PSID A national study of socioeconomics and health over lifetimes and across generations

Technical Series Paper #95-01

Unfolding Brackets for Reducing Item Nonresponse in Economic Surveys

Steven G. Heeringa, Daniel H. Hill, David A. Howell
Survey Research Center - Institute for Social Research
University of Michigan

1995

This project was supported by funding from the National Science Foundation (SES 9022891).

Unfolding Brackets for Reducing Item Nonresponse in Economic Surveys[1]

Steven G. Heeringa*
Daniel H. Hill**
David A. Howell***

Survey Research Center
Institute for Social Research
University of Michigan
426 Thompson Street
Ann Arbor, Michigan 48106-1248

---

[1]*Head, Sampling Section, Survey Research Center, ** Associate Research Scientist and Project Manager: Health and Retirement Survey, Survey Research Center; ***Programmer Analyst: Health and Retirement Survey, Survey Research Center. This research was sponsored, in part, by NIA Grants U01AG09740 and P01AG10179. 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 brackected 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 >= $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

---

[1]The 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^2s 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(β) is not always well behaved. If the y_i are 'lumpy' then WSS(β) 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 $\lambda \to 0$, $y^* \to \ln(y)$, whereas as $\lambda \to 1$, $y^* \to y-1$. 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

[2]This 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 purpuses 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^2s 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

[3]These 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 λ = .20, the first and second breakpoints are quite low ($164 and $669) and aresult 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 λ = .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 <u>series</u> 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

[4]The "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

$$\exp(Y_i) = \exp(Y_i^*) \qquad \textit{iff} \;\; \exp(Y_i^*) > 1$$

$$\exp(Y_i) = 1 \qquad \textit{iff} \;\; \exp(Y_i^*) \leq 1$$

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 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 Items[5]

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 +

[5]The 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?

1. YES | 2. YES, MORE THAN ONE | 5. NO → GO TO M7

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

$________ | X96. NOTHING | X97. REFUSED | X98. DON'T KNOW

GO TO M7

> M6a. Would it amount to $50,000 or more?
>
> 1. YES | 5. NO | 8. DON'T KNOW
>
> GO TO M7
>
> M6b. $500,000 OR more?
>
> 1. YES | 5. NO | 8. DK
>
> M6c. $10,000 or more?
>
> 1. YES | 5. NO | 8. DK

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

1. YES | 5. NO ——> NEXT PAGE, M10

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

$________ | X97. REFUSED | X98. DON'T KNOW

> M8a. Would it amount to $25,000 or more?
>
> 1. YES | 5. NO | 8. DK —> GO TO M9
>
> M8b. $50,000 or more?
>
> 1. YES | 5. NO | 8. DK
>
> Go to M9
>
> M8c. $100,000 or more?
>
> 1. YES | 5. NO | 8. DK
>
> M8d. $5,000 or more?
>
> 1. YES | 5. NO | 8. DK
>
> Go to M9

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

$________ IN 1991 | X96. NOTHING

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 n	Total %	Actual Value	Bracketed Value	Range Card Value	Missing 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

