surveys; Item Nonresponse; Missing Data

Unfolding Brackets for Reducing Item Nonresponse in Economic Surveys

I. Introduction

Survey questions that ask respondents to report amounts -- particularly dollar values for financial variables such as income and assets, liabilities, transfers -- are subject to high rates of item missing data. (Juster and Smith, 1994). As an alternative to simply accepting high rates of item missing data for financial variables, researchers are making increased use of special questionnaire formats that are designed to collect an interval-scale observation whenever a respondent is unable or unwilling to provide an exact response to a financial amount question. (Juster, Heeringa, and Woodburn 1992). Loosely termed "bracketing questions", these new question formats are a type of the more general class of unfolding question sequences that are developed for improving survey measurements of complex characteristics.

The use of interval scale measures for financial items is not new to survey research. The simple income questions included in many survey questionnaires are often designed to measure amounts on an interval scale (e.g. \$0-4999, \$5000-\$14,999, etc.). In face-to-face interview situations, "show cards" or other visual devices enable respondents to map an underlying cardinal-valued response item onto an interval or ordinal scale. In surveys where cardinal-scale measurement of financial variables is necessary or preferred, the Survey Research Center (SRC) has historically provided its interviewers with a "range card" which enabled them to record an interval scale response code for cardinal scale items. Unlike show cards, the range card was not designed to be used each time the question was asked but served as an interviewer aid in

cases where it was clear that the respondent would not report an actual amount. To avoid confusion on the part of the interviewer, a single set of fixed range card categories was applied to all financial measures regardless of their underlying distribution in the population. In large part, the frequency and accuracy of range card responses to financial amount items was determined by the individual interviewer.

Bracketing question sequences for measuring financial variables first appeared in the special wealth supplement to the 1984 Panel Study of Income Dynamics (PSID), see Curtin, Juster and Morgan (1989). Bracketed measurement of 1984 PSID households' financial assets served to: 1) standardize the process for recovering interval scale observations for missing amounts; 2) adapt the interval scales to the population distribution for the financial variable of interest; and 3) enable the collection of interval scale measures in a telephone interview format. The use of bracketing question sequences was repeated in the 1989 and 1993 wealth supplements to the PSID. This paper will draw heavily on data and field experience with bracketed question items used in the Health and Retirement Survey (HRS). HRS questions on important household assets are specially designed to recover interval valued data whenever the respondent refuses or is unable to report actual amounts. Through the use of special question formats the rate of completely missing data for HRS asset amount variables is significantly reduced; however, the resulting measures are a mixture of single valued responses, "bracketed" or interval valued responses, and completely missing data.

II. Background

As noted in the introduction, financial surveys are particularly apt to encounter serious

item non-response. Table 1, adapted from Juster and Smith, 1994, shows the item nonresponse rates for six financial variables obtained in five major national studies with substantial financial sequences. Overall, the item missing data rates are highest for financial assets such as "Checking and Savings Accounts" and "Stocks Bonds and Trusts" where roughly a quarter of the reports are missing in the 1981 National Longitudinal Survey (NLS) and in the 1979 Retirement History Survey (RHS). Fully a third of the 1984 SIPP observations on the value of real estate other than the primary home were missing. The item nonresponse rates are lowest for equity in primary residence and in the amount of consumer debt, especially in the 1979 RHS and the 1989 Survey of Consumer Finances (SCF). The 1989 SCF did use the range card response option described in Section I, above. Largely as a consequence of the unfolding bracket method, the 1992 HRS has item missing data rates on these financial components which are half to one-fourth as large as comparable items in the NLS, RHS, and the SIPP.

[Table 1]

II.B Bracketing of Amounts

Figure 1 illustrates the format of the bracketing question sequence for two asset items: equity in a business and combined value of IRA and Keogh accounts. For these and seven other key asset items, if a respondent could not recall or refused to report the exact value for the item, the HRS Wave 1 questionnaire followed up with a short sequence of questions designed to "bracket" the true response value. The question sequences open by asking if the household owns the asset in question (e.g., a business). If the asset is owned, its exact value is requested. If the exact value is not reported, the questionnaire routes the respondent through a series of dichotomous response questions which attempt to "bracket" the value of the asset. Taking the

business asset and IRA/KEOGH account value question sequences as examples, the finest level of bracketing attainable through the questions is shown in Table 2 below.

[Table 2]

Routing the respondent through the nested series of bracketing questions does not guarantee that a specific bracket will be identified for the unreported amount. In some cases, no additional information will be obtained. In other cases, the responses will indicate that the true value lies in one of three brackets, but not precisely which of the three brackets. By example, a respondent may indicate that the value of their IRA or Keogh account is \geq \$25,000 but cannot/will not indicate if it is \$25,000-\$49,999, \$50,000-\$99,999, or \$100,000+.

Table 3 summarizes the data problem for each of the nine household assets. The left-hand panel of Table 3 identifies the individual asset (A) components in question. The central panel, labeled "Does item apply?", provides estimates of the percentage of HRS sample households (unweighted) that reported having each asset (i.e., a nonzero amount value is assumed). For example, of the $n=7608$ respondent households included in this summary, 23.1% report owning real estate other than their personal residence. For households that report owning a particular asset or having a particular type of debt, the right-hand panel of Table 2 describes the distribution of response types: actual value, bracketed value,¹ range card value, or missing data value.

