Determinants of wealth fluctuation: changes in hard-to-measure economic variables in a panel study

Pfeffer, Fabian T.; Griffin, Jamie

methods, data, analyses | Vol. 11(1), 2017, pp. 87-108 DOI: 10.12758/mda.2016.015

Determinants of Wealth Fluctuation: Changes in Hard-To-Measure Economic Variables in a Panel Study

Fabian T. Pfeffer & Jamie Griffin
University of Michigan

Abstract

Measuring fluctuation in families' economic conditions is the *raison d'être* of household panel studies. Accordingly, a particularly challenging critique is that *extreme* fluctuation in measured economic characteristics might indicate compounding measurement error rather than actual changes in families' economic wellbeing. In this article, we address this claim by moving beyond the assumption that particularly large fluctuation in economic conditions might be too large to be realistic. Instead, we examine predictors of large fluctuation, capturing sources related to actual socio-economic changes as well as potential sources of measurement error.

Using the Panel Study of Income Dynamics, we study between-wave changes in a dimension of economic wellbeing that is especially hard to measure, namely, net worth as an indicator of total family wealth. Our results demonstrate that even very large between-wave changes in net worth can be attributed to actual socio-economic and demographic processes. We do, however, also identify a potential source of measurement error that contributes to large wealth fluctuation, namely, the treatment of incomplete information, presenting a pervasive challenge for any longitudinal survey that includes questions on economic assets. Our results point to ways for improving wealth variables both in the data collection process (e.g., by measuring active savings) and in data processing (e.g., by improving imputation algorithms).

Keywords: wealth, panel study, active savings, measurement error, imputation

1 Motivation

Our understanding of families' economic wellbeing depends not only on how well we capture their current socio-economic conditions but also their movement within the economic hierarchy across time. In fact, the measurement of fluctuation in families' economic conditions could be considered the primary *raison d'être* of household panel studies (Duncan, 1984). In this research note, we reveal some of the factors that contribute to or jeopardize the ability of family household panel studies to accurately capture the changing economic fortunes of families. Doing so is particularly pressing in the context of an emerging new field of empirical inquiry: After decades of research on the dynamics of family income, recent scientific and public debate is increasingly focused on family wealth, or net worth, as a different and important dimension of economic wellbeing (e.g., Pfeffer & Schoeni, 2016; Piketty, 2013). The dynamics of wealth are of particular interest, for instance, to understand families' ability to smooth consumption during times of economic distress (Deaton, 1991) and to provide intergenerational support both in terms of investing in the young and caring for the elderly (Conley, 2001).

However, wealth information can be challenging to collect, and panel surveys that seek to measure its fluctuation over time face additional challenges (Bucks & Pence, 2015). In particular, a number of researchers have noted that wealth data tend to be noisier than many other economic data and have suggested that extreme fluctuation in wealth may result from measurement error (Bosworth & Smart, 2009; Hill, 2006; Venti, 2011). Here, we assume that even extreme wealth fluctuation is driven partly by real economic changes and partly by measurement error and our empirical analyses demonstrate the relative role of potential factors on both sides. Specifically, we consider households' demographic changes, economic behaviors and circumstances on the one hand, and two potential sources of measurement error on the other hand: "Observational errors" that might stem from a change in survey respondents and, more importantly, "errors of non-observation" (Groves, 2004) that arise from item nonresponse and its handling in the data processing phase.

We analyze data from the Panel Study of Income Dynamics (PSID, 2015), a study that not only has the distinguished record of being the world's longest-running nationally representative household panel study, but that also – and impor-

Acknowledgments

Funding for this research was provided by the National Institute on Aging (R01AG019802). The authors thank Vicki Freedman and Charlie Brown for helpful comments on an earlier version of this manuscript as well as the editors and reviewers.

Direct correspondence to

Fabian T. Pfeffer

Institute for Social Research, 426 Thompson Street, Ann Arbor, MI 48104, USA

E-mail: fpfeffer@umich.edu

tantly for this project – began to field a detailed survey module on families' assets in 1984. Our analyses identify some of the successes and limitations of the PSID asset module and, more generally, inform both data collection and data processing strategies for other household panel studies.

We begin by briefly summarizing some of PSID's main strategies for collecting, editing, and processing wealth data. We then describe our sample, main variables, and analytic approach. Our empirical findings address the distribution of between-wave changes in net worth and their predictors. We conclude by discussing the implications these findings have for the longitudinal collection of high-quality wealth information utilized by a rapidly growing field of empirical research.

2 Wealth Measurement in the PSID

The PSID started in 1968 and has collected a large set of socio-economic indicators for families and their descendants every year until 1997 and every other year since then. In 1984, it implemented a detailed module to measure families' assets. This module was repeated every five years until 1999 and every wave since then, amounting to a total of 12 waves of wealth data by 2015. The specific assets that form part of these data are listed later; here, we describe some of the strategies PSID employs during data collection, editing, and processing to reduce measurement error in its wealth variables. Many of these strategies were implemented in the first wealth survey of 1984 and were then state-of-the art. Some of these strategies still are; however, as we will show, others might be ripe to revisit given more recent methodological advances.