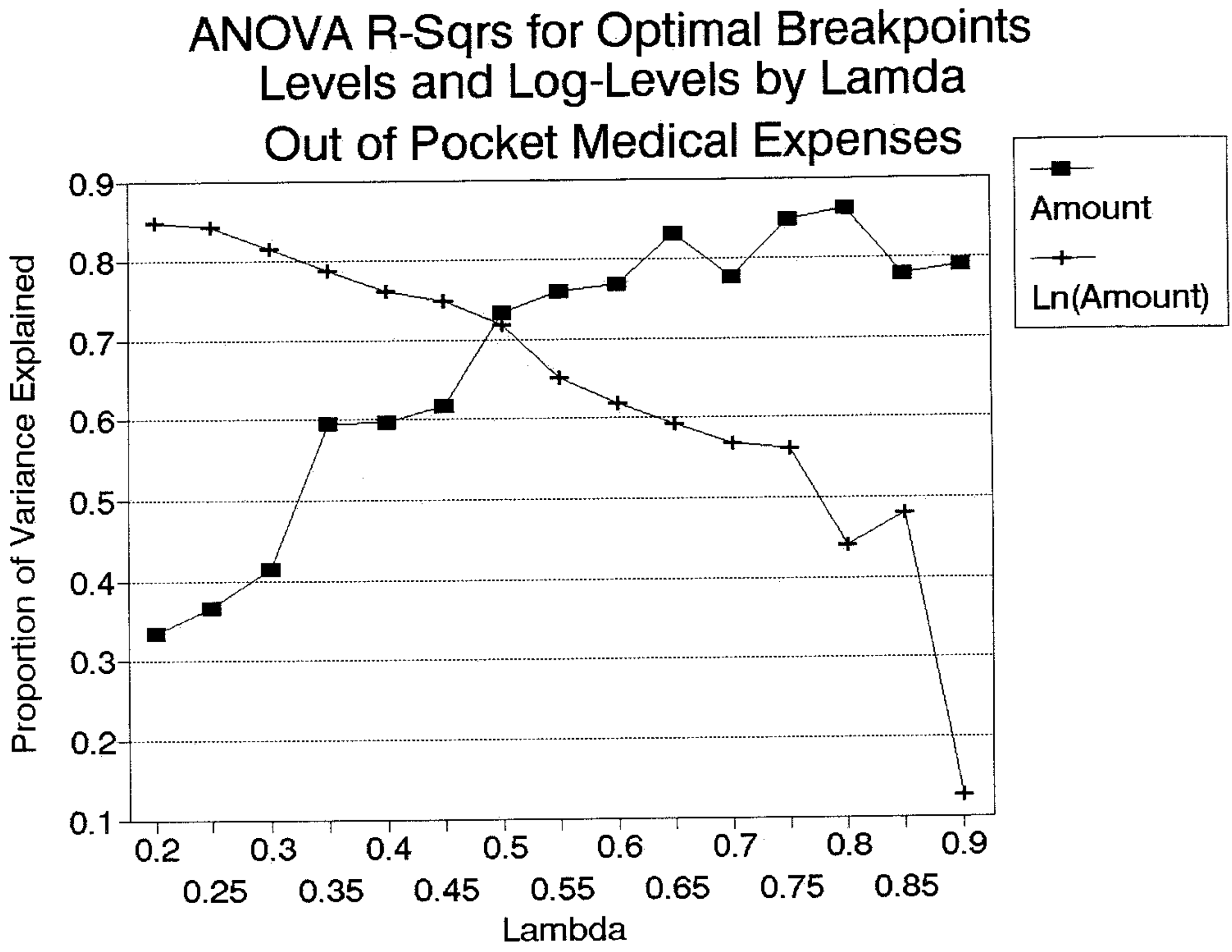
ANOVA R-Sqrs for Optimal Breakpoints
Levels and Log-Levels by Lamda
Out of Pocket Medical Expenses
Proportion of Variance Explained
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9
Lambda
Amount
Ln(Amount)

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

Pre_Accounts Exposure		Post-Accounts Exposure: Yes	Post-Accounts Exposure: No
Yes	ρ^2	-	8.65%
	n		1,974
No	ρ^2	9.70%	12.83%
	n	88	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
Common Parameters (β's)					
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

Economic Behavior Program
Survey Research Center
Institute for Social Research
The University of Michigan

James N. Morgan
Greg J. Duncan
June 11, 1976 DRAFT

AFTER TEN YEARS, WHAT?
AN INITIAL DISCUSSION TO PROVOKE THOUGHT

The Michigan Panel Study of Income Dynamics was first intended as a five-year study, then as a ten-year study. A large study of this kind takes time to wind down, and since extensive commitments made to staff and respondents go beyond the contract year, there needs to be at least a general understanding about the future of the study. If the project is to continue, it would be highly desirable to have a longer horizon than is possible within our present system of annual contracting. Some certainty about the future would allow longer term planning of the questionnaires in order to cover a series of coordinated objectives. However, if the last data collection is to be in 1977, we should now be planning a final clean up, archiving, and some in-house analysis to give future users a base from which to work.

There are many people (and not just at Michigan) who think the study should continue, and it seems useful to make a list of some of the research objectives that would benefit from a longer series of data. What research objectives are met with fifteen years of panel data that could not be met with the ten already committed?

1. One of the most interesting groups of individuals in the sample consists of those who began as children in their parental families in the first year of the study (1968) and have since "split off" to form households of their own. By the eighth year of the study, there are well over 2,000 of them. For these splitoffs, we have a variety of interesting outcome measures such as educational attainment, early occupation and earnings, fertility, mobility, and so on. More importantly, we have measures of the present situation and background history of the parents <u>as reported by the parents themselves</u>. Other data sets have been forced to rely on the child's memory for information about the parents. This sample of splitoffs accomodates a variety of analyses on intergenerational transmission of values and status. We have already found that the father's attitudes about <u>ideal</u> family size have more to do with completed family size expected by the next generation than does the father's <u>actual</u> completed family size. In another analysis we found that a measure of five-year average parental income was considerably <u>less</u> important in explaining the educational attainment of the children than were measures of the educational

attainment of the parents. Furthermore, father's education level had the stronger influence on the education level of the son, while mother's education level was most important for completed education of the daughters.