[Table 3]

Among financial assets, the percentage of actual value reports ranges from 67.4% for

¹The bracketed value category includes cases in which, due to nonresponse or uncertainty, the boundary values for the amount may span two or three of the actual bracket ranges for the item question.

stocks and mutual funds to 87.4% for combined value of vehicles and other personal property. Depending on the asset, the percentage of bracketed responses ranges from 8.2% for property to 21.3% for business value. Even though a bracketing question sequence was provided for these asset items, from 2.4% to 6.4% of bounded response values were recorded as choices from the range card. The rates of completely missing data -- proportions of cases where no real information on bounding values is available -- range from 1.9% of responses for the vehicle and property question to 10.6% for value of bonds.

III Variance Recovered with Brackets

Clearly, the bracketing method is quite effective in reducing completely missing reports of financial variables. The question remains, however, as to how effectively the bracketed responses are increasing our information on the underlying financial variable. A relatively straight forward way of assessing this is to see how much of the variance of exact reports is explained when we partition them via the brackets. This can be estimated using ANOVA and is the ratio of between-bracket sum of squares to the total sum of squares. Of course, the answer will depend on the metric used for the observed data. For some analytic purposes (e.g. accounting) the appropriate metric is simply the level of the asset or income component, whereas for others (e.g. economic behavioral modeling) the most appropriate metric will be its natural logarithm. Table 4 presents the ANOVA R^2 s obtained for actual observations and log-transforms of nine net worth components measured in the first wave of the HRS. The degrees of freedom which corresponds to the number of breakpoint questions employed (i.e. the number of brackets minus 1) are also provided in Table 4.

[Table 4]

It is quite apparent from Table 4 that even with only three or four breakpoints, substantial fractions of the total variance in the underlying variable can be explained. The brackets for Wave 1 of the HRS appear to have been set so as to maximize the amount of variance of logarithms components in mind since more than seventy percent of the variance in log-levels is explained for each net worth component. These same brackets explain generally less of the variance in asset levels, and for some components (i.e. vehicles, savings accounts and other assets) the amount of variance explained by the brackets is quite small.

The more important financial studies are longitudinal (e.g. HRS, PSID, SIPP) and much of their analytic power comes from their ability to measure or model changes in financial measures. It is quite possible, in theory, for a given set of bracket values to do a very good job in explaining cross-sectional measures of the levels or log-levels of financial measures and yet do a poor job of capturing wave-to-wave change in these same measures.

Careful consideration of the types of wave-to-wave changes in types of reports which are possible, however, reveals that this could only occur if substantial numbers of respondents provide bracket responses in two consecutive waves of a panel. For respondents who provide exact reports (including "don't own") in both waves the bracket breakpoints are irrelevant. Theoretically, for those respondents with extra-marginal changes (i.e. from not owning to owning or visa versa) brackets which do well in explaining levels or log-levels will do exactly as well explaining change in levels or log-levels--in this case the change is identical to the level. Changes of this sort will tend to be larger relative to intra-marginal changes and will tend to dominate overall longitudinal change. For those respondents providing exact reports in one

wave and a bracket report in the other, good brackets for cross-sectional observations will also be good brackets for analyzing change.

Table 5 presents report types for two net worth components from the 1984 and 1989 PSID. It is clear for assets in the form of Real-estate and Business of Farm reports involving brackets are predominantly extra-marginal in nature. Therefore, in practice it appears that for bracket breakpoints which are optimized for cross-sectional measures will also be near-optimal for longitudinal measurement as well.

IV. Optimal Bracket Breakpoints

From Table 4, above, it is clear that there can be wide variation in how well the brackets recover information about the missing observations. The R^2 's there varied from a low of 10.8% for the value of vehicles and personal property to a high of 90.5% for the value of certificates of deposit. This variance is due to variation in the empirical distributions and the number and precise placement of the breakpoints defining the brackets. In this section we will present a method of setting the breakpoints in such a way as to maximize their explanatory power. The method presented here presumes that micro-level data on the variable to be bracket is available.

To see how optimal breakpoints can be constructed let us assume that we have N_e "exact" observations of the variable of interest y . We can express the within group sum of squares as a function of a vector of breakpoints (β) defining a set of brackets as:

$$WSS = WSS(\beta) = \sum_i \sum_j (y_{ij} - \bar{y}_{\beta_j(i)})^2$$

where $\bar{y}_{\beta_j(i)}$ is the mean of the exact reports in the interval β_j to β_{j+1} . Assuming that the underlying distribution of the missing reports is the same as the exact reports, optimal

breakpoints can, in theory, be obtained by setting them in such a way as to minimize WSS.² Since WSS is not differentiable β (or even continuous), optimization requires a non-Newtonian computer intensive method such as the DownHill Simplex which we will discuss shortly.

The question of which metric to use for y is an important one which will depend on the intended analytic uses of the final data. If variation at the top of the distribution is important then it is generally best to optimize (1) using levels of y , whereas if variation at the bottom is important then log-levels is the better choice. If there are a number of intended uses then we would want to chose β which do a good job in explaining both levels and log-levels. We can imagine minimizing (1) twice--once for levels and once for log-levels--and then setting β at some sort of mid-point of the two optimal vectors. Such a procedure could be tedious, however, since with finite (or even small) N_e , $WSS(\beta)$ is not always well behaved. If the y_i are 'lumpy' then $WSS(\beta)$ can be quite sensitive to small changes in the β_j and finding a good compromise may require repeated trial and error calculations.