2.1 Data Collection: Unfolding Brackets

For each asset, respondents are first asked whether they own such asset (e.g., a home). For those who answer yes, the follow-up question asks about the value of the asset, sometimes with separate questions about the gross value (e.g., current home value) and the outstanding debt held against the asset (e.g., mortgages). To minimize the incidence of missing data in the collection of asset values, the PSID introduced a surveying technique that has become known as the "unfolding bracket" approach and that is now in use in a range of other major surveys (e.g., the Health and Retirement Study [HRS]). Respondents who report that they do not know an exact asset value receive a series of follow-up questions that ask them to report whether the value falls within certain pre-specified ranges ("brackets") (Juster & Smith, 1997). These brackets "unfold" as the interviewer asks for a dependent sequence of threshold values (e.g., "Does [X] amount to $10,000 or more?" If yes: "Does it amount to $50,000 or more?" If no: "Does it amount to $1,000 or more?"). In the PSID, this technique helps keep the prevalence of item nonresponse across a variety of different assets relatively low (reported below in Table 1). However, it also requires the assignment of a continuous value within those reported brackets, especially if individual asset components are to be cumulated to create a measure of total net worth.

2.2 Data Editing: Individual Lookups

In the data editing process, the PSID attempts to correct errors of observation that arise from either respondents or interviewers by investigating outlying responses and reconciling them with other information collected in the same or prior waves through individual lookups – a labor- and time-intensive process. Importantly, the outlying values for a given variable are defined only with respect to the distribution of that variable within the current survey wave. Conversely, other studies incorporate prior-wave information in the editing stage or even during data collection. For instance, the HRS preloads wealth values from the prior-wave interview and asks respondents to reconcile conflicting responses between the current and prior wave. In 2012, this procedure identified a small number of cases (≤ 2.5%) who corrected errors in either the prior or current wave.

2.3 Data Processing: Imputation

Finally, and most important for the purpose of this contribution, the PSID applies imputation procedures to fill in missing continuous asset values arising from item nonresponse and bracketed responses. Random hot-deck imputation procedures are used in the following sequence of steps (see also PSID, 2013, pp. 41-42): First, when a respondent does not report whether or not an asset (debt) is held, a yes or no value is randomly assigned with probabilities equal to the distribution of observed yes or no values. Second, for those reporting neither a continuous nor a bracket response for the value of the asset (debt), a bracket (e.g., $10,000 - $50,000) is randomly selected with selection probabilities equal to the distribution of observed brackets. Finally, all respondents who do not provide a continuous value for the asset (debt) (steps 1 and 2) are assigned a continuous value by randomly selecting an observed value within a given bracket and with selection probabilities equal to the distribution of observed continuous values within the respective bracket.[1]

Table 1 reports the share of cases with unknown continuous asset values for each wealth component, that is, those to which the described imputation procedure is applied (in years 2005 and 2007 for reasons described later). The extent of imputation differs substantially across wealth components (upper panel of Table 1), with

1 The imputation approach differs somewhat for home equity as described in detail elsewhere (PSID, 2013, pp. 55-56).

Table 1 Item Nonresponse in Wealth Components and Net Worth
N=7,051

Wealth Component	Variable Names	Share of Item-Nonresponse (%) 2005	2007	Overall
Vehicles/motor homes/trucks/etc.	S713A / S813A	7.9	9.0	14.0
Checking/savings/money order/etc.	S705A / S805A	7.1	7.3	11.6
Retirement wealth (annuity/IRA)	S719A / S819A	4.9	4.3	7.7
Home equity (value-mortgages)	S720A / S820A	4.2	4.7	7.6
Stocks/mutual funds/etc.	S711A / S811A	4.7	4.6	7.6
Other financial assets (bond funds/estate/etc.)	S715A / S815A	4.1	3.9	7.2
Farm and business wealth	S703A / S803A	3.0	3.5	5.3
Other debt (credit card/student loans/etc.)	S707A / S807A	1.5	1.6	2.8
Other real estate	S709A / S809A	1.5	1.6	2.8
Across all components (= net worth measure)	(S717A / S817A)			
Average		4.3	4.5	7.4
Cumulative				
Zero		75.8	74.0	62.4
One		15.6	17.0	19.1
Two		4.7	5.5	8.5
Three or more		3.9	3.5	10.0

Note: The overall column reports the share of cases with a specific wealth component imputed in either 2005 or 2007 (or both) and the total number of components missing across both years; N=7,051

the largest share of cases requiring imputation for the continuous value of vehicles (8-9%) and the lowest for real estate and other debt (less than 2%). On average, less than five percent of asset values are imputed in a given year (lower panel of Table 1). However, for the assessment of total net worth, the number of cases subject to imputation cumulates across wealth components: For about one quarter of cases, at least one wealth component that is part of total net worth is imputed. For about four percent of cases, three or more wealth components are imputed. Finally, in assessments of longitudinal changes (e.g., between two survey waves), the number of cases affected by imputation cumulates across years (see "Overall" column): only 62% of cases require no imputation of any wealth component in either year.