While the actual educational attainment and expected completed family size can be measured for a sufficient number of splitoffs with ten years of data, an additional five years could provide information on early earnings and occupational experiences, actual family size, marital stability and mobility, and for a larger number of splitoffs. This allows for studies of some of the processes by which an individual's initial (or inherited) status and personal attributes are transformed into a relatively permanent economic status, residence, marriage, and social well-being. Of special interest are the progression through jobs into, perhaps, a reasonably permanent occupation, the nature of discrimination by race and sex, decisions about marriage and divorce, and choices about residential location. None of these processes are well understood yet; we barely know the basic empirical dimensions, let alone the causes or long-run effects.

Ten years is probably barely long enough to begin to understand some of this, but for the splitoffs we have, on average, far less than ten years. Extending the study to 15 years would provide much better longitudinal information on the splitoffs.

2. Generalization from studies in any historical period is dangerous if that period has not been average or normal or has had so little variation in some variables that their effects cannot be seen. One of our main findings to date has been that people's own attitudes and behavior patterns do not seem to have much effect in altering their economic status. But this has been a period of rising unemployment and inflation, not the optimal period to test the Horatio Alger theory. With longer periods of data, more of the environmental state and county variables can be used in the analysis, and they will have varied more and perhaps in both directions so that we can estimate their effects and remove them from the effects of individual variables. Indeed, it is a policy issue of major importance whether altering the macroenvironmental variables (unemployment, inflation) or the individual behavioral variables produces the desired results most effectively.

With each successive wave, the potential for efficient use of the time-series aspect of the panel data increases. In particular, continuation of the panel for a longer time span would greatly enhance future research on the structure of relations between employment, economic growth and poverty. While past researchers were limited primarily to analyzing the relationship between the overall proportion of families with incomes below a poverty threshold and some inidicator of overall economic activity, the lengthening of the panel would allow focusing on the household as the unit of analysis. That is, the relationship of family economic well-being to economic activity could be explored by pooling the time-series and cross-section observations from the Panel Study.

The results from some preliminary investigation in this area indicated a differential sensitivity of household income to changing economic conditions between blacks and whites. (Dissertation by John Holmes) While blacks experienced movements in family money income roughly proportional to the growth in GNP, the incomes of whites, especially those with a female head, were markedly less responsive to cyclic variations in GNP. A similar contrast was observed between blacks and whites in the responsiveness of family income to changing local labor market conditions. The incomes of blacks were inversely related to movements in the level of the county unemployment rate and the effect was highly significant. By contrast, the incomes of whites were not responsive to local employment opportunities. Thus, it appears that sustained economic growth and tight labor markets play an important role in enabling blacks to improve their economic position.

3. Some decisions people make are made only intermittently, but we nevertheless have theories that deal with them. Examples are having another child, moving, or changing jobs. There are methods of testing the hypothesis that the population is made up of two groups--one of which never changes, the other of which has some probability of changing--but only extended information for relatively long periods allows good investigation of the individual outcome patterns and what influences them. If we think women are disadvantaged because they move with their husbands to maximize the husband's income, then extended periods provide more and more cases to study where there have been moves and job changes. It would not be enough just to study movers or changers; one needs the "control groups" and the pre-move expectations.

4. Many effects take time to work out, and since there are many idiosyncratic forces at work plus some measurement error, only after some years will the changes be large enough to show significantly over the "noise" even for

the original panel families. The very high correlations of change (negatively) with initial level, reflect either random error or cyclic departures from and returns to normal, while the results of people's attitudes and behavior, or the payoff to investment in human capital, produces small marginal changes whose cumulative effect takes time to appear over the "waves." Longer periods, even though they allow more other events to alter outcomes, still promise better estimates of the effects of background variables and of behavior patterns.

5. One advantage of looking at longitudinal data for extended families (the families of both parents and children) in different historic periods, is that one can assess the extent to which the extended family is an actual or potential source of help and the extent to which isolation from the extended family may be an actual or potential source of social problems. Longer periods also make it possible to study the assortative mating process by which new unions are formed, since marriage is perhaps the major source of improvement in economic status for the individuals involved.

6. Perhaps the overriding rationale for continuing the panel is to keep options open in a rapidly changing world, for monitoring major changes as they affect individual households. The desire for social indicators can be met in part by improved repeated measurements without the costs of panel studies, but such data need to be complemented and interpreted by the use of panel data on the impact of change on individuals and families. Sometimes this is possible using the repeated measures already being collected. In other cases it may require small additions. And the analysis is further enriched if similar data are intermittently available from independent national cross-sections.

For example, we have asked a question about expecting to be better off over the next few years, in 1968, 1969, 1970, 1971 and 1975, and that question has been asked quarterly in national samples since 1950 or so. Some self-ratings on sense of personal efficacy (fate control) were asked in 1968-72, 1975, and of wives in 1976, and have been used in a number of other national samples and in the Parnes panels. In 1976 we ask both husbands and wives questions about plans to keep working, desire to keep working, chances of getting another job, all questions which have been asked in other national studies. For such questions, three-way comparisons are possible, as well as analysis of changes over time under the impact of national, local-environmental, and individual experiences.