An alternative procedure for finding optimal breakpoints which also provides a good compromise between levels and log-levels is to employ the Box-Cox transform of y :

$$y_i^* = \frac{y_i^\lambda - 1}{\lambda}$$

As $\lambda \rightarrow 0$, $y^* \rightarrow \ln(y)$, whereas as $\lambda \rightarrow 1$, $y^* \rightarrow y-1$. If we use y^* in place of y in minimizing (1), λ can be varied to attain a set of breakpoints which yields an acceptable

²This assumption is equivalent to assuming that the data are coarsened at random (see Heitjan and Rubin, 1991). Elsewhere we have found evidence that this assumption is not true for most financial items in the HRS--reports are more apt to be missing for wealthier respondents. Never the less, the MCAR assumption is a good first approximation for setting breakpoints.

goodness of fit for both levels and log-levels.

As noted above, since WSS is non-differentiable in the bracket breakpoints (β) we need an optimization routine which does not rely on gradient information. The Downhill simplex algorithm provided by Flannery, et al., 1989, pp 326-330 is one of the most robust and efficient. We know of no better explanation of the method than theirs and we refer the interested reader to it. To use their algorithm for our purposes we must adapt it to the appropriate dimensionality and program the objective function to be minimized. In our case this is given by Equation (1) (Pascal source code is available from the authors upon request).

The major complication of our application over that presented by Flannery et al. is that we also wish to optimize over λ --the Box-Cox parameter. We can imagine a composite objective function $F(WSS_x, WSS_{\ln(x)})$ which is implicitly a function of both the β and λ . In theory we could then optimize this with respect to all $k+1$ parameters. In practice, however, the relative scaling of WSS_x and $WSS_{\ln(x)}$ is itself a function of λ and this complicates optimization appreciably. The alternative we employ is less elegant but feasible and relatively efficient. Specifically, we systematically search over λ until we find a value which yields acceptable R^2 's in both metrics. This search is aided by the fact that (barring extreme clumping) if $R^2_x < R^2_{\ln(x)}$ then we can generally improve the overall performance by decreasing λ (i.e. by placing more emphasis on larger observations). This is illustrated graphically in Figure 2 which presents the ANOVA R^2 's for the annual amount of out of pocket medical expenditures and their logs as a function of λ .³ While the plots of these R^2 's are neither smooth nor even monotonic

³These data are taken from Wave 1 of the AHEAD survey conducted by the Survey Research Center in 1993.

(a result of finite clumpy data), it is clear that there is an overall trade-off between levels and log-levels. For small λ the optimal bracket R^2 's for log-levels are much larger than for raw levels. This is because small λ correspond more closely to logarithms and emphasis is placed on variation at the bottom of the distribution. For large λ , on the other hand, the optimal bracket R^2 's for levels exceed those for log-levels. In this case the Box-Cox transform is closer to levels and the algorithm stress variation at the top of the distribution.

Tables 6a and 6b present, respectively, the optimal dollar breakpoints and the optimal distribution of observations into the implied brackets for medical expenditures for three values of the Box-Cox parameter λ . For $\lambda = .20$, the first and second breakpoints are quite low (\$164 and \$669) and result in a relatively even distribution of the cases into the brackets. This is because the Box-Cox transform in this case is closest to a logarithmic transform and emphasis is placed at the bottom. With $\lambda = .80$, the opposite is true. This makes sense because the transform is closer to the level and in this case those very few cases with extreme expenditures (over \$21,293 per year out-of-pocket) dominate the overall variance.

Table 7 presents the R^2 's obtained using the bracket breakpoints from the HRS Wave 1 brackets and those which would have been obtained using the optimized breakpoints for four of the net-worth components in the HRS. In three of the four cases the R^2 's for levels were increased as a result of the optimization with only modest reductions in the R^2 's for the log-level values. As was the case with medical expenditures, detailed examination of the breakpoints (not shown) reveals that most of the improvement for levels came about by increasing the upper most breakpoint. This has the effect of isolating a very few cases at the top of the distribution into the upper bracket. For level, of course, it is just such cases which contribute the most to the

variance. In the fourth case "Business Assets", the R^2 's for both levels and log-levels were increased by optimization.

V Effects of Unfolding Brackets on Response Quality

In the preceding sections we have seen that the unfolding bracket methodology can reduce item nonresponse considerably and that good break-points can lead to a minimal loss in information. These conditions are necessary if we are to conclude that the unfolding method results in better data overall. But the apparent variance reductions alone are not sufficient to justify the general use of the unfolding bracket method. We must also know if exposure to the methodology significantly decreases respondents' quality standards in reporting. The reason this might happen is that exposure to the brackets (or range cards) sends the respondent the message that great precision in reporting is not necessary--order of magnitude reports are perfectly acceptable.

The ideal method of addressing this issue would be to expose respondents to the methodology at random and compare the accuracy of subsequent reports (via comparisons to validating data) for those exposed versus those not exposed to the method. Unfortunately, we lack both random assignment and validating data and must rely on observational methods to address the question.