The described random hot-deck imputation was a state-of-the art method in the 1980s. In contrast, modern approaches incorporate covariates to increase the precision of the imputations, e.g., in a regression-switching framework, a technique that would have been all but impossible to implement back then given the limited computing power. The quality of imputed data is known to vary across different imputation approaches (Frick, Grabka, & Sierminska, 2007); the hot-deck imputation approach currently applied by PSID might be particularly prone to inflate estimates of wealth fluctuation, calling for the type of methodological assessment provided here.

3 Analytic Approach, Measures, Methods

3.1 Analytic Approach

We assess the relationship between large wealth fluctuation and potential sources of measurement error, including the number of imputed wealth components. However, we also investigate the extent to which actual changes in households' socio-economic circumstances predict large wealth fluctuation. It is necessary to pursue both aims at the same time. By jointly estimating the conditional role of imputation as a potential source of measurement error on the one hand and substantively meaningful changes on the other, we take into account that the two might be interrelated. For example, item nonresponse might be correlated with turbulences in a household's socio-economic conditions if a respondent is less likely to recall or disclose asset information if he recently lost his job and now consumes out of his family's assets.

It is important to note that our analyses cannot provide a strict comparative adjudication between the total "signal" and "noise" underlying large wealth fluctuation. Although our analyses include another potential source of measurement error, an indicator noting whether there was a change in respondent between waves, we cannot claim to exhaustively capture all possible "noise," nor, for that matter, all possible "signals." Instead, we reveal *some* of the predictors of large wealth fluctuation that likely indicate measurement error to motivate further improvements in data collection and processing. At the same time, we reveal substantively meaningful sources of changes in household wealth, which might – especially if they account for a significant share of large wealth fluctuation – caution against the premature conclusion that large wave-to-wave fluctuation in hard-to-measure economic variables is inherently problematic.

3.2 Sample and Measures

For this methodological project, we use PSID's imputed net worth variables that cumulate all measured asset and debt components (see Table 1) to examine net worth fluctuation between the 2005 and 2007 waves. We selected these two waves to circumvent strong period effects in subsequent waves brought about by the Great Recession in the form of substantial shocks to the wealth holdings of many American families (see Pfeffer, Danziger, & Schoeni, 2013). Our main analytic sample comprises 7,051 households with the same household head at both time points.[2]

Our outcome measures are based on the following six different specifications of wealth changes:

(a) absolute gains and absolute losses in net worth between 2005 and 2007, i.e., W_{2007}-W_{2005} ("absolute gain/loss");

(b) gains and losses in net worth between 2005 and 2007 relative to 2005 net worth among those with positive net worth in both years, i.e., $(W_{2007}-W_{2005})/W_{2005}$ ("relative gain/loss (to net worth)"); and

(c) gains and losses in net worth between 2005 and 2007 relative to 2005 household income among those with positive net worth in both years, i.e., $(W_{2007}-W_{2005})/I_{2005}$ ("relative gain/loss (to income)").

Though each of these measures has its advantages and disadvantages,[3] as we will show, they yield similar overall conclusions about the determinants of wealth fluctuation.

Determining the degree of wealth fluctuation that is large enough to raise suspicion about its sources is ultimately based on a subjective decision about what constitutes "too" extreme of a change. In this contribution, we define extreme gains and extreme losses as cases within the top five percent of the overall distribution of wealth gains and losses, respectively. Results based on just the top 2.5% yield similar results and are available upon request.

2 Drawing the analytic sample based on household heads observed in both waves is one common and necessary strategy to identify households across waves. It does, of course, condition on an important aspect of demographic changes in household structure (namely, the dissolution or formation of a household with a new household head) and, as such, provides a conservative estimate of the role of demographic changes in accounting for large wealth fluctuations.

3 For example, households with greater wealth should be more likely to experience large absolute changes (e.g., losing more than \$200,000) whereas households with lower wealth should be more likely to experience large changes relative to their baseline net worth (e.g., double their wealth by moving from \$100 to \$200 net worth). Additionally, the measure of change relative to baseline household income is also intended to address these distributional concerns (e.g., a wealth gain of \$10,000 for a household with an income of \$50,000 is treated the same as a wealth gain of \$50,000 for a household with an income of \$250,000).

Predictors of large wealth fluctuation (i.e., independent variables), include

(a) indicators of measurement characteristics, including the *number of imputed wealth components across both waves* (see Table 1, bottom panel) and whether there was a *change in respondent*,[4]

(b) an encompassing list of demographic characteristics (*age*, *sex*, and *race of household head* and *baseline wealth*) and changes in socio-economic circumstances between 2005 and 2007, including changes in *household composition*, *asset portfolios ("active savings")*, *labor market participation*, and *health conditions* (see Appendix A for a detailed list).