7. There are also small additions that can be made to round out the information. Some items useful in approximating the net worth of families is an important example. We can estimate the value of income-earning assets from income reports, and net equity in home by repeating a question on mortgage balance, and we could ask about ownership of nonearning assets (e.g., empty land) and life insurance, but the biggest and most disparate missing item is equity in retirement programs. With the pension reform act in place and the great current sensitivity to Social Security and retirement plans, it is likely that people approaching retirement or with any substantial equity in retirement programs would have some notion what kind of retirement income they could count on.

With some assumptions about life expectancy at 65, appropriate discount rates, and when the pension accumulation started (based on years of experience and tenure), we can approximate the present value of pension rights, and its correlations with other measures of well-being.

8. Coordination with other studies is increasingly possible, often with relatively minor changes in each study. We have been doing this with the Parnes labor force panel studies, and the Wisconsin and Michigan panels of youth, and will endeavor to do it with the Dohrenwends' study of stressful life events (where the original focus was psychiatric).

Panel Obsolescence

How can a panel study avoid obsolescence? If one starts with a representative sample and follows it indefinitely, the members age, and die off. Clearly, the simple answer is that replacements are required. If one starts with a probability sample, an obvious unbiased replacement is the children born to that sample. This is what the Michigan Panel Study does. In practice it means that we follow members of the original panel families as they leave home and start families of their own. If the panel went on indefinitely, it would mean identifying the children of panel members, with some complexity, as panel members marry nonmembers and have children, i.e., one might want to count only children of female panel members, or weight down children when only one parent was a panel member. But currently, those starting new units are children who were already in the existing families as family members in the original sample.

The part of the original sample that reinterviewed families from the SEO study excluded those over 65 and above twice the poverty line, but the weights in the merged sample weight up the appropriate groups in the SRC sample making the combined weighted sample properly representative in 1968 and now. Hence, there is no obsolescence problem in general. The cumulative loss after the first year or two has been extremely small--two or three percent a year. Such a record would be difficult or impossible to maintain without annual contact.

The End of Contact?

If the study is to end with the tenth wave, there are a series of alternatives as to the amount of effort put into rounding out and cleaning up the files. By rounding out we mean the possibility of attempting to find and interview some sample members who have been lost from the panel over the years, particularly younger splitoffs and people involved in divorce. Even if intervening data are missing, a final-year measurement on a more complete sample would allow a great deal of useful analysis. While there are relatively few losses after the second year, they are a more substantial fraction of some subgroups such as young men leaving home. Such an attempt would require advanced planning and preparation of sample materials so that in the course of the regular interview those remaining in the sample would be asked about the whereabouts of missing relatives.

Regardless of that decision, there should be a final cleaning of the data file and some methodological investigations of the extent of discrepancy in answers that cross-validate (age, occupation if job did not change, etc.). Several years of analysis funded at Michigan would also be rewarding since there is a professional staff intimately familiar with the data. (They would have to spend some time helping other anlaysts with problems in any case, and locating one analysis effort at Michigan would allow a continuation of a coordinated reporting of all the analysis going on with the data.)

Compromises

Even if the study does not continue on an annual basis, ending all contact with the respondents is not the only alternative. It would be possible to tell respondents we might want to recontact them in five or ten years, and to secure information that would facilitate that. The advantage of this is that it leaves options open, including revisiting only some part of the sample, or different proportions of different strata. A disadvantage is that it would require keeping in some secure place a mass of sensitive (name and address) information over a long period of time. There would certainly be many losses; one should not extrapolate the success in relocating Air Force veterans, college graduates, or even high school graduates from a limited number of schools, to the probable success in locating this widely scattered, heavily poor, and variegated sample.

Another possibility would be to reduce either the frequency or the extent of the communication, or both. Every-other-year interviews make for some (modest) saving in cost but would require contacts of some sort between interviews if we were not to lose substantial numbers from the panel.

We have little faith in the possibilities of using mail alone for keeping contact--at least telephone and sometimes personal contacts would be required for many cases. It seems doubtful to us that the saving in cost would be commensurate with the loss of information. Failure to secure at least annual measurement of income earnings, work hours, family composition, the location of splitoffs, reasons for changes in job and location and plans to change, would make much analysis difficult or impossible. Indeed, there has been a movement toward securing dating of events within each year. At a minimum, going to an every-other-year procedure would require advance commitment since only if the respondents were notified, and procedures set up for maintaining contact (and even paying for such responsibility), would we stand a chance of maintaining the sample.

Reducing the <u>content</u> of an annual data collection would save some interviewer time, some data processing, and some other marginal costs, but the savings would be quite small relative to the total budget. It would be difficult in a period of inflation, for instance, to reduce the interview payment from $7.50 to $5.00. Indeed, while the costs of additional waves seem large, they are marginal in comparison with the investment already made, or the huge costs of attempting alternative data collections (particularly starting a new panel).