As a substitute for accurate validating data we will rely on the empirical validity of data. By empirical validity we mean the strength of association of the measures in question with covariates which theory suggest they should be associated. In other words, we specify theoretically plausible models and then judge data quality on the basis of goodness of fit of the

data to those models.

Lack of randomized exposure of respondents to the unfolding bracket method is a more serious problem in our analysis. Since there is a series of questions for which unfolding bracket followups were available, we can classify responses to items toward the end of the series by whether or not the respondent had encountered the bracketing method in a prior item. Weaker association of the variable to its theoretical covariates for those exposed would be consistent with the hypothesis that exposure reduces the respondents' response quality--with the resulting increased noise in the report attenuating true associations.

Such a finding, however, would also be consistent with the hypothesis that sloppier or less informed respondents are more apt to become exposed to the unfolding brackets. In this case the lower empirical validity of the data is merely a reflection of these poor reporters being disproportionately represented in the exposed group via self selection.

An alternative which gives us some purchase on the exposure versus self selection question is to focus on an item in the middle of the series of questions for which unfolding brackets are available and compare the non-exposed cases with two pseudo-experimental groups-- 1) those exposed prior to; and 2) those exposed only after the item in question. If the empirical validity declines for group 1) only, then we could conclude that exposure to the method decreases data quality. If, on the other hand, the empirical validity of the item for both groups 1 and 2 is lower than the completely unexposed reference group, then we would conclude that the decline is due to self selection.

Unfortunately, the number of items in Wave 1 of the HRS for which the unfolding brackets are available is rather limited. It is therefore difficult to find a single item in this

sequence which is both common and sufficiently close to the middle of the sequence to have a rich set of items before and after it to provide a balanced design. The closest we can come is the item concerning the value of assets in "Savings/Checking Accounts". Most of the HRS respondents had such accounts, and it is more or less in the middle of the asset sequence being preceded by five items and followed by four items for which brackets are available. Table 8 presents the distribution of HRS Financial Respondents by pre- and post-"Accounts" exposure to the unfolding bracket method. Because they introduce additional considerations, cases in which the brackets were used for the "accounts" response, itself, are distinguished in the table as "Concurrent" exposures cases.

[Table 8]

From the table it is clear that the reference group of 4,098 cases which were never exposed to the unfolding brackets is the largest single group of respondents. Furthermore, there were a total of 3,122 respondents who were exposed to the method prior to the "accounts" question and 1,536 (1,275+261) after the question.⁴ While it would be tempting to use all of these cases in our comparisons, most of them are contaminated. The "concurrent" cases are contaminated because the mere fact that they use the unfolding brackets to obtain the dependent variable introduces measurement error which would be correlated with the treatment variable. This, in turn, would introduce a downward bias in the empirical validity of the treatment groups of unknown magnitude. Similarly, it would be tempting to include the 173 cases which were

⁴The "concurrent" cases fall into the post accounts group, because their exposure is after the reading of the question.

exposed before and after the "accounts" question in both the pre- and post- treatment groups. This, however, would bias test of structural differences between pre- and post- "accounts" exposure models toward acceptance of the null hypothesis of no difference.

As a result of these considerations, the actual samples to be used in our evaluation of the effects exposure and self-selection are those in the shaded cells of Table 8. The very small number of clean post-"accounts" exposure cases means that the power of our tests of self-selection versus true exposure effects is considerable reduced.

IV.1 Empirical Model and Relative Empirical Validity Results

Of course, before we can assess the empirical validity of the accounts data and test for the effects of exposure and self-selection on it, we need to specify an empirical model. Ours is based on the proposition that the (natural logarithm of) the desired holdings of liquid assets is a function of household financial position (logs of income and net worth other than liquid assets), household demographics (age, sex, race and gender of the financial respondent), health of household members, and the education and cognitive ability of the financial actors. Thus we can express this desired liquidity as:

$$Y_i^* = \alpha + \beta'X_i + \epsilon_i$$

where X_i is a vector of the factors listed above, β is a vector of parameters relating these characteristics to the log of desired liquidity (Y_i^*) and ϵ_i is a random disturbance term. While the desired liquidity might be arbitrarily small, actual liquidity is bounded at \$0 from below. The observed liquidity therefore can be represented as:

where Y_i is the observed log of holdings in checking and savings accounts. Because of the

$$\exp(Y_i) = \exp(Y_i^*) \quad \text{iff } \exp(Y_i^*) > 1$$

$$\exp(Y_i) = 1 \quad \text{iff } \exp(Y_i^*) \leq 1$$

truncation in the dependent variable explicit in equation 4.2, we employ a Maximum Likelihood Tobit model to obtain estimates the parameters (β) and the measures of goodness of fit which in our case will be the pseudo R^2 or likelihood ratio index ρ^2 .

Table 9 presents this measure of empirical validity for the reference group and for the two control groups defined above. Clearly, the empirical validity of the accounts data from respondents who were never exposed to the unfolding brackets is substantially higher than that from respondents who were ever exposed. The goodness of fit for those exposed prior to the accounts question is only two thirds ($.67 = 8.65/12.83$) as large as the never exposed group while that for those exposed after the accounts question is only three-quarters as large as the unexposed reference group. These results suggest that it is self-selection rather than exposure to brackets per se which is driving the results.

