WEALTH DYNAMICS: REDUCING NOISE IN PANEL DATA

DANIEL H. HILL*

The Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI 48104, USA

SUMMARY

Although the asset data from the Health and Retirement Study (HRS) is of very high quality, there is sufficient noise to frustrate attempts to study saving behaviour by examining wave-to-wave change in wealth. In this research, we attempt to reduce noise by means of reactive-dependent interviewing in which respondents with large inexplicable changes in assets between 1998 and 2000 are called back by HRS interviewers, presented with their prior reports and asked to reconcile the data. We achieved reconciliation for 1255 households (2479 net-worth components) and, as a result, the variance in measured change for the entire sample of 11,583 households with the same financial respondents in both waves was cut in half. The empirical validity of the data also appears to have been improved. The correlation of gross change in net worth and income, for instance, increased from an insignificant negative to a highly significant positive value. Although reconciliation of large asset changes marginally improves the goodness of fit of multivariate models, there remains sufficient noise in the asset-change data to require analysts to employ additional methods to reduce the influence of outliers. Copyright © 2006 John Wiley & Sons, Ltd.

1. INTRODUCTION

One of the most important goals of the Health and Retirement Survey is to track household economic status as individuals age from pre-retirement to death. Saving behaviour before and during retirement will have a critical impact on the economic well-being of the old over the next several decades. Panel data provide unique opportunities for the analysis of saving as well as other behavioural dynamics by measuring change at the individual level (see e.g. Lansing and Morgan, 1971; Hsiao, 1986). While saving can, in theory, be estimated from wave-to-wave change in net-worth components and aggregates, spurious change in these measures has frustrated many attempts at estimating saving from the panel data. To see why this might be so, note that the observed net worth (W) at any wave (t) of a panel for each individual (j) can be expressed as the sum of the true value of net worth (W^*) and measurement error (ε). Gross saving (S) between one wave (t) and the next ($t + 1$) can then be expressed as:

$$S_{t+1,i} = [W^*_{t+1,i} - W^*_{t,i}] + [\varepsilon_{t+1,i} - \varepsilon_{t,i}] \tag{1}$$

The first bracketed terms represent the true change in net worth (the signal) and the second bracketed terms represent the change in measurement error (the noise). Our inferences on the relationships of saving to other variables are based on the variance in measured gross saving.

* Correspondence to: Daniel H. Hill, The Survey Research Center, Institute for Social Research, University of Michigan, 426 Thompson Street, Room 3136, Ann Arbor, MI 48104, USA. E-mail: dhhill@isr.umich.edu

Assuming for simplicity that the measurement errors are uncorrelated with the true levels (a hypothesis we will find reason to reject below), the variance in measured gross saving is:

$$
\begin{aligned}
\sigma^2_{S_{t+1}} = {}& [\sigma^2_{W^*_{t+1}} + \sigma^2_{W^*_t} - 2\rho_{W^*_{t+1}, w^*_t} \sigma_{W^2_{t+1}} \sigma_{W^2_t}] \\
& + [\sigma^2_{\varepsilon_{t+1}} + \sigma^2_{\varepsilon_t} - 2\rho_{\varepsilon_{t+1}, \varepsilon_t} \sigma_{\varepsilon_{t+1}} \sigma_{\varepsilon_t}\}]
\end{aligned}
\tag{2}
$$

where, again, the first bracketed terms represent the variance in signal and the second the variance in noise. The fact that in survey data the σ^2_ε tend to dominate the $\sigma^2_{W^*}$ means that noise dominates signal—something that is only intensified by the fact that there is good reason to expect the correlation in the error components to be much weaker and, perhaps, even negative[1] at the same time as there is good reason to expect that the correlation in the signal components will be strongly positive.

While noise in change in panel designs tends to dominate signal, panel designs also provide a means of reducing the noise in change by allowing the transfer of information from one interview wave to the next (see e.g. Mathiowetz and McGonagle, 2000). Various forms of dependent interviewing, in which reports from prior interviews are presented to the respondents in the current interview, have been employed in panels and a consensus is beginning to form that the benefits of reduction in 'false positive' indications of change far outweigh the costs of increased 'false negative' indications (see e.g. Hill, 1995; Dibbs et al., 1995). An important question is whether this same conclusion holds for wave-to-wave changes in net-worth components and aggregates.

In this research we introduce a form of dependent interviewing to the HRS and attempt to evaluate its impact on the quality of net-worth measures. The form of dependent interviewing we use is what has been called 'reactive-dependent interviewing' (see e.g. Mathiowetz and McGonagle, 2000; Brown et al., 1998), in which the respondent is reminded of the prior report only when it is seriously at odds with the current report. The data collection phase of the study consisted of calling back in 2001 a subsample of HRS respondents whose reported net-worth components differed substantially and inexplicably between 1998 and 2000 and asking them to reconcile the differences. Although we chose this design primarily for reasons of practicality, it also has the advantage of preserving the independent measures we have been collecting since 1992 while adding the dependent measures. This is a property that we have incorporated in the dependent asset reconciliation in our 2002 computer assisted interviewing (CAI) questionnaire. That is, we first ask about asset values without benefit of past records and then, at the end of the interview, conduct a reconciliation based on the just collected and the preloaded prior-wave asset values.

This paper is organized in five sections. In Section 2, I will briefly describe the criteria used in selecting households and net-worth components to be reconciled. I then present, by asset and response type, the numbers identified, reconciled and corrected as a result of the callback. In Section 3, the effects of reconciliation on the first four distributional moments and various quantile breakpoints are presented for representative net-worth components and for the aggregate levels and wave-to-wave change. The effects of length of recall and changes in economic conditions from the time of the original data collection to the reconciliation callback are also discussed briefly in Section 3. Section 4 explores the effects of reconciliation on the empirical validity of the asset-change data. Finally, I summarize my conclusions and present recommendations for future reconciliation efforts in Section 5.