Finally, there are always demands for special or extensive additional data from part or all of the sample. Medical examinations have been suggested, particularly for those who report some disability or an erratic work experience. Remeasuring the coping behaviors and relevant attitudes secured in the first five years would allow a more serious study of the dynamic interplay of attitudes, behaviors, and economic success/failure.

Application of a test (sentence completion?) to the current heads and wives in the sample, and perhaps also a measure of achievement motivation, would also help tighten the explanation of changes in economic fortunes, but might well require personal interviews with most or all of the sample, at substantial expense.

We do not think the payoff to adding questions on family background, fertility and contraceptive history, detailed job history, is very high. We already have extensive information on family background of both the head and wife, eight years of recent experience, and some summary reports of the wife's pattern of working before and after marriage and before and after the first child. Memory of detailed sequences more than ten years previous is treacherous and requires an extensive and complex interviewing process to pin down dates.

There are numerous special objectives for which subparts of the sample seem appropriate. The costs of "using up" subsamples in this way are very large, however. That subsample is likely to feel they have completed their obligation, or to respond differentially to subsequent approaches, which means that the rest of the sample is no longer a representative sample, except of other nonintersecting subgroups. In addition, these special demands are likely to be seen as an added and unexpected and unrelated burden imposed on respondents, and to violate the doctrine of informed consent. In general, we regard such possibilities as of dubious merit. At the point where the panel is to be closed down, however, the issue of whether various nonoverlapping subsets should be exposed to requests for additional cooperation on at least partially related and relevant topics becomes an open one.

The respondents themselves are one of the main assets of the study. It has been possible to interview the same families and their splitoffs for nine years with very few losses, but this has required a year-round succession of contacts, payments, gratitude and persuasion. If this routine is broken, even for a year, it is unlikely that it will be possible to reassemble the sample in anything like its present representative form.

APPENDIX

A VERY CURSORY OVERVIEW OF WHAT IS IN THE PANEL STUDY OF INCOME DYNAMICS

It is a panel of 5,700 families as of now. It consists of 16,000 individuals who were in an original sample in 1968 plus those who have married, been born to, or otherwise joined them. Looking backward, we have annual interviews with the heads of whatever families each person was in for each year back through the spring of 1968. They represent a replacing panel with the children replacing those who die off. We are currently collecting ninth year interviews with both "heads" and wives. The original sample oversampled low-income families, but wieghts provide unbiased representative estimates. The cumulative response rates have been sufficiently high so that we still have more than half the original sample individuals. Most of the losses were in the first two years (deaths and those respondents too ill to cooperate are included in nonresponse).

The original sample combined a national cross section of about 2,800 families with some 1,800 families previously interviewed as part of the Census' Survey of Economic Opportunity and who had incomes less than twice the poverty level. Hence, comparisons of low-income families with others, or of blacks with others, have increased reliability since the relevant sample is larger than it would be on a random sample.

The data are optimal for studying change in all its details and interactions, not to estimate aggregates but to understand what is behind changes in aggregates.

Year after year the basic content is repeated: details of income, occupation, labor force participation, and family composition. Thus, change can be studied by looking at changes in these repeated measures, as well as by answers to questions about <u>expected</u> changes (in job, location, family) and explanations for past changes in job or residence. Patterns of employment, unemployment and commuting within the family can also be studied.

Many changes are offset by other changes, so the panel is ideal for studying these offsetting changes. For example:

- Changes in heads' wage rates offset by changes in hours worked.
- Changes in heads' earnings offset by changes in earnings of wife and others.
- Changes in family earnings offset by changes in transfer income.
- Changes in family income offset by changes in family needs (size and structure).
- Changes in location (commuting costs, housing costs) offset by changes in earnings.

Extensive family background information has also been collected, including the education of head and wife and all four of their parents, measures of college quality (where relevant) for head and wife, and the counties where the head and wife grew up.

The interplay of attitudes, behavior patterns, and economic fortunes also can be studied: A set of measures of coping behaviors and economically relevant attitudes were measured in each of the first five years. They were partly remeasured in 1975 for head and in 1976 for wives.

A variety of well-offness measures can be derived from the data. Components of nonmoney income are measured as well as some of the costs of earning income such as commuting and union dues. We also record the existence of the most important and variable fringe benefit, coverage by an employer pension plan. In 1975 we added yet another measure of well-offness by asking about the quality of housing and neighborhoods.

The data would be ideal for assessing the impact of alternative definitions of poverty that combine earnings and free time if we were to move to an income maintenance system that took account of work effort as well as earnings.

The data also allow one to study where needs are temporary and where they are permanent since the design of programs to cover temporary emergencies may well be different from those designed to deal with more lasting problems. Data on rents and house values, on miles driven, and on food expenditures (and food stamp usage) allow studies of the impact of rising prices and taxes, actual or expected.