[Table 9]

Of course, the ρ^2 in Table 8 are themselves random variables and it is possible that their differences are more apparent than real. These results were obtained by estimating the model separately for members of each of the three groups. One way of formally testing the differences in the fits of the model is to estimate the model for the combined sample incorporating a sequence of equality constraints on the various parameters. The resulting declines in combined goodness of fit can be tested via likelihood-ratio tests and the significance of various aspects of similarity and differences can be assessed.

The first hypothesis to be tested is whether exposure to the unfolding bracket method

biases the structural parameter estimates (i.e. the β 's). To test this we estimate the model on the pooled sample constraining $\beta_{00} = \beta_{10} = \beta_{01}$. This yields a log-likelihood value of -13,917.2 which when compared to the combined unconstrained log-likelihood of -13,904.8 implies a likelihood ratio χ -square of 24.8 with 22 degrees of freedom. This is well below the critical χ -square of 33.9 and thus we can not reject the hypothesis that any apparent differences in the β 's due to exposure to bracketing are due solely to chance.

The second meaningful hypothesis is that the mean and residual variance for those exposed early are really the same as for those exposed late. This is essentially the "self selection" hypothesis. It is implemented by imposing the restraints $\alpha_{01} = \alpha_{10}$ and $\sigma_{\epsilon 01} = \sigma_{\epsilon 10}$. With these imposed the log-likelihood drops to -13,922.9--implying a likelihood-ratio χ -square of 11.4 with 2 degrees of freedom. Since the critical χ -square for 2 degrees of freedom is 5.99, we can reject the null hypothesis of no significant differences in the mean and residual variance of the early and late exposure groups. We will discuss the implications of this in conjunction with the point estimates presented in Table 10 below.

The final hypothesis is that there is no effect of either early or late exposure on the mean and residual variance and hence no effects on the empirical validity. This hypothesis is implemented by constraining $\alpha_{00} = \alpha_{01} = \alpha_{10}$ and $\sigma_{\epsilon 00} = \sigma_{\epsilon 01} = \sigma_{\epsilon 10}$. Adding these two further restrictions results in the log-likelihood declining from -13,922.9 to -13,933.5. The resulting likelihood ratio test statistics of 21.2 with two degrees of freedom is highly significant--thus there is little question that exposure to unfolding brackets is associated with decreased data quality.

Given the above, the most parsimonious model which does not do significantly decrease

Draft--6/7/95

the goodness of fit is that which constrains the effects of predictors of assets in the form of accounts to be equal for those exposed and unexposed to brackets but allows the mean and residual variance to vary across groups. The parameter estimates for this specification are presented in Table 10. The most powerful predictors of the value of assets in checking and savings accounts are net-worth (other than accounts) and income followed closely by education--all of which have a positive effect. Since net worth and income are specified in their logarithmic form, the .856 coefficient on income is interpretable as the percent increase in accounts value associated with a 1 percent increase in income. It is thus an elasticity estimate and it is quite large--more than three times as large as the corresponding net-worth elasticity. All the other predictors in the model are entered in level or dummy form and therefore do not have the elasticity interpretation. The coefficient of .329 for education means that each year of education is associated with a 33% increase in assets in the form of checking and savings accounts. Older married and healthier respondents also have higher accounts balances whereas African American respondents have substantially lower balances. The "Immediate Recall" and "Similarity" variables refer to the score on measures of cognitive ability and the positive and significant coefficients suggest that the more able have more assets in the form of checking and savings accounts than do the otherwise similar less able.

[Table 10]

VI. Conclusions

In this paper we have shown that the unfolding bracket methodology can substantially

reduce item missing data in financial surveys. Furthermore, a large proportion of the variance in the underlying measure can be recovered with as few as three additional questions. We have also shown that use of the Box-Cox transform and the downhill simplex minimization algorithm can yield optimal breakpoints for the brackets which recover much of the variance in both levels and log-levels of the financial variables. Finally, while bracketing is associated with lower empirical validity of the data, it appears that this is a result of self-selection rather than a consequence of bracketing itself.

References

- Curtin, R. T., Juster, F. T., and Morgan, J. N. (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, The University of Chicago Press.
- Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1989), "Numerical Recipes in Pascal: The Art of Scientific Computing," Cambridge University Press, pp. 326-330.
- Heitjan, D. F., and Rubin, D. B. (1991), "Ignorability and Coarse Data," *The Annals of Statistics*, Vol. 19, No. 4, pp. 2244-2253.
- Juster, F. T., Heeringa, S. G., and Woodburn, R. L. (1992), "The 1989 Survey of Consumer Finances: A Survey Design for Wealth Estimation," in *Statistics of Income and Related Administrative Record Research: 1990*, eds. B. Kilss and B. Jamerson, Washington, DC: Statistics of Income Division, Internal Revenue Service, Department of the Treasury, pp. 107-131.
- Juster, F. T., and Smith, J.P. (1994), "Improving the Quality of Economic Data: Lessons from the HRS", HRS Working Papers Series #94-027, presented at NBER Summer Institute on Health and Aging, Cambridge, MA, July, 1994.