3.3 Methods

To analyze the determinants of large wealth fluctuation, we estimate logistic regression models for each of our six outcome variables (large gains and losses as absolute, relative to net worth, and relative to income changes). All of our analyses are weighted using the 2005 PSID family weight. All regression coefficient estimates are displayed as average marginal effects in Appendix A. For ease of presentation and interpretation, we display a selection of the main estimates in the form of predicted probabilities – more specifically, as discrete changes based on average marginal effects. We also briefly discuss model fit based on a pseudo-R2 for logistic regressions following McKelvey and Zavoina (1975), a measure that has been shown to best approximate the "percent explained variance" interpretation commonly used in OLS regressions (Hagle & Mitchell, 1992; Windmeijer, 1995).[5] All estimates are produced using the `margin` and `spost` commands in Stata 14 (Long & Freese, 2014).

---

4 The PSID does not necessarily interview the same respondent in both years, even in households with no composition change since the prior wave. For instance, a husband might be the respondent in one year whereas his wife might be the respondent in another year.

5 This interpretation requires us to assume a latent trait underlying our outcome variables (Long & Freese, 2014). Such an assumption seems justified in this application because we are more interested in evaluating the latent trait of "wealth fluctuation" than in evaluating the observed trait of specifically falling into the outlying gains/losses of the wealth change distribution. The fit statistics reported here are based on unweighted regressions.

Table 2 Distribution of Wealth Fluctuation

	Absolute Change	Relative Change (to baseline wealth)	Relative Change (to baseline income)
Percentile 1	-1,026,278	-0.97	-30.96
Percentile 5	-218,700	-0.80	-5.50
Median	7,800	0.20	0.33
Percentile 95	525,000	7.00	10.17
Percentile 99	1,885,300	53.55	47.26
N	7,051	5,329	5,323

4 Results

4.1 Distribution of Wealth Fluctuation

Table 2 displays the distributions of our main measures of between-wave wealth fluctuation: absolute change as well as relative change among those with positive net worth in both years (relative to baseline wealth and relative to baseline income). The median wealth change is $7,800 in absolute terms, 20% relative to net worth, and 33% relative to income. Inflation accounts for at least some of the increase in the first two measures; however, we do not adjust for inflation because we are more interested in the accuracy of respondents' reports than relating wealth to changing macro-economic conditions. The typical degree of wealth fluctuation reported here indicates that wealth tended to increase leading up to the crash, a finding consistent with prior research based on the same data (Pfeffer et al., 2013).

Our main interest here is in the tails of the distribution of wealth fluctuation. As shown in Table 2, the largest five percent of wealth losses and gains, which we designate as large fluctuation for the purpose of this contribution, are losses of $218,700 or more and gains of $525,000 or more, respectively. Large fluctuation relative to net worth includes losses of 80% or more and gains by a factor of 7 or more. Large wealth fluctuation relative to income includes cases experiencing a loss of wealth that is at least 5.5 times as high as their baseline income or a gain of wealth that is at least 10.2 times as high as their baseline income.

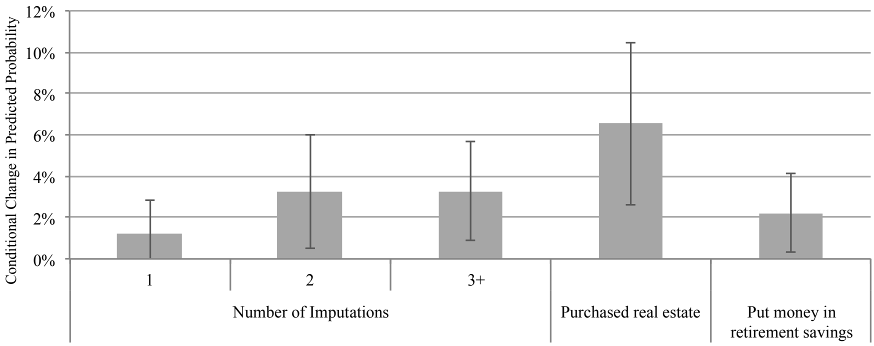
4.2 Predictors of Large Wealth Fluctuation

Table A.1 reports regression estimates for the prediction of large gains and large losses across the different specifications of wealth change (six separate regressions) and reveals the main categories of variables that independently and consistently predict large wealth changes (besides the expected association between baseline level of wealth and wealth changes; see footnote 4): (1) the number of imputations as an indicator of potential measurement error, (2) changes in asset portfolios, and (3) changes in household composition (though we remind the reader that our sample necessarily conditions on some fundamental changes in household composition; see footnote 3). Here, we report some of the main results in graphical form to facilitate interpretation. Specifically, we illustrate those predictors that are generally the largest and most consistent predictors of wealth gains or losses (see Figure 1a for predictors of large absolute gains and Figure 1b for predictors of large absolute losses). Although we only display associations with *absolute* gains and losses, graphs showing (in many cases even larger) associations with relative wealth changes are available upon request.