---

[1] A negative error correlation would result if there is strong regression to the mean—large positive errors in one wave being followed by smaller or negative errors in the next.

2. HISTORY OF SELECTION RULE AND THE DISTRIBUTION OF THE RECONCILIATION SAMPLE

As noted above, the purpose of this research is to reconcile inexplicable differences in asset reports between interviewing waves. The basic design of future asset-change reconciliation efforts was set during a meeting of the HRS Economic Status Working Group (ESWG) on September 25, 1998. This group consists of experts on economic status from around the country. The Working Group decided early in the discussions that not all asset changes should be reconciled. Selection of the households and the individual asset components to be reconciled was guided by two primary considerations. First, we did not want to risk future respondent cooperation. Bothering respondents with trivial asset changes or with substantial asset changes that are easily explainable by things already known from prior interviews might well do this. Thus, after lengthy discussion of the merits of absolute versus proportional asset-change thresholds, we decided to set a minimum threshold for an individual net-worth component change of $50,000 for it to qualify for reconciliation. Also, we did not want to bother the respondent with reconciling apparent asset changes that result simply from changes in categorization of the asset. For example, it often happens that a respondent will classify an asset as a farm in one wave (perhaps because he planted some corn) and as real estate in the next wave (perhaps because he is considering subdividing the land for development). This would result in a large increase in real estate assets offset by a large decrease in the reported value of the farm. For this reason we set a threshold for aggregate net-worth change of $150,000—i.e. the absolute value of the wave-to-wave change in total household net worth must be at least this amount for the household to be eligible for reconciliation.

The second consideration is that we want our selection of households and net-worth items to be consistent with what it would be if we were to incorporate the reconciliation as an on-the-fly component of our CAI instrument in future waves of the HRS. One result of this is that we would not attempt to reconcile asset changes when there is a change in the financial respondent.[2] The asset values from one wave are preloaded only for the financial respondent for that wave and, as a result, there would be no preloaded asset values for a new financial reporter.[3]

After implementing these rules (technical details of implementation are available from the author as a research memorandum), we identified 2955 asset items in 1581 households (out of 11,583 households with the same financial respondents in 1998 and 2000) to be reconciled. We began contacting and interviewing these on March 26, 2001 and closed out the field effort on June 13, 2001. We obtained 1255 completed interviews. Once we remove the deceased (28), those we were unable to find or contact during the relatively brief field period (53), and those who were too sick or cognitively impaired to respond (25), our response rate was just over 91%. 40% of the non-respondents were cases in which the household was contacted but the interview could not be completed within the field period. Some of these would have been completed if the field period were as long as that of the main study (roughly 11 months).

---

[2] In households with two primary adults the financial respondent is identified by asking which of the adults is most knowledgeable of financial matters and plans for retirement.

[3] We are also very reluctant to reveal to one member of a household what another told us previously—something we would have to do if we attempted reconciliation when the financial respondent has changed.

2.1. Alternative Selection Criteria

Although a variety of research interests were represented in the ESWG, it is fair to say that interest in aggregate wealth accumulation and depletion was dominant. The selection rules chosen concentrate rather narrowly on those households and asset changes that have major impacts on aggregate wealth estimates. Other researchers may, however, be interested in the saving and consumption behaviours in the broader population. In this case selection criteria based on relative wave-to-wave change in reported assets might be more appropriate. Finally, a selection criterion based on some combination of absolute and relative change might offer a good compromise for the broad community of researchers. In Table I, I present various statistics for three specific selection criteria we might have used. These are:

1. Absolute: 1998 to 2000 change of $150,000 in aggregate and $50,000 in some specific category (i.e. the criterion we actually used).
2. Relative: 1998 to 2000 change of 175% or more.
3. Mixed: 1998 to 2000 change of $50,000 + absolute and/or 98% + relative.

The relative change percentages were selected to provide roughly equal numbers of eligible cases under the three alternative criteria. The first column of statistics pertains to the absolute change criterion we actually used. This yields 55 eligible respondents (3.7% of all eligibles) in the lowest quintile of the 2000 wealth distribution and 961 (65.9%) in the top wealth quintile. This is clearly a very wealthy subsample—accounting for more than half (57.5%) of all 2000 wealth and 91.1% of the variance in 1998–2000 wealth change. Column 2 of Table I presents the corresponding statistics for the sample derived using the relative change criteria. The reconciliation subsample under this rule would have been considerably poorer, with more than four in ten households being in the bottom wealth quintile, and would cover roughly one-fifth of the total wealth in 2000. Nevertheless, (a) the wealthiest households are selected at slightly more than their population prevalence, (b) more than three-quarters of the variance in absolute change is covered and (c) nearly 95% of the variance in relative change would be eligible for reconciliation. Finally, the mixed change criterion (column 3) would appear to offer a nice compromise. It results in over-representation of households at both extremes of the wealth distribution, covers more than a third

Table I. Characteristics of reconciliation subsample under three selection criteria (11,583 households with sample financial respondent in 1998 and 2000)

	Selection criterion		
	1. Absolute change	2. Relative change	3. Mixed change
Eligibles in bottom wealth quintile	55	625	361
	(3.7%)	(53.5%)	(25.2%)
Eligibles in top wealth quintile	961	295	558
	(65.9%)	(20.5%)	(37.5%)
Share of 2000 wealth	57.5%	20.9%	37.5%
Share of squared deviations of raw change	91.1%	77.8%	81.7%
Share of squared deviations of relative change	5.5%	93.5%	90.7%
Subsample size (% of total sample)	1581	1537	1589
	(12.6%)	(12.2%)	(12.7%)

Table II. Distribution of eligible, reconciled and erroneous changes by asset type