Actual and potential help patterns can be studied among the now separated part of families, by reports on payments given or received (usually minimal), and by studies of changing family composition.

To help explain changes in individuals' employment and income, studies of the impact of environment on individuals are possible by using county data and national aggregate data on employment. Differential impacts in different occupations and industries also speak to issues about structural, aggregate, and individual sources of economic difficulty for families.

We have attempted to simplify use of the data by providing documentation where one can use an alphabetic index to variables. Questionnaires with the variable numbers of the answers noted can be read, reference can be made to the worksheets to see precisely how composite or edited variables were treated, or one can use the concordance to check for the yearly availability of a particular variable. Finally, one can turn to the actual codes for detailed categories and the sample distribution across them. If questions still arise, Jock Lane can usually answer them or will know who can.

We attempt to provide some analysis each year and also a summary of as much of the analysis done elsewhere as we can find out about, and we are eager to have others use the data.* The family data tape is made available as soon as we are reasonably sure it is clean (January, for data collected in March-July), and the individual data tape similarly--but a few months later. As the length of the panel and the amount of data per individual expands, the time required to produce merged files has expanded, and we are pushing the bounds of technology. If it seems important, the latest year's data tape can be made ready unmerged somewhat earlier (though some corrections arise when the merging takes place).

We are eager to know of work going on, willing to read drafts, plans, and so forth. We'll try to answer questions about using the data if you'll send them in writing so we have time to think and look things up.

* The findings and some methodological appendices are in Vols. I-IV of Five Thousand American Families and the documentation in Vols. 1-2 and Wave VI-Wave VIII supplements of A Panel Study of Income Dynamics, Institue for Social Research, Ann Arbor.

SN ✓
MH ✓

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
OFFICE OF THE SECRETARY
WASHINGTON, D.C. 20201

August 20, 1976

Dr. James N. Morgan
Institute for Social Research
University of Michigan
Ann Arbor, Michigan 48106

Dear Jim:

I have been meaning to give you my reaction to your letter of July 2 to Richard Coleman objecting to the way he wrote up his analysis of the PSID interviews. I will get this letter off to you before you leave; then probably continue the dialogue with your staff.

I disagree with your idea that one should not try to put some intuitively meaningful flesh on the bones of survey reporting by writing up vignettes of life in colorful language when they illustrate general points. It seems to me this is very much in the tradition of social psychology and survey reporting - from Sam Stouffer, through Dan Lerner to Lee Rainwater. Rainwater's last book is partly titled "the social meanings of income." Weaving together juicy case study material with survey findings is the way we know what survey findings mean in human terms, in my opinion. It is true, I guess, that economists generally think that meaning comes from policy relevance or prediction and control of one variable by another. I suppose one could argue about what is substantively and scientifically useful. As you say in your letter, it would be best if we can avoid that. What I want to emphasize is that Coleman is not doing anything at all unusual for the sociology-social psychology tradition. So I quite disagree with your point that "evaluative terms applied to the data, like purchase of a dream house,"..." are unacceptable for public reporting."

I gather that one reason this kind of write-up bothers you (correct me if I am wrong) is that it does not show proper respect for respondents, or it may hurt your ability to interview respondents in the future, if they read it and are offended by it. I don't find it disrespectful; certainly no more so than "Oscar" and "Hortense" and their "disastrous" marriage in your Interviewer Guide. The question of whether it will cause difficulties for interviewers in the future is then a purely practical one in my mind, and I think it very unlikely it will hurt the response rate considering the number of people who might read it and the numbers of them who might be offended.

For there is no question of anyone being identified. Let's be sure that is clear. There may be some questions about what respondents were told about who would see the data and how it would be used. And there may be some questions about ever having gathered the "interviewer's observations" in addition to the responses to questions. But there is no question about anybody being individually identifiable from Coleman's write-up. I trust you agree on that. If there were, I'm sure we'd all want to change it.

On the question of who gets to see what kind of data, you say, "We assure our respondents that the people who analyze the information which they help us to assemble have access only to coded information from data tapes..." How can this be, since somebody obviously has to read their answers to code them? (Indeed I've heard respondents object more to the loss of their meaning which is involved in coding than I've heard them object to be being read word for word). Our letters to respondents say no person can ever be indentified from the publications based on the study. Using an anonymous case as an example of a general point does not seem to me to violate our assertion that we only analyze groups of people. Do you actually tell people that analysts only have information to coded data, and if so, where?

On the question of using the interviewer's observations ("thumbnail sketch"), I would think we may have a problem, but perhaps only with respect to data collected since the applicability of the privacy act, which doesn't include the years Coleman looked at, does it? You say that if interviewers comments go beyond what is necessary to clear up misunderstandings about the questions actually asked, they go beyond your intentions. But the interviewer guide asks them to add anything they think will help us understand the "respondent's situation" and this seems reasonable, and slightly broader than the questions actually asked. We are using the questions actually asked to study their economic and social "welfare" and anything that pertains to that seems OK.