Figure 1a shows that, conditional on all other observed factors, the imputation of one wealth component (in either year) is associated with an increase in the probability of observing a large absolute wealth gain by 1.2 percentage points, though not statistically significant ($p>.05$). The imputation of two or more wealth components is associated with a statistically significant increase in the probability of observing large wealth gains by about 3 percentage points ($p<.05$). That is, we are 3.3 percentage points more likely to observe a large wealth gain for households with at least two imputed wealth components compared to otherwise similar households for whom we observe all wealth components. Because the baseline probability of experiencing a large wealth gain, as defined in this study, is 5 percentage points, an increase of 3 percentage points is substantial. We return to a substantive interpretation of these associations below.

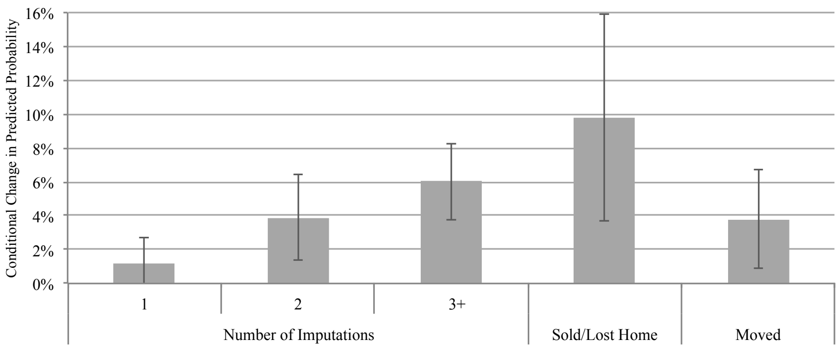
Figure 1a also displays results for two examples of substantively meaningful predictors of large wealth gains: purchasing real estate and saving for retirement. Specifically, everything else equal, the probability of observing a large absolute wealth gain is 6.6 percentage points greater for households that purchased real estate (other than their main residence) and 2.2 percentage points greater for those who put money into retirement savings (private annuities and Individual Retirement Accounts). We then observe that some substantive indicators – such as the purchase of real estate – are more predictive of large wealth gains than the imputation indicator chosen here.

Figure 1b displays the independent predictors of large absolute wealth *losses* and reveals quite similar conclusions. Specifically, everything else equal, the probability of observing large wealth losses is 3.9 percentage points and 6 percentage points greater among those with two imputed wealth components and those with three or more imputed wealth components, respectively, compared to those with no wealth imputations. Furthermore, large wealth losses are also associated with substantively meaningful changes in household characteristics, including the transition from home ownership to non-ownership (with an increase in probability of 9.8 per-



Note: Discrete change (based on logistic regression reported in A.1); 95% confidence intervals

Figure 1a Predictors of Large Positive Wealth Fluctuation (Absolute Gains)



Note: Discrete change (based on logistic regression reported in A.1); 95% confidence intervals

Figure 1b Predictors of Large Negative Wealth Fluctuation (Absolute Losses)

centage points) and, separately, the household moving to a different residence (an increase of 3.8 percentage points).

We judge all conditional associations shown here to be of considerable size. But how do we interpret them in substantive terms? We designated as "substantively meaningful predictors" the various aspects of active savings that are independently associated with large wealth fluctuation, including the purchase of real estate, putting money into retirement savings, and selling or losing a home (see Table A.1 for others, such as the purchase of stocks or home improvements). We believe that these indicators are likely to reflect true fluctuation in households' economic profiles: Some households experience both large wealth shifts and shifts in their wealth portfolio and investment behavior together. However, we do not claim that these factors exert a causal effect; in fact, for many of these factors, it is unclear whether they should be thought of as determinants of a large wealth change (e.g., selling a house in a bad market might trigger a substantial loss of net worth) or a consequence (e.g., the involuntary loss of a house, such as through foreclosure, might be caused by preceding socio-economic troubles and asset losses). Either way, we believe that changes in active savings and several other household characteristics listed in Table A.1 are sources of meaningful wealth fluctuation.

In contrast, we believe that the independent association between wealth fluctuation and the presence of imputations suggest that the imputation algorithm currently applied might be a source of measurement error underlying large wealth changes.[6] Having described the nature of the hot-deck imputation algorithm above, this interpretation seems quite probable to us. Of course, theoretically, the imputation indicator might also be a proxy for selective nonresponse. That is, even with the ample list of observable control variables included here, it is possible that reports on wealth components might not be missing at random (MAR). However, the structure of selective nonresponse would have to be quite peculiar to produce the patterns observed here: similarity in the associations between the imputation indicator and large wealth gains and wealth losses as well as the monotonic increase in the probability of large fluctuation across the number of imputed components.