Net-worth item	# Eligible	# Reconciled	# Correct	# Erroneous
Debts	28	22	8	15
	1.0%	78.6%	36.5%	63.6%
Trusts	318	281	156	125
	10.8	88.3	55.5	55.5
Other assets	163	136	56	90
	5.5	83.5	33.8	66.2
Vehicles	35	30	13	17
	1.2	85.7	53.3	56.7
CDs	157	126	69	57
	5.3	80.3	52.8	57.2
Accounts	179	155	82	63
	6.1	81.0	56.6	53.5
Bonds	87	76	50	39
	3.0	87.5	58.7	51.3
Stock	552	367	202	165
	15.0	83.0	55.0	55.0
IRAs	295	252	135	118
	10.0	85.5	53.2	56.8
Business or farm	259	203	88	115
	8.5	81.5	53.5	56.6
Real estate	370	315	135	180
	12.6	85.1	52.9	57.1
First home	327	273	182	91
	11.1	83.5	66.7	33.3
First mortgage	91	85	56	29
	3.1	93.5	65.9	35.1
Second mortgage	6	5	3	2
	0.2	83.3	60.0	50.0
Equity loan	7	5	5	0
	0.2	71.5	100.0	
Mobile home	12	8	2	6
	0.5	66.7	25.0	75.0
Second home	155	128	65	63
	5.2	83.1	50.8	59.2
Loans second home	25	22	13	9
	0.8	88.0	59.1	50.9
Total	2955	2579	1298	1181
	100.0%	85.2%	52.5%	57.6%

of year 2000 wealth, and would result in reconciliation of households accounting for eight-tenths and nine-tenths of variance in absolute and relative change, respectively.

2.2. Distribution of the 2001 Reconciliation Sample

As noted above, Criterion 1 was used in selecting the 2001 asset-change reconciliation sample. Table II presents the distribution of eligible, reconciled and erroneous asset changes by asset type. By *reconciled* we mean that the discrepant asset values were presented to the respondent and the respondent had the opportunity to dispute and correct the record. By *erroneous* we mean that the respondent was presented with the discrepant values and said that the 1998 and/or 2000 value was wrong. Large and inexplicable changes between 1998 and 2000 were flagged for reconciliation in each of the 18 net-worth categories investigated. Given the volatility of the stock market over the

last few years it is not surprising that stocks are the most common 1998–2000 asset type eligible for reconciliation, with 552 changes flagged. This is 19% of all eligible asset changes observed. Of these we were able to reconcile 367 (83%) and found that 165 or nearly half were, according to the respondents, wrong. This fraction (55%) of reconciled asset changes which respondents identify as erroneous is typical of most asset types. On the low side the exception is first home value and on the high side are first mortgage and other debts. For the first home and first mortgage variables the percentages of questionable changes that were wrong are only 33.3% and 35.1%, respectively. This probably reflects the fact that big changes in house values over a two-year period really do occur.

Similarly large swings in remaining mortgage principal really do happen as well when respondents pay off or refinance their mortgages. At the other extreme the error rates for large changes in 'other assets' and debts are 66.2% and 63.6%, respectively. I suspect that respondents are very uncertain about the value of 'other assets', which include such things as jewels, coin and stamp collections, precious metal, etc. Although there may be similar confusion in the respondent's mind on how much they really owe at a point in time, I am less satisfied with this explanation for the high error rate in the original reports.

Another way of looking at the incidence of questionable asset changes, reconciliation rates and error rates is to categorize them not by the asset type but by the response types they involve. In each wave the response can be of three forms: (1) a zero report, (2) a bracket report and (3) a cardinal ('exact') report. For two waves there are eight interesting combinations of these report types (the zero–zero combination is not of interest here).

Table III presents the frequency of eligible and erroneous asset changes for each combination. The combination with the most eligible asset changes is the 'exact' 1998–'exact' 2000, with 1315 (55.6%) reconciliation-eligible asset changes. This is not surprising since the vast majority of all (non-zero) response combinations are of this type. What is a bit more surprising is that the error rate for this combination (29.5%) is substantially lower than that for any other combination. The second most common response type for reconciliation-eligible asset changes is 'exact' 1998 and zero 2000, with 635 eligible changes. This is a case in which a previously existing asset disappears without a trace. This is just one form of 'extra-marginal' change involving a zero in one wave and a non-zero response in the other. If we combine the elements of the first row with those of the first column we obtain a total extra-marginal count of 1362, so extra-marginal changes are slightly more prevalent than the 'exact–exact' combination.

Table III. Distribution of eligible and erroneous asset changes by response type in 1998 and 2000 [erroneous counts and percentages in brackets]

HRS 1998	HRS 2000		
	Zero	Bracket	Cardinal 'exact'
Zero	—	50 (1.5%) [18 (55%)]	635 (21.5%) [300 (57.2%)]
Bracket	121 (5.1%) [58 (57.9%)]	30 (1.0%) [12 (50.0%)]	171 (5.8%) [65 (38.0%)]
Cardinal 'exact'	566 (19.2%) [252 (55.5%)]	67 (2.3%) [35 (50.7%)]	1,315 (55.6 [388 (29.5%)]

There are two other notable patterns of asset change types in Table III. First, with respect to extra-marginal changes, asset disappearances are just about as frequent ($675 = 635 + 50$) as are asset appearances ($687 = 566 + 121$). This suggests a certain symmetry in asset omissions and/or acquisitions and dispositions across the waves. Secondly, suspicious changes involving brackets[4] are significantly more frequent for the 2000 wave ($322 = 121 + 30 + 171$) than for the 1998 wave ($137 = 50 + 30 + 67$). Why this should be the case is not immediately clear.