It may be that with the coming of the privacy act, we cannot record anything in this government-sponsored survey which is not explicitly down on the questionnaire as something to be recorded (as we now do with race). So maybe thumbnail sketches are out with the privacy act, but such information gathered before the privacy act may be usuable.

The thumbnail sketch issue seems important for getting Coleman's analysis into the light of day because he says he relied so heavily on it. To suggest that he be able to publish what he concluded without letting him describe the procedures he went through and what information he relied on is very troublesome.

I know you have a legitimate and proper concern for the integrity of your dealings with respondents and we both have a concern with adhering to the privacy act. But I do think that your desire to eliminate colorful vignettes from individual lives goes beyond what is necessary, and is quite contrary the usual way surveys have been written up in sociology and social psychology.

Sincerely,

Jock

Jonathan P. Lane

Institute for Social Research
Survey Research Center
Economic Behavior Program

September 28, 1976

MEMORANDUM

TO: Members of the "After Ten Years, What?" Committee
FROM: Greg Duncan

I have calculated 11th wave (1978) and 12th wave (1979) budgets for the following alternative project designs:

A. Interview entire sample in both 11th and 12th waves.

B. Interview three-fourths of the sample in both 11th and 12th waves, drop the other one-fourth.

C. Interview one-half of sample in 11th wave and the <u>same</u> half sample in the 12th wave, drop the other half of the sample.

D. Interview one-half of sample in 11th wave and then interview the <u>other</u> half sample in the 12th wave.

E. Make no contact with sample in 11th wave, interview entire sample in 12th wave.

The mail service between Ann Arbor and Agua Amarga, Spain is not swift and so I have not had a chance to get Jim Morgan's reaction to these numbers. Most, however, come directly from the heads of the various service sections in the Center (e.g., coding, field, sampling).

As a summary, Table 1 shows the total costs of these various options, with the 10th wave budget included for reference. The cheapest option is to omit the 11th wave altogether, largely because some research salaries (and overhead costs) are saved that way. The two-year costs of following one-half of the sample are less than $100,000 more expensive than the skipping a year option. The two-year cost of following the entire sample is about $350,000 more than the cheapest option.

not in the table

The numbers in Table 1 rest upon various assumptions regarding project costs. Considerably more detail is provided in Tables 2-4. I have divided costs into those which vary with the number of interviews (Tables 2 and 3) and those independent of the number of interviews (Table 3). The only exception to this rule is that research salaries (shown as a fixed cost in Table 4) drop by 25 percent when <u>no</u> interviews are taken in the 11th year.

There is a story behind each of the numbers on Tables 2-4. I will try to anticipate questions you might have by commenting on the most important assumptions and facts:

1. The research salaries assume continuing with our present number of data managers (coding supervisor, editing supervisor, computer programmers,

secretaries and respondent masseuse) and cutting back the number of data analyzers. The latter (denoted "senior staff in Table 4) are composed of

.5 of James Morgan
1.0 of a study director
2.0 of graduate student research assistants.

I have not yet received a reaction to these assumptions from Jim Morgan.

2. The "computer specialist salaries" are set by Institute-wide formula based on extrapolation of previous years into the future. The "analysis data processing" item on Table 3 includes not only expenditures for analysis but also the costs of analysis runs and constructing special tapes for outside users whom we cannot bill directly.

3. The sample size is assumed to continue to grow at its current rate, reaching 6,350 in the 11th wave and 6,550 in the 12th. It is also assumed that the length of the questionnaire increases by three minutes (10 percent) when we interview every other year and obtain certain retrospective information for the omitted year.

4. Coding salaries are roughly a linear fraction of the number of interviews. Field costs, on the other hand, are not. The per interview field costs of interviewing three-fourths of the sample are 13 percent higher than for the full sample, and average costs of interviewing one-half of the sample are about 25 percent higher than for the full sample.

explain why

Alternating interviews with two half samples further increases costs by about 10 percent (above the costs of interviewing the same half sample) because the respondents will be more difficult to find and because it is assumed that the retrospective questions will add three minutes to total interviewing time. The costs of the extra interviewing minutes are much smaller than the cost of locating the respondents (and the splitoffs).

5. According to our samplers, the only way to "thin" the sample by one-fourth or one-half (other than simply following only the original SRC subsample or only the SEO subsample) is to begin with the original sample design, and divide each subsample with each sampling point. All of the individuals within a given 1968 family, regardless of current geographic location or current family status, will be allocated into or out of the thinned sample. A $3,000 charge is added to the sampling salaries in the 11th wave for options which call for a thinned sample.