4.3 Accounting for Large Wealth Fluctuation

In a final step, we evaluate whether the observed household characteristics and potential measurement artifacts studied here account for an appreciable share of the variability in wealth fluctuation. This assessment is based on the estimated pseudo R2 reported in the bottom panel of Table A.1. Across all outcomes, our full models account for a substantial share of the variability in wealth fluctuation and, for half of the models, the majority of the variability (row 1). Indicators of demographic and changes in socio-economic characteristics alone (row 2) explain between one quarter and one half of the variance in wealth gains (38% of absolute gains, 48 % of gains relative to wealth, and 27% of gains relative to income) and up to four fifths of the variance in wealth losses (80% of absolute losses, 31% of losses relative to wealth, and 50% of losses relative of income). As a single predictor, the number of wealth components imputed (row 3) explains up to 11% of the variance whereas a change in respondent explains far less (row 4). However, conditional on the predictors of meaningful wealth fluctuation (row 1), the contribution of measurement error indicators is quite modest (compare rows 1 and 2): The additional variance

6 We also note that our other tested indicator of measurement error, a change in respondent, is a less consistent predictor of wealth fluctuation. Specifically, a change in respondent independently predicts extreme changes in relative gains but not other specifications of change.

explained by taking into account indicators of potential measurement error is less than 5% for all models and far less in most (about 1%).

5 Conclusion

We have studied between-wave changes in family net worth as an increasingly important indicator of economic wellbeing that is also particularly hard to measure. Using PSID data from 2005 and 2007, we sought to differentiate between substantively meaningful predictors of wealth fluctuation (specifically, changing socioeconomic and demographic conditions of households) and potential measurement error arising from wealth imputations and a change in respondent.

Deciding what degree of wealth change is large enough to qualify as suspicious is arbitrary; here, we focused on the five percent of households that experienced the largest absolute and relative gains and the five percent that experienced the largest losses. Using this definition, we were able to account for between 31% and 80% of large wealth losses (depending on whether measuring absolute or relative losses) and between 29% and 52% of large gains based only on households' demographic and socio-economic characteristics and changes therein. In other words, the mere fact that a household's wealth in one wave is radically different from its wealth in the prior wave should not automatically trigger concerns about the presence of measurement error. Instead, the best explanations for such extreme fluctuation (other than the household's baseline level of wealth) are changes in asset portfolios. For example, a change in home ownership is highly predictive of experiencing large wealth fluctuation as are other asset portfolio changes, such as the purchase of real estate or investments in businesses.

However, we have also shown in detail that, whereas the imputation strategy currently implemented by PSID contributes only a small additional portion to the overall explained variance in wealth fluctuation, having more imputed wealth components is clearly and independently associated with large wealth fluctuation. This finding suggests that the random hot-deck imputations that were the state-of-the-art approach when the PSID began collecting wealth data in the 1980s could be updated to accommodate covariates, including information from prior and subsequent waves (Moldoff et al., 2013; Westermeier & Grabka, 2015). In particular, including the changes in life circumstances identified here (e.g., changes in home ownership and active savings behaviors) appears to be a promising next step in improving the wealth data provided by PSID and perhaps other surveys.

Generally speaking, multivariate multiple imputation methods have been demonstrated to be superior to univariate single imputation methods. For example, in an evaluation of methods for imputing bracketed survey data on household wealth in the Health and Retirement Study, Heeringa, Little, and Raghunathan,

(2002) found that a Bayesian approach to multiple imputation was more effective than complete-case analysis, mean or median substitution, and multiple imputation based on a univariate hot deck (see Heeringa, 1999 for earlier simulation work demonstrating the utility of the method). More recent research directly addresses the effectiveness of incorporating longitudinal information in the imputation of panel data, considering the effects of imputation on both cross-sectional accuracy (e.g., trends, distributions, and measures of inequality) and longitudinal accuracy (e.g., distributional accuracy of wealth mobility). Although Frick & Grabka (2007) found that imputations incorporating longitudinal information were superior to those that did not, Kennickell (2011) found no meaningful differences between different methods and Westermeier & Grabka (2015) found that no single method was best for all scenarios. To that end, future methodological work should explore the effectiveness of a variety of these latest imputation techniques given the particulars of PSID. In the meantime, analysts are able to utilize imputation flags provided by the PSID to re-impute wealth information themselves and, in the process, ensure that their imputation models mimic their specific analytic models (Allison, 2002; Rubin, 1987).

References

Allison, P. D. (2002). *Missing data* (Vol. 136). Thousand Oaks, CA: Sage Publications, Inc.

Bosworth, B. P. & Smart, R. (2009). *Evaluating micro-survey estimates of wealth and saving* (Working Paper 2009-4). Center for Retirement Research at Boston College, Boston.

Bucks, B., & Pence, K. (2015). Wealth, pensions, debt, and savings: Considerations for a panel survey. *Journal of Economic and Social Measurement, 40,* 151-175.

Conley, D. (2001). Capital for college: Parental assets and postsecondary schooling. *Sociology of Education, 74*(1), 59-72.

Deaton, A. (1991). Saving and liquidity constraints. *Econometrica,59,* 1221-48.