3. RECONCILIATION AND THE DISTRIBUTIONAL MOMENTS FOR LEVELS AND CHANGE

The main purpose of the 2001 asset-change reconciliation callback is to see to what extent quasi-dependent interviewing would affect our estimates of asset levels within waves and asset change across waves. There is little doubt that the asset measures contain considerable amounts of error that results in over-dispersion. From examining the interviewer written comments it is also clear that errors sometimes occur at the very extremes of the distributions, and when they do they can be very serious.[5] Another indication that error is concentrated at the extremes of the distributions is that initial item non-response in the form of 'don't knows' are subsequently very often found to be in the highest bracket of the unfolding bracket sequence used to reduce the item non-response. Such errors can seriously skew the distributions and will reduce the relative frequency of intermediate deviations from the mean, leading to increased kurtosis.

3.1. House Value

Table IV(a) presents the distributional moments for house value for those reconciliation respondents with a positive value in either 1998 or 2000.[6] Housing is the most common form of non-pension net worth and for most households the most important. The first column of Table IV(a) refers to the original house value recorded in the 2000 interview and the second is the 2000 value recorded after reconciliation in 2001. The third and fourth columns are the corresponding 1998 values and the final two columns relate to the 1998–2000 change in the original and reconciled values. The first thing to note about the 2000 values is that the average is not greatly affected by reconciliation. It is roughly $200,000 both before and after the callback. The higher moments of the distribution, however, are all reduced tremendously by reconciliation. The variance is more than cut in half, the skewness is quartered and the kurtosis is reduced by more than a full order of magnitude. Surprisingly, the various percentile breakpoints are not affected, which can only mean that the effect of the reconciliation must be taking place either below the first percentile or above the 99th. For housing assets the same general pattern of results holds for the 1998 values. There is some evidence here for shrinkage of the upper tail of the distribution, but otherwise the effects of reconciliation of housing assets in 1998 is the same as in 2000. As we will see below, this is the

---

[4] Bracket reports in the HRS occur when a respondent will not or cannot provide a cardinal report. In such cases they are asked a series of questions of the form 'would it amount to less than, more than or about $x'. See Juster and Smith (1997) for details.

[5] In the 1994 HRS, for instance, one interviewer entered the value of $3,500,000 for a respondent's total medical expenditures and wrote in the comment field that the expenditures were $35,000. In the response field she had erroneously entered dollars and cents rather than simply dollars.

[6] Respondents who stated in 2000 that they either bought or sold a home were excluded from the asset-change reconciliation sample. This means that changes in house value should reflect appreciation or depreciation of the house.

exception rather than the rule. Perhaps the most interesting results of reconciliation in Table IV(a) are those related to the 1998–2000 change measures. Here the mean is substantially increased as a result of reconciliation and the variance is reduced to one-ninth its original value. The kurtosis is also reduced by more than one order of magnitude. The percentile breakpoints indicate that most of the effect for change is at the lower end of the distribution. Subsequent examination of the reconciled data shows that this is due mostly to extra-marginal changes—six respondents in 2000 failed to report home ownership in the original interview but upon reconciliation did 'fess up' to owning one; one of these omitted homes was allegedly worth $900,000.

Table IV. Distributional moments

(a) For house value [those reconciliation respondents owning a home in either 1998 or 2000]

	2000 Original level	2000 Reconciled level	1998 Original level	1998 Reconciled level	1998–2000 Original change	1998–2000 Reconciled change
Mean	222	217	208	195	13.5	22.2
Variance	135e + 9	57.8e + 9	132e + 9	52.5e + 9	167e + 9	18.0e + 9
Skewness	17.17	5.35	17.95	5.08	0.010	1.55
Kurtosis	535.5	32.53	555.5	53.7	511.2	36.22
1st Centile	0	0	0	0	−599	−325
10th Centile	50	50	60	60	−95	−55
50th Centile	150	150	150	150	10	10
90th Centile	500	500	375	350	100	100
99th Centile	1500	1500	1500	1000	500	500

(b) For stock assets ($1000s) [those reconciliation respondents having stocks in either 1998 or 2000]

	2000 Original level	2000 Reconciled level	1998 Original level	1998 Reconciled level	1998–2000 Original change	1998–2000 Reconciled change
Mean	335	285	279.6	273.5	55.0	10.2
Variance	1570e + 9	31.5e + 9	659e + 9	501e + 9	1,660e + 9	501e + 9
Skewness	16.36	5.15	8.93	8.67	9.78	−8.11
Kurtosis	355.7	53.75	98.71	100.6	232.0	118.2
1st Centile	0	0	0	0	−1200	−1200
10th Centile	0	0	0	0	−263.5	−200
50th Centile	75	75	75	80	0.750	0
90th Centile	900	800	600	600	350	290
99th Centile	5000	2750	3000	3000	2500	1500

(c) For total net worth ($1000s) [all 1255 reconciliation respondents]

	2000 Original level	2000 Reconciled level	1998 Original level	1998 Reconciled level	1998–2000 Original change	1998–2000 Reconciled change
Mean	1386	1258	1052	1057	335	201
Variance	9220e + 9	3690e + 9	2680e + 9	2280e + 9	7510e + 9	2020e + 9
Skewness	10.55	5.80	5.61	5.01	9.52	5.33
Kurtosis	156.3	59.7	31.2	25.5	155.8	87.7
1st Centile	−1.5	−1.6	−2.5	8	−5317	−3225
10th Centile	108	120	153	151	−805	−597
50th Centile	659	690	581	585	227	150
90th Centile	2968	2837	2328	2506	1285	996
99th Centile	11,500	8602	9565	8230	8607	5563

Table IV. (*Continued*)

(d) For total net worth ($1000s) [all 11,583 HRS households with the same financial respondent in 1998 and 2000]