Table 1
Item Missing Data Rates on Major Household Surveys
Percent Item Non-response on Asset and Liability Items

Item	1981 NLS	1979 RHS	1984 SIPP	1983 SCF	1989 SCF	1992 HRS
Home Equity (Primary Residence)	11.3	13.5	NA	7.8	4.8	6.8
Other Real Estate	13.6	12.8	33.5	9.2	8.1	5.0
Checking / Savings Accounts	25.9	14.2	13.3/ 16.8	9.6/ 14.1	7.9/ 2.8	5.2
Stocks, Bonds and Trusts	29.6	28.1	25.9	24.7	13.8	5.8
Savings Bonds	32.6	23.6	24.9	17.4	5.7	6.7
Consumer Debt	13.5	1.1	NA	5.6	4.0	5.8

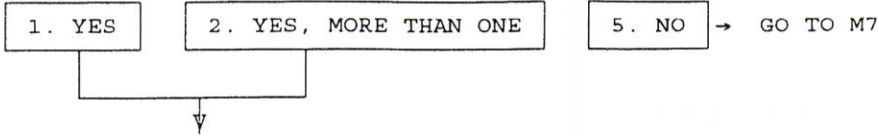
Table 2
 Examples of Response Bracket Ranges for HRS Asset Items⁵

Bracket	Business Value Response	IRA, KEOGH Response
1	\$1 - \$9,999	\$1-4,999
2	\$10,000 - \$49,999	\$5,000 - \$24,999
3	\$50,000 - \$499,999	\$25,000 - \$49,999
4	\$500,000 +	\$50,000 - \$99,999
5	Inapplicable	\$100,000 +

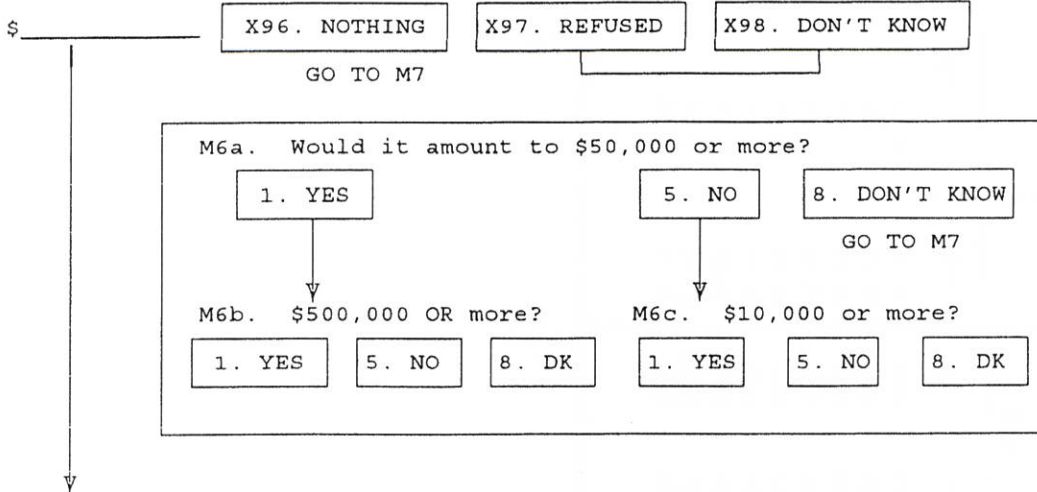
⁵The number of brackets and the associated dollar amounts vary to reflect differences in the range of the underlying asset distribution.

Figure 1

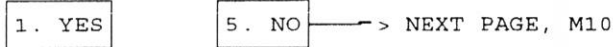
M5. Do you [or your (husband/wife/partner)] own part or all of a business?



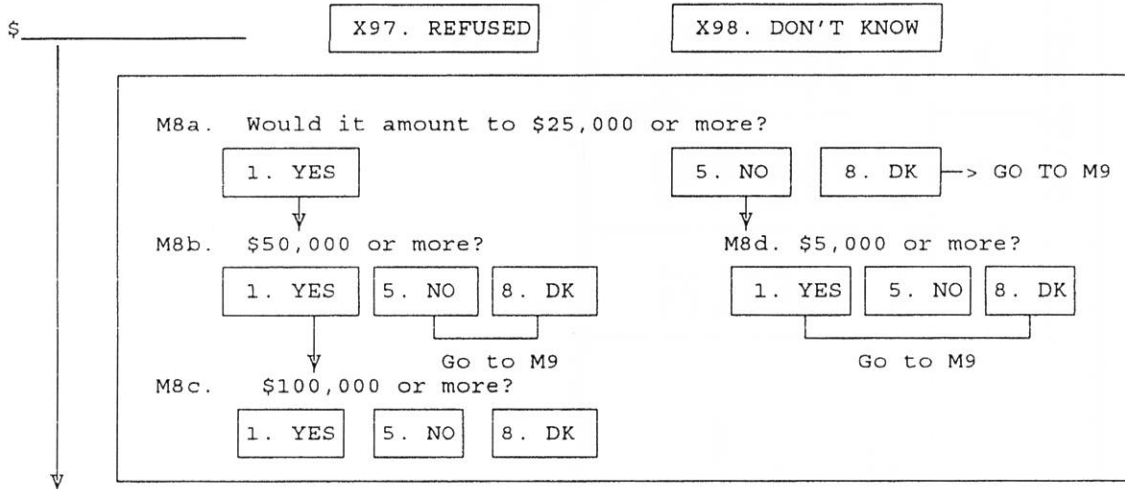
M6. If you sold (all of) the business(es) and paid off any debts on (it/them), how much would you get?