Duncan, G. J. (1984). *Years of poverty, years of plenty: The changing economic fortunes of American workers and families.* Ann Arbor, MI: Institute for Social Research, University of Michigan.

Frick, J. R., & Grabka, M. M. (2007). *Item non-response and imputation of annual labor income in panel surveys from a cross-national perspective* (Discussion Paper 763). DIW Berlin, Berlin.

Frick, J. R., Grabka, M. M., & Sierminska, E. M. (2007). *Representative wealth data for Germany from the German SOEP: The impact of methodological decisions around imputation and the choice of the aggregation unit* (Discussion Paper 672). DIW Berlin, Berlin.

Groves, R. M. (2004). *Survey errors and survey costs*. Hoboken, NJ: Wiley-Interscience.

Hagle, T. M., & Mitchell II, G. E. (1992). Goodness-of-fit measures for probit and logit. *American Journal of Political Science, 36*, 762-784.

Heeringa, S. G. (1999). *Multivariate imputation and estimation for coarsened survey data on income and wealth* (Doctoral dissertation). University of Michigan, Ann Arbor, MI.

Heeringa, S. G., Little, R. J. A., & Raghunathan, T. E. (2002). Multivariate imputation of coarsened survey data on household wealth. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 357-371). New York: John Wiley & Sons.

Hill, D. H. (2006). Wealth dynamics: Reducing noise in panel data. *Journal of Applied Econometrics, 21*, 845-860.

Juster, F. T., & Smith, J. P. (1997). Improving the quality of economic data. Lessons from HRS and AHEAD. *Journal of the American Statistical Association, 92*, 1268-1278.

Kennickell, A. B. (2011). Look again: Editing and imputation of SCF panel data. In *JSM Proceedings*, Survey Research Methods Section. Alexandria, VA: American Statistical Association.

Long, S., & Freese, J. (2014). *Regression models for categorical dependent variables Using Stata*. College Station: Stata Press.

McKelvey, R. D., & Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology, 4,* 103-120.

Moldoff, M., Meijer, E., Chien, S., Hayden, O., Hurd, M., Main, R., Miu, A., Rohwedder, S., & St. Clair, P. (2013). *RAND HRS income and wealth imputation, Version M*. Rand Center for the Study of Aging, Santa Monica, CA.

Panel Study of Income Dynamics (2013). PSID main interview user manual: Release 2013. Ann Arbor, MI: University of Michigan.

Panel Study of Income Dynamics (2015). Public use dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI.

Piketty, T. (2014). *Capital in the twenty-first century*. Cambridge, MA: Belknap Press.

Pfeffer, F. T., & Schoeni, R. F. (eds.) 2016. Wealth inequality: Economic and social dimensions. *Russell Sage Journal of the Social Sciences, 6(2).*

Pfeffer, F. T., Danziger, S. H., & Schoeni, R. F. (2013). Wealth disparities before and after the great recession. *Annals of the American Academy of Political and Social Science, 650*(1), 98-123.

Rubin, D. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

Venti, S. F. (2011). Economic measurement in the Health and Retirement Study. *Forum for Health Economics & Policy, 14* (3), Article 2.

Westermeier, C., & Grabka, M. M. (2015). *Longitudinal wealth data and multiple imputation: An evaluation study* (SOEPpapers on Multidisciplinary Panel Data Research 790). DIW Berlin, Berlin.

Windmeijer, F. A.G. (1995). Goodness-of-fit measures in binary choice models. *Econometric*

APPENDIX A

Predictors of Large Wealth Fluctuation

Table A.1 Logistic Regressions

Average marginal effects (standard errors in parentheses)

	Gains (top 5%)			Losses (top 5%)		
	Absolute	Relative (to net worth)	Relative (to income)	Absolute	Relative (to net worth)	Relative (to income)
Wealth Imputation						
Number of asset components imputed (reference = none)						
One	0.012	0.020 *	0.005	0.012	0.001	0.013
	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)	(0.010)
Two	0.033 *	0.044 **	0.04 *	0.039 **	0.009	0.045 **
	(0.014)	(0.014)	(0.016)	(0.013)	(0.014)	(0.016)
Three or more	0.033 **	0.078 ***	0.067 ***	0.060 ***	0.009	0.071 ***
	(0.012)	(0.019)	(0.017)	(0.012)	(0.013)	(0.015)
Change in Respondent						
Whether respondent changed between 2005 and 2007	0.022	0.026 *	0.045 ***	-0.009	0.012	-0.012
	(0.014)	(0.010)	(0.013)	(0.017)	(0.015)	(0.020)
Demographics						
Age of HH head	0.000	0.000	0.001 ***	0.000	-0.001 *	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Sex of HH head (reference = female)	0.024 *	0.005	-0.005	0.009	-0.017	-0.031 **
	(0.011)	(0.008)	(0.011)	(0.010)	(0.009)	(0.010)