	2000 Original level	2000 Reconciled level	1998 Original level	1998 Reconciled level	1998–2000 Original change	1998–2000 Reconciled change
Mean	368	354	332	333	36.3	21.75
Variance	1770e + 9	1140e + 9	804e + 9	761e + 9	1170e + 9	579e + 9
Skewness	22.08	18.62	15.70	15.86	25.40	27.48
Kurtosis	734.9	635.9	506.0	543.5	1073	1680
1st Centile	−10	−10	−18	−16	−1356	−1223
10th Centile	0	0	1.3	1.5	−173	−144
50th Centile	110	110	121	121	0	0
90th Centile	774	792	736	740	204	179
99th Centile	4208	3880	3285	3286	1798	1477

A useful exercise is to solve the variance formula presented in the Introduction for the unknown true and error correlations. We can do this by assuming, for illustrative purposes, that reconciliation reduces the error variances to zero. When we do so, we obtain an estimate of the true wave-to-wave correlation of housing values of 0.83 and of wave-to-wave error correlation of only 0.1070. Furthermore, the ratios of error to true variance are 1.35 and 2.11 for the 2000 and 1998 original reports, respectively. Since some measurement error certainly remains after reconciliation, the actual correlations and ratios are probably even more extreme.

3.2. Value of Stocks

The effect of the reconciliation on the distributional moments for house value is somewhat atypical in that both the 2000 and 1998 distributions were substantially affected. In this respect stocks provide a more typical picture of the effects of reconciliation. Table IV(b) presents the distributional moments and percentile breakpoints for stock assets. As with home value, reconciliation resulted in major reductions in the first four moments of the 2000 stock value distribution as well as in the centile breakpoints at the upper end of the distribution. When we solve the variance formula as in the preceding subsection we obtain an estimate of the correlation of the true component of 0.52 and of the error components of 0.55. The 1998 distribution, however, was hardly affected at all. This is typical of most of the other net-worth components (tables for all 18 net-worth components are available from the author). What is not typical, however, is the estimated error-to-true variance estimate of nearly 50 in 2000 and the large reduction in the 1998–2000 change in stock values that comes about as a result of the 2001 reconciliation.

The fact that the callback was conducted anywhere from six to 18 months after the original 2000 data collections may introduce a bias which would not be encountered with an immediate on-the-fly reconciliation. This potential bias is that respondents may tend to value the asset not at its worth at the time of the original interview, but at a more recent time. The wording of the questions left the timing of the valuation unclear. A typical question, for instance, might begin: 'According to our records in 2000 you had stocks worth $x...'. In retrospect we should have said: 'According to our records in June of 2000, when we last talked with you, you had stocks worth...'. The results in Table IV(b) for 'Stocks' look suspicious to me from this point of view. Most of the original stock data for 2000 were provided prior to July 2000—a period when the stock market was near its all-time high. Most of the reconciled data were provided after April

2001, after many stocks had suffered dramatic declines in price. The reduction in the 1998–2000 increase in stock value from $55,000 in the original data to $10,200 in the reconciled data may be more a reflection of the 2000–2001 declines in the market than a reflection of better reporting in the dependent callback. This may also be responsible for decreases in estimated growth of value of IRAs and even Trusts, since the portfolios of these often have considerable assets in the form of stock. I will return to this in Section 4, below.

3.3. Total Net Worth

Although the individual net-worth components are interesting, the bottom line on judging the efficacy of reconciliation requires us to examine the total of the 18 non-pension wealth components. Tables IV(c) and (d) provide the distributional moments for these totals—Table IV(c) does so for the 1255 reconciliation sample respondents and Table IV(d) for the entire HRS sample of 11,583 respondents with the same financial respondent in both 1998 and 2000. For the reconciliation respondents, reconciliation resulted in a modest decrease in the mean level of wealth in 2000 but hardly any change in 1998 measured wealth. Even more dramatic is the reduction in the higher-order distributional moments in 2000. The variance is cut to roughly 50% of its original value, the skewness to about 56% of its original value and the kurtosis to less than 50%.

The mean estimated 1998–2000 change in net worth was reduced from 335 to 201 (i.e. by 50%) and the variance from 7510 to 2020 (i.e. by 73%). The higher-order moments of the change distribution are also greatly reduced by reconciliation. The centile breakpoints suggest that most of the effect of reconciliation is taking place in the tails of the distribution. Surprisingly, we obtain almost the same conclusions with respect to the mean and variance of change when we examine the entire sample of 11,583. The mean change (not shown) drops from 36.6 thousand dollars to 21.8 thousand (a 40% decrease) and the variance drops by 51% (from $1.17e + 12$ to $0.58e + 12$). This is a remarkable result. Reconciling less than one-ninth of the sample cuts the variance in wave-to-wave change in net worth in half.

3.4. Length of Recall and Changes in Economic Environment

Another aspect of the effects of reconciliation apparent in Tables IV(b) and (c) (as well as in corresponding tables for most other asset types) is that it affects the most current data (2000) much more than the older data (1998). Both probit and Tobit analyses of reconciliation (results not shown) reveal strong and significant negative effects of the number of months between the 1998 report and the reconciliation report in 2001—i.e. the more distant the report the less likely the respondent is to correct the record and the smaller is the absolute value of the correction if they do. This effect does not appear for the 2000 reconciliation.[7] Furthermore, changes in the Dow-Jones Industrial Average between the date of the original report and the date of reconciliation, *something that should be totally irrelevant*, exert a strong and significant positive effect on the probability of changing the report after reconciliation for 1998 but not for 2000. My interpretation of these results is that the respondent's uncertainty of the truth increases with the length of recall and, when uncertain, respondents tend to accept the previously recorded amounts. They are also more apt to be influenced by irrelevant changes in outside conditions. This leads me to

---

[7] Very similar results were found in the 1983 Panel Study of Income Dynamics Validity Study (see Duncan and Hill, 1985).

conclude that timing is critical to the success of reactive-dependent interviewing in improving reports of asset amounts—the sooner a report is reconciled with previously reported data, the better.