M7. Do you [or your (husband/wife/partner)] have any Individual Retirement Accounts, that is, IRA or Keogh accounts?



M8. How much in total is in all those accounts?



M9. How much did you put into (this/these) account(s) last year, 1991?



Table 3.
HRS Wave 1 Net Worth Components
Distribution of Responses by Response Type (n = 7608 respondent households)

Asset or Liability Item	Does Item Apply?			If Item Applies To Household						
	Total	Yes	No	DK	Total			Range Card Value	Missing Value	
					n	%	Actual Value			Bracketed Value
A: Real Estate (not home)	100%	23.1%	76.3%	0.6%	1759	100%	74.7%	16.2%	5.9%	3.4%
A: Vehicles, Pers Property	100%	-	-	-	7608	100%	87.4%	8.2%	2.4%	1.9%
A: Business	100%	16.1%	83.4%	0.5%	1226	100%	68.2%	21.3%	4.7%	5.8%
A: IRA, KEOGH	100%	37.1%	62.2%	0.7%	2825	100%	73.7%	16.3%	5.2%	4.9%
A: Stock, Mutual Funds	100%	26.5%	72.6%	0.9%	2015	100%	67.4%	21.0%	5.6%	6.0%
A: Checking, Savings	100%	77.5%	21.5%	1.0%	5895	100%	73.3%	16.4%	5.1%	5.2%
A: CDs, Sav Bonds, T-Bills	100%	24.6%	74.3%	1.1%	1870	100%	70.6%	16.2%	6.4%	6.8%
A: Bonds	100%	5.9%	93.2%	1.0%	445	100%	69.7%	13.3%	6.5%	10.6%
A: Other Assets	100%	15.0%	83.9%	1.1%	1143	100%	72.1%	16.3%	5.3%	6.4%

Table 4
Percentage of Variance Explained by Brackets
HRS Wave 1 Net Worth Components

Net Worth Component	degrees of freedom	R ₂ Asset Level	R ² Log(Asset Level)
Real Estate	3	37.9%	72.9%
Vehicles, Personal Property	3	10.8%	74.8%
Business	3	52.8%	80.8%
IRA	3	55.4%	87.2%
Stocks	4	74.2%	75.3%
Savings Accounts	4	28.7%	86.7%
Certificates of Deposit	4	38.1%	90.5%
Non-Gov't Bonds	4	80.7%	79.2%
Other Assets	3	14.0%	78.2%

Table 5
Longitudinal Bracketing in PSID

Real-Estate				
Type of Report	Exact 1989	Bracket 1989	Don't Own 1989	Total 1984
Exact 1984	439	42	309	790
Bracket 1984	34	10	35	80
Don't Own 1984	373	43	3,999	4,416
1989 Total	846	95	4,343	5,284
Business or Farm				
	Exact 1989	Bracket 1989	Don't Own 1989	Total 1984
Exact 1984	250	39	154	443
Bracket 1984	48	19	37	104
Don't Own 1984	264	57	4,418	4,739
1989 Total	562	115	4,609	5,286

ANOVA R-Sqrs for Optimal Breakpoints Levels and Log-Levels by Lamda Out of Pocket Medical Expenses