6. The current Institute overhead rate is 69.0 percent for Ann Arbor salaries and 43.8 percent for off-campus salaries. Our best estimate of these rates for the 1978-1979 period is 72.0% and 45.0%, respectively.

Table 1

TOTAL COST OF 11th AND 12th WAVES UNDER VARIOUS ASSUMPTIONS ABOUT SAMPLE SIZE AND DESIGN

	(A) Entire Sample Both Years	(B) Three-fourths Sample Both Years	(C) Same Half Sample in Both Years	(D) One-half Sample in 11th Wave, Other Half Sample in 12th Wave	(E) No Interviews in the 11th Wave Full Interviews in the 12th Wave
11th Wave	$ 712,000	$ 656,600	$ 583,500	$ 584,300	$ 298,000
12th Wave	752,200	679,300	592,400	616,500	781,900
Total Cost	$1,464,200	$1,335,900	$1,175,900	$1,200,800	$1,079,900

Table 2

ELEVENTH WAVE VARIABLE COSTS UNDER VARIOUS ASSUMPTIONS ABOUT SAMPLE SIZE AND DESIGN

	11th Wave (1978) Variable Costs				
	(A) Entire Sample Both Years	(B) Three-fourths Sample Both Years	(C) Same Half Sample in Both Years	(D) One-half Sample in 11th Wave, Other Half Sample in 12th Wave	(E) No Interviews in the 11th Wave Full Interviews in the 12th Wave
Salaries (including fringe benefits)					
On Campus:					
Coding Salaries	$ 36,300	$ 27,200	$ 18,200	$ 18,200	$ 0
Central Field Office	11,000	9,400	7,900	7,900	0
Computer Specialist	7,500	6,500	4,700	4,900	0
Sampling	1,500	4,500	4,500	4,500	1,500 why?
Off Campus:					
Field Supervisors	12,300	10,400	8,500	8,500	0
Interviewers	72,900	59,900	43,900	43,900	0
Nonsalary Costs					
Payments to Respondents	78,400	66,500	54,600	54,600	30,800
Data Processing	17,900	15,700	11,300	11,700	0
Supplies, Printing, Postage & Communication	15,700	13,700	11,800	11,800	0
Travel (supervisors and interviewers)	27,300	24,500	16,700	16,700	0
Indirect Costs					
72% (on campus salaries)	40,500	34,300	25,400	25,600	1,000
45% (off campus salaries)	38,300	31,600	23,600	23,600	0
TOTAL BUDGET	$359,600	$304,200	$231,100	$231,900	$ 33,300

Table 3

TWELFTH WAVE VARIABLE COSTS UNDER VARIOUS ASSUMPTIONS ABOUT SAMPLE SIZE AND DESIGN

	12th Wave (1979) Variable Costs				
	(A) Entire Sample Both Years	(B) Three-fourths Sample Both Years	(C) Same Half Sample in Both Years	(D) One-half Sample in 11th Wave, Other Half Sample in 12th Wave	(E) No Interviews in the 11th Wave Full Interviews in the 12th Wave
Salaries (including fringe benefits)					
On Campus:					
Coding Salaries	$ 39,300	$ 29,500	$ 19,700	$ 21,600	$ 43,200
Central Field Office	11,400	10,000	8,400	9,200	12,800
Computer Specialist	8,200	7,000	5,000	5,300	8,600
Sampling	1,500	1,500	1,500	1,500	1,500
Off Campus:					
Field Supervisors	13,100	11,000	9,000	9,900	14,600
Interviewers	79,700	65,500	47,900	57,400	89,300
Nonsalary Costs					
Payments to Respondents	80,900	60,700	40,500	40,500	80,900
Data Processing	19,800	17,300	12,400	13,000	20,700
Supplies, Printing, Postage & Communication	16,500	14,400	12,400	12,400	16,500
Travel (supervisors and interviewers)	29,400	26,300	18,000	21,200	32,300
Indirect Costs					
72% (on campus salaries)	43,500	34,600	24,900	27,100	47,600
45% (off campus salaries)	41,800	34,400	25,600	30,300	46,800
TOTAL BUDGET	$385,100	$312,200	$225,300	$249,400	$414,800

Table 4

FIXED COSTS OF 11th AND 12th WAVES

	11th Wave All Options Except No Interview	11th Wave No Interview	12th Wave
Salaries (including fringe benefits)			
Senior Staff (includes graduate students)	76,800	57,600	80,500
Support Staff	105,100	78,800	108,900
Computer Specialists	8,600	6,400	9,000
Nonsalary Costs			
Data Processing	12,200	9,100	12,800
Research Travel	1,000	1,000 why	1,000
Supplies, Printing, Postage and Communication	10,500	7,900	11,000
Consultants	1,000	1,000	1,100
Indirect Costs			
72% (on campus salaries)	137,200	102,800	142,800
TOTAL	$352,400	$264,600	$367,100