4. EFFECTS OF RECONCILIATION ON MEASURES OF ASSOCIATION

The descriptive statistics just presented show that the reconciled asset data are substantially different from the original. The question is: are they any better? Lacking direct external validating data there is little we can do beyond examining the effects of reconciliation on correlations with other survey measures and on regression coefficients. *Ceteris paribus*, if reconciliation is actually improving the quality of the wealth measures, and especially of the wealth change measures, we would expect the strength of these associations and effects to increase. Table V presents the correlation of change in non-pension net worth with a host of other HRS survey measures. With only a couple of exceptions the reconciled change data are more strongly correlated with nine of the 13 variables examined. The main exceptions to this are for marital status (whether married) and education, for which reconciliation has little effect, and for receipt of pension income and the

Table V. Correlations of original and reconciled net-worth change [11,578[a] respondents: *p*-values in parentheses]

	Original 1998–2000 change in total net worth	Reconciled 1998–2000 change in total net worth
Age	−0.0251**	−0.0298**
	(0.0096)	(0.0013)
Gender (female)	−0.0235*	−0.0256**
	(0.0113)	(0.0059)
Married	0.0235*	0.0212*
	(0.0116)	(0.0226)
Education (years)	0.0377**	0.0291**
	(0.0000)	(0.0017)
Family income	0.0101	0.0555**
	(0.2775)	(0.0000)
Receiving social security	−0.0225*	−0.0276**
	(0.0155)	(0.0029)
Receiving pension income	−0.0201*	−0.0152
	(0.0153)	(0.1027)
Covered by Medicare	−0.0285**	−0.0389**
	(0.0022)	(0.0000)
Covered by employer-provided health insurance	0.0266**	0.0362**
	(0.0052)	(0.0001)
Subjective probability of living to 85+	0.0170+	−0.0151
	(0.0666)	(0.1297)
Subjective probability of $100k bequest	0.0555**	0.0523**
	(0.0000)	(0.0000)
Subjective probability of nursing home	0.0111	0.0191*
	(0.2320)	(0.0398)
Change in household work hours: 2000–2002	−0.0053	−0.00190+
	(0.5951)	(0.0565)

[a] Five cases were eliminated from this analysis because of missing data on age.
[†], [*], [**] Significant at the 90, 95 and 99 percent level of confidence, respectively.

respondent's subjective probability of surviving to age 85. For these there is a lower significance level for the reconciled change than for the original change.

4.1. Wealth Change as an Independent Variable

The effect of reconciliation on the association of change in work hours and change in wealth, seen in the last entries of Table V, while not extremely strong, is of particular interest since much of the measurement error literature deals extensively with the effects of error in independent variables on regression parameter estimates. We can imagine a model of the form:[8]

$$\Delta H_i^t = \alpha + \beta \Delta W_i^t + \varepsilon_i^{\Delta_t} \tag{3}$$

Here exogenous changes in wealth are hypothesized to reduce work hours (i.e. $\beta < 0$).[9] Indeed, much has been made recently, at least in the popular press, about the wealth effect of the run-up (and decline) of stock prices on retirement aspirations, plans and decisions. Under classical assumptions (the validity of which we will investigate below), measurement error in the ΔW_i^t will result in β being biased towards zero. Table VI presents regression coefficients of a version of this model, in which explicit control is made for capital gains in various asset types, i.e.

$$\Delta H_i^t = \alpha + \beta \Delta W_i^t + \Gamma K_i^t + \varepsilon_i^{\Delta_t} \tag{4}$$

Table VI. Change in labour supply as a function of gross saving and capital gains: concurrent and lagged by reconciliation status

	Concurrent change[a] (HH hours 2000—HH hours 1998)		Next period change[b] (HH hours 2002—HH hours 2000)	
	Δ Original 1998–2000	Δ Reconciled 1998–2000	Δ Original 1998–2000	Δ Reconciled 1998–2000
Gross saving ($\Delta W_{98,2k}$)	−0.0213+	−0.0583**	0.0008	−0.0501*
	(−1.69)	(−2.59)	(0.06)	(−2.08)
Capital gains				
Farm/business	0.1182**	0.0885*	0.0068	0.1191*
	(3.00)	(2.57)	(0.17)	(2.06)
Stock	0.0385*	0.0650*	−0.0150	0.0391
	(1.98)	(2.29)	(−0.70)	(1.30)
Real estate	0.0267	0.0280	−0.0051	−0.0155
	(0.92)	(0.52)	(−0.15)	(−0.25)
Housing	0.0131	0.0602	0.0093	−0.0525
	(0.30)	(1.12)	(0.21)	(−0.75)
Adj. R-square	0.05%	0.03%	−0.05%	0.05%
F	2.25*	1.67	0.19	2.10+

[a] Sample = 11, 583 stable households with same financial respondent 1998 and 2000.
[b] Sample = 10, 135 stable households 1998 and 2000 surviving to 2002.
[†], [*], [**] Significant at the 90, 95 and 99 percent level of confidence, respectively.