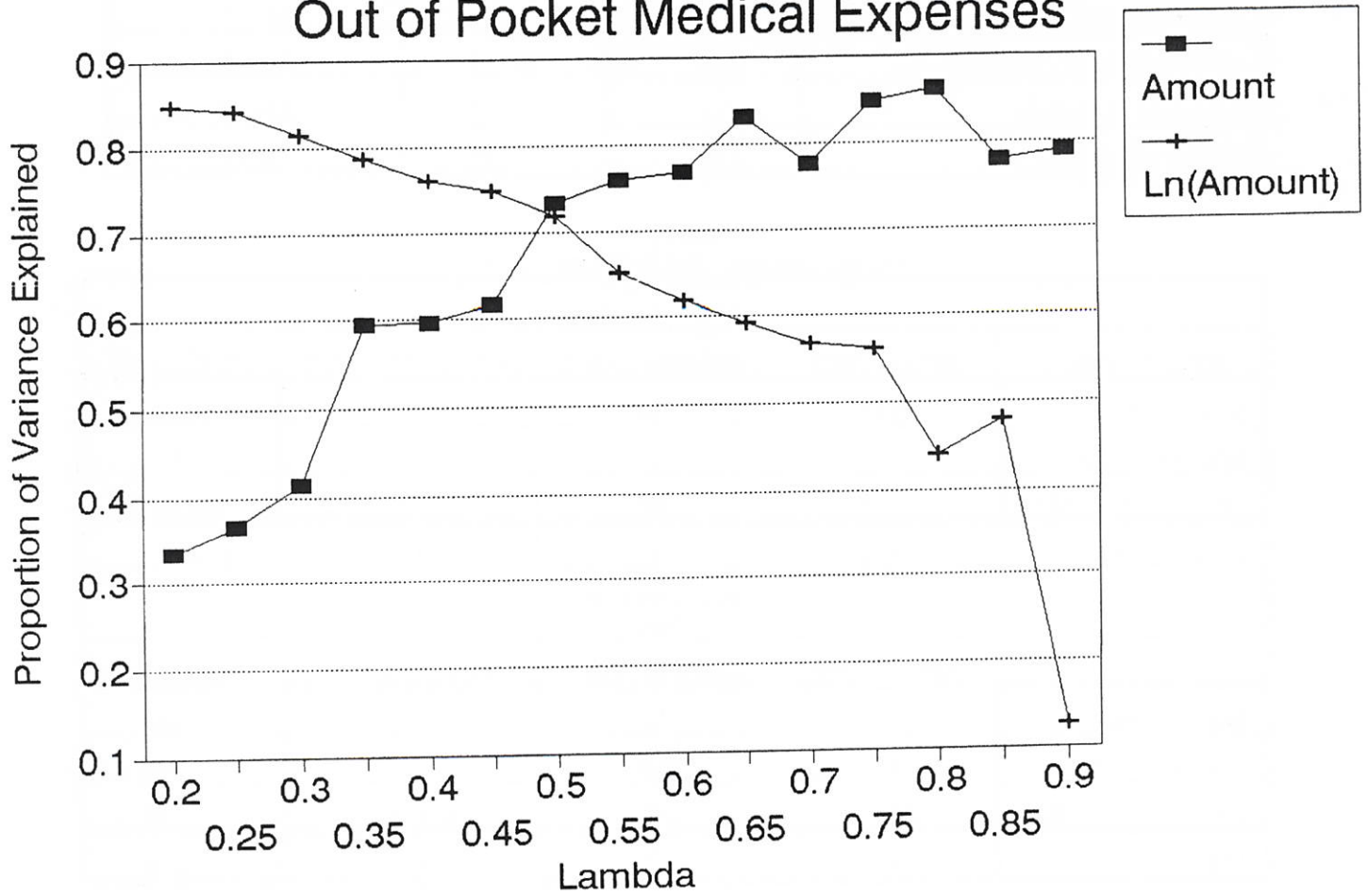


Table 6a
Optimal Dollar Breakpoints
for Medical Expenditures and Three Values of λ

λ	Lowest Breakpoint	Middle Breakpoint	Upper Breakpoint
.20	\$164	\$669	\$2,289
.50	\$420	\$1,814	\$11,472
.80	\$1,251	\$5,761	\$21,293

Table 6b
Distribution of Cases into Optimal Brackets
for Medical Expenditures and Three Values of λ

λ	Lowest Bracket	Second Bracket	Third Bracket	Highest Bracket
.20	975	1,305	859	288
.50	1,812	1,187	415	13
.80	2,726	580	109	12

Table 7
Goodness of Fit with Optimized Breakpoints

Net-Worth Component	Level Wave 1	Level Optimized	Log-Level Wave 1	Log-Level Optimized
IRA	55.4%	87.3%	87.2%	82.7%
CD	38.1%	81.1%	90.5%	76.7%
Other Assets	14.0%	53.0%	76.2%	68.2%
Business	52.8%	60.7%	80.8%	96.6%

Table 8
Distribution of HRS Wave 1 Financial Respondents
by Exposure to Unfolding Brackets

Post-Accounts Exposure

Pre-Accounts Exposure	Yes	Concurrent	No	Total
Yes	173	975	1,974	3,122
No	88	300	4,098	4,486
Total	261	1,275	6,072	7,608

Table 9
Empirical Validity of Accounts Data
by Exposure to Unfolding Brackets

		Post-Accounts Exposure	
Pre_Accounts Exposure		Yes	No
Yes	ρ^2 n	-	8.65% 1,974
No	ρ^2 n	9.70% 88	12.83% 4,098

Table 10
Effects of Bracket Exposure on Accounts Model

1 Constant	-14.91316**	s.e.	0.82710	BHHH T	-18.03062
2 exp(Sigma)	1.31094**	s.e.	0.01430	BHHH T	91.69112
14 early Constant	-15.12717**	s.e.	0.81503	BHHH T	-18.56019
15 early exp(Sigma)	1.42709**	s.e.	0.02049	BHHH T	69.66245
16 late Constant	-14.13905**	s.e.	0.96761	BHHH T	-14.61227
17 late exp(Sigma)	1.16944**	s.e.	0.07985	BHHH T	14.64497
<u>Common Parameters (β's)</u>					
3 Net Worth	0.24835**	s.e.	0.01069	BHHH T	23.23031
4 Income	0.85603**	s.e.	0.03434	BHHH T	24.93055
5 Black	-1.84249**	s.e.	0.12964	BHHH T	-14.21235
6 Male	0.10900	s.e.	0.10968	BHHH T	0.99386
7 Education	3.28816**	s.e.	0.17996	BHHH T	18.27163
8 Married	0.51509**	s.e.	0.11558	BHHH T	4.45655
9 Health	-0.39249**	s.e.	0.04516	BHHH T	-8.69086
10 Age/10	7.56397**	s.e.	1.11239	BHHH T	6.79976
11 Proxy	-0.39828	s.e.	0.37443	BHHH T	-1.06370
12 Recall Immediate	0.09403**	s.e.	0.02137	BHHH T	4.40073
13 Similarity	0.10782**	s.e.	0.01971	BHHH T	5.47060
Log-Likelihood	- 13917.177	ncases = 6160			
19: 2:27.56	1	4/24/1995			