---

[8] See Bound *et al*. (2001) for a review of this literature. We should be careful not to take this model too seriously, since we can just as easily imagine causation running in the opposite direction—pending retirement may present a strong incentive to save.
[9] An alternative hypothesis is that wealth changes reduce labour supply with a lag. Thus the effect of wealth changes between 1998 and 2000 will show up in hours changes between 2000 and 2002. We will examine both of the implied specifications.

where Γ is a vector of parameters relating the vector of capital gains (K), of various sorts, to changes in work hours. This allows us to test for differential effects of active saving (β) and the various forms of capital gains on changes in labour supply. The first two columns of Table VI present estimated effects of wealth changes on concurrent labour supply, while the second two columns present estimated effects if wealth changes operate with a lag. The first column of results for concurrent hours change pertains to the asset changes as originally measured, and the next column pertains to the reconciled asset data. Comparing the two we see that reconciliation results in a much stronger 'effect' of total wealth change on hours change. Since we are controlling for capital gains we are tempted to interpret the coefficient on gross saving as the association of *active* saving on changes in household work hours supplied. Assuming the reconciled data to be accurate, the quite noticeable attenuation of the coefficient obtained with the noisier unreconciled data implies a 'reliability ratio' of less than 36% for unreconciled gross asset change.[10] Interestingly, with both types of data in the concurrent model as well as with reconciled data in the lagged model the effect of saving on hours change is significantly negative and significantly different from the positive effects of capital gains in stocks and businesses or farms. There is even some evidence in the concurrent model for a stronger positive effect of capital gains in housing assets.

Table VII. Regression analyses of original and reconciled gross saving ($W_{2000} - W_{1998}$) [11,578[a] households with same financial respondent: *t*-ratios in parentheses]

	OLS (dependent variables in $1000s)		Median regression
	Original gross saving 1998–2000	Reconciled gross saving 1998–2000	Original gross saving 1998–2000
Constant	−77.5989	−3.1709	5.8873
	(−0.78)	(−0.05)	(1.61)
Age	−1.3173	−1.1578	−0.2597**
	(−1.33)	(−1.52)	(−7.81)
Education	7.7055*	1.7076	0.2669**
	(2.55)	(0.77)	(2.76)
Married	23.6585	3.3790	3.3052**
	(1.10)	(0.22)	(5.99)
Family income	−0.1778	0.2525**	0.1595**
	(−1.38)	(2.80)	(37.86)
Probability of bequest $100k+	1.5128**	0.9083**	0.1821**
	(5.11)	(5.36)	(20.03)
Probability of nursing home	0.3587	0.3955	−0.0330*
	(0.77)	(1.21)	(−2.32)
Probability of Living to 85+	0.3297	0.1511	0.0099
	(0.90)	(0.55)	(0.89)
Adj. R-square	0.37%	0.39%	0.38%
Root MSE	1079.80	759.56	—
1% Influence index	92.5%	90.9%	37.5%

[a] Five cases were eliminated from this analysis because of missing data on age.
*, ** Significant at the 95 and 99 percent level of confidence, respectively.

---

[10] See Fuller (1987) for a thorough discussion of classical measurement-error models and for definitions of various statistics such as the reliability ratio ($= \sigma^2_{true}/(\sigma^2_{true} + \sigma^2_{error})$).

4.2. Wealth Change as the Dependent Variable

Although reconciliation improves the empirical validity of the asset-change data, the influence of extreme changes is still sufficiently large to obscure the effects of most covariates. Table VII presents regression analyses of the 1998–2000 net-worth change for the original and reconciled data. While the regression for the reconciled data is significantly better than for the original data (primarily because of the coefficient on income), its fit is still not very impressive. The bequest result is interesting, but should not be taken too seriously—it is probably a reflection of the fact that the respondent has to have $100,000 to potentially bequeath. The final row in the table presents the '1% Influence Index' of cases in the top and bottom percentile of the distribution.[11] The 92.5% figure for the original data can be interpreted as saying that the top and bottom 1% of the original-change distribution accounted for 92.5% of the OLS objective function (the total sum of squares) for the constant-only model. This means that these few cases dominate the regression analysis. Reconciliation reduces the tail-influence index to 90.9%, but the few extreme cases still dominate the analysis. To see the structure of net-worth change any more clearly, the analyst must do something to reduce the influence of outliers. The final column of figures in Table VII presents the results of a median-regression version of the saving model in which the objective function is the sum of absolute deviations rather than of squared deviations. The 1% tail influence index is reduced to 37.5% and fit of the model is improved—virtually all of the covariates become significant, primarily because the standard errors (not shown) are greatly reduced. One interesting substantive result in this specification is that the effect of anticipating entering a nursing home becomes significantly negative—i.e. those anticipating entering a nursing home do not save as much as others or may even dis-save. To the best of my knowledge this is the only evidence of the asset-spend-down hypothesis that has been seen with HRS data.

4.3. Error Correlations

Finally, Bound *et al.* (2001) strongly encourage researchers to test the 'classical' assumptions of measurement error models whenever validating data is available. While far from perfect, our reconciled data is a form of validating data and with it we can construct crude estimates of measurement error according to $\hat{\varepsilon}_i^t \equiv W_{t,i}^O - W_{t,i}^R$. Table VIII presents the correlation of these estimated errors for 1998 and 2000 with each other, with their components and with other variables of interest to analysts. The correlation of errors across waves is negative and significant which,

Table VIII. Measurement error correlations

	ε^{2000}	W_{1998}^O	W_{1998}^R	W_{2000}^O	W_{2000}^R	$\Delta W_{1998,2000}^R$	Family income$_{1998}$
$\hat{\varepsilon}^{1998}$	−0.0802*	**0.2366***	−0.0553*	−0.0597*	−0.0188*	0.0253*	0.1310*
$\hat{\varepsilon}^{2000}$	1.0000	0.0666*	0.0915*	**0.5965***	**0.0510***	−0.0333*	0.0136
$\hat{\varepsilon}^{2000} - \hat{\varepsilon}^{1998}$	0.9598*	−0.0128	0.0985*	0.5703*	0.0530*	−0.0385*	−0.0286*

* Significant at the 95 percent level of confidence.

[11] This is simply the fraction of total sum of squares or of absolute deviations accounted for by the cases in the 1.0% tails of the distribution.

as we can see from the variance of change formula in the Introduction, will clearly reduce the signal-to-noise ratio in wealth-change data.

Furthermore, by construction we would expect the correlations printed in bold in Table VIII to be strongly positive and negative for original and reconciled reports, respectively, and roughly equal in absolute size. The error correlations with the original measures are as expected, but those with the reconciled measures are small (but significantly) negative for 1998 and small (but significantly) *positive* for 2000. This means that those with high 'true' wealth in 1998 tended to reduce their reports from the original value upon reconciliation (in 2001), whereas those with high wealth in 2000 tended to increase their reports from the original value. Why this should be is a mystery to this author, but what is clear is that the classical assumptions of measurement error models are highly suspect.

The last column in Table VIII presents the correlations of error and error change with family income in 1998. The correlation of the change in error and family income explains the effects of reconciliation on the income coefficient noted in our discussion of Table VII. The relevant regression coefficient is:

$$\hat{\beta}_{s,\gamma} = (Y'Y)^{-1}(Y'S^{orig}) = [Y'Y]^{-1}[Y'(S^R + \Delta\varepsilon)] = [Y'Y]^{-1}[Y'S^R] + [Y'Y]^{-1}[Y'\Delta\varepsilon]] \qquad (5)$$

where Y is income and S is saving. The rightmost term ($Y'\Delta\varepsilon$) corresponds to the lower right entry in Table VII (-0.0286) and the result is that the regression coefficient of saving on income is biased towards $-\infty$ by the negative correlation of income and measurement error in wealth change. It is important to note that this bias is *not* a manifestation of measurement-error induced attenuation discussed so extensively in the measurement-error literature. Instead, it is a result of bias due to violations of the classical measurement-error assumptions of that literature and could be of either sign and of any magnitude. Interestingly, this is exactly the same conclusion reached 20 years ago in our analysis of the Panel Study of Income Dynamics Validity Study (see Duncan and Hill, 1985, p. 529).

5. CONCLUSIONS

Panel measures of change are often so noisy as to seriously affect their utility in statistical models. The change in reporting error (noise) dominates change in true values (signal). In this paper I have shown that asking a small subsample (12%) of Health and Retirement Study respondents with large and inexplicable wave-to-wave changes in net-worth components to reconcile the recorded differences reduces the variance in total net worth by half for the entire sample. I argue that most of this reduction is the result of reduction in error variance, and show that the reconciliation interview (a form of reactive-dependent interviewing) substantially improves the empirical validity of the data. This improvement shows up in the form of increased magnitude and significance of simple correlations of net-worth change (gross saving) with other survey measures, with which theory suggests it should be related. It also shows up in multivariate models, both when net-worth changes or gross saving is an independent variable and when it is the dependent variable. Classical measurement-error models show that under a rather strict set of assumptions error in independent variables will attenuate regression coefficients. My findings, when gross saving measured by wave-to-wave change in total net worth, are consistent with this. Under the classical assumptions, measurement error in the dependent variable should simply reduce the estimated goodness of fit and the significance of independent variables. Again there is some evidence of this in my analysis—the

root mean square error of the model using reconciled data is 30% lower than that using unreconciled data, and the adjusted R-square of the model does increase slightly after reconciliation. The data also, however, cast serious doubts on the classical measurement-error model assumptions. Taking the reconciled reports as valid and using them to construct measurement error estimates I obtain strong and significant correlations of these errors with the level of gross saving and its component, as well as with a key covariate—family income. This can result in a bias which is quite distinct from the attenuation bias of the classical model in that it can be of either sign and of any magnitude.

As a result of this research the HRS now incorporates a form of reactive-dependent interviewing in its main computer assisted questionnaire. Reconciliation is now triggered whenever the current report of a net-worth component is seriously at odds with the prior-wave report and is implemented at the end of the biennial interview after the normal net-worth questions. This research also suggests, however, that the HRS might want to re-evaluate the rules used in triggering reconciliation. These are currently based on absolute change in asset values, but a mixed trigger based on both absolute and relative change may be superior.

REFERENCES

Bound J, Brown C, Mathiowetz N. 2001. Measurement error in survey data. In *Handbook of Econometrics*, vol. 5, Heckman J, Leamer A (eds); Amsterdam: North Holland; pp. 3705–3843.
Brown A, Hale A, Michaud S. 1998. Use of computer assisted interviewing in longitudinal surveys. In *Computer Assisted Survey Collection*, Couper M, Baker R, Bethlehem J, Clark C, Martin J, Nicholls W, O'Reilly J (eds). John Wiley & Sons: New York.
Dibbs R, Hale A, Loverock R, Michaud S. 1995. Some effects of computer assisted interviewing on the data quality of the Survey of Labour and Income Dynamics. In *Proceedings of the International Conference on Survey Measurement and Process Quality*. American Statistical Association: pp. 175–177.
Duncan GJ, Hill DH. 1985. An investigation of the extent and consequences of measurement error in labor-economic survey data. *Journal of Labor Economics* **3**(4): 80–101.
Fuller WA. 1987. *Measurement Error Models*. John Wiley & Sons: New York.
Hill DH. 1995. The relative empirical validity of dependent and independent data collection in a panel survey. *Journal of Official Statistics* **10**: 359–380.
Hsiao C. 1986. *Analysis of Panel Data*. Cambridge University Press: Cambridge.
Juster FT, Smith JP. 1997. Improving the quality of economic data: lessons from the HRS and AHEAD. *Journal of the American Statistical Association* **92**(440): 1268–1278.
Lansing J, Morgan J. 1971. *Economic Survey Methods*. Survey Research Center: Ann Arbor, MI.
Mathiowetz N, McGonagle K. 2000. An assessment of the current state of dependent interviewing in household surveys. *Journal of Official Statistics* **16**: 501–518.