	Gains (top 5%)			Losses (top 5%)		
	Absolute	Relative (to net worth)	Relative (to income)	Absolute	Relative (to net worth)	Relative (to income)
Race of HH head (reference = white)						
Black	-0.003 (0.012)	0.003 (0.008)	0.022 (0.016)	0.024 (0.018)	0.015 (0.011)	0.030 (0.019)
Other	-0.013 (0.012)	0.016 (0.010)	0.015 (0.016)	-0.007 (0.012)	0.033 * (0.016)	0.005 (0.017)
Baseline wealth quintile						
1st	reference	reference	reference		reference	
2nd	0.004 (0.008)	-0.436 *** (0.053)	-0.044 (0.038)	reference	0.016 (0.012)	reference
3rd	0.002 (0.006)	-0.593 *** (0.055)	-0.069 (0.037)		0.018 (0.013)	0.008 * (0.004)
4th	0.019 * (0.008)	-0.621 *** (0.056)	-0.045 (0.037)	0.011 *** (0.003)	0.025 (0.014)	0.025 *** (0.007)
5th	0.081 *** (0.011)	-0.627 *** (0.056)	-0.022 (0.038)	0.216 *** (0.020)	0.046 ** (0.016)	0.145 *** (0.016)
Change in household composition						
Any change in household structure	0.002 (0.008)	0.007 (0.007)	-0.019 (0.011)	0.000 (0.009)	0.012 (0.008)	-0.013 (0.011)
Change in sum of assets and/or debts due to mover-in or mover-out (reference = else)						
Net gain	0.050 (0.026)	0.021 (0.018)	0.061 (0.032)			
Net loss				0.021 (0.018)	0.040 * (0.017)	0.016 (0.024)

	Gains (top 5%)			Losses (top 5%)		
	Absolute	Relative (to net worth)	Relative (to income)	Absolute	Relative (to net worth)	Relative (to income)
Whether household member entered college	0.027 *	0.009	0.011	-0.003	-0.007	-0.014
	(0.011)	(0.014)	(0.016)	(0.014)	(0.014)	(0.018)
Whether family moved	0.000	-0.002	0.012	0.030 **	0.033 **	0.037 **
	(0.012)	(0.007)	(0.013)	(0.010)	(0.011)	(0.013)
Changes in asset portfolio						
Whether change in home ownership status (reference = else)						
Now owns home	0.023	0.056 ***	0.057 ***			
	(0.018)	(0.008)	(0.016)			
No longer owns home				0.065 ***	0.094 ***	0.064 ***
				(0.016)	(0.013)	(0.018)
Whether sold home used as main dwelling	-0.034 *	0.021	-0.027	-0.040 **	-0.028	-0.051 *
	(0.017)	(0.011)	(0.021)	(0.015)	(0.016)	(0.022)
Whether purchased real estate other than main home	0.045 ***	0.037 *	0.023	0.009	0.008	-0.009
	(0.010)	(0.017)	(0.015)	(0.011)	(0.017)	(0.016)
Whether sold real estate other than main home	-0.004	-0.009	0.000	0.01	0.016	0.024
	(0.015)	(0.027)	(0.019)	(0.015)	(0.027)	(0.022)
Whether made home additions or improvements	0.026 ***	0.015	0.019	-0.013	-0.057 **	-0.015
	(0.007)	(0.011)	(0.011)	(0.008)	(0.019)	(0.011)
Whether purchased non-IRA stock	0.021 *	-0.003	0.007	-0.014	-0.006	-0.018
	(0.009)	(0.012)	(0.013)	(0.009)	(0.014)	(0.011)
Whether sold non-IRA stock	0.009	0.014	-0.002	0.008	-0.015	0.003
	(0.009)	(0.014)	(0.015)	(0.010)	(0.018)	(0.012)
Whether put money into private annuities or IRAs	0.019 **	0.027 **	0.010	-0.019 **	-0.059 **	-0.022 *
	(0.007)	(0.009)	(0.010)	(0.007)	(0.018)	(0.010)

	Gains (top 5%)			Losses (top 5%)		
	Absolute	Relative (to net worth)	Relative (to income)	Absolute	Relative (to net worth)	Relative (to income)
Whether cashed in any part of pension, private annuity, or IRA	0.017 (0.010)	-0.001 (0.015)	0.010 (0.012)	-0.016 (0.012)	-0.017 (0.017)	-0.025 (0.014)
Whether invested in business or farm	0.020 * (0.010)	0.029 ** (0.010)	0.023 (0.014)	0.010 (0.010)	-0.003 (0.015)	0.013 (0.013)
Whether sold business or farm	-0.020 (0.022)	0.008 (0.026)	-0.016 (0.030)	0.013 (0.025)	0.009 (0.035)	-0.013 (0.031)
Whether received gift or inheritance >=$10k in last two years	-0.002 (0.012)	0.040 ** (0.014)	0.014 (0.020)	0.010 (0.012)	-0.067 ** (0.025)	0.012 (0.014)
Whether received large settlement or inheritance in last year	0.009 (0.012)	0.013 (0.014)	0.010 (0.018)	-0.011 (0.012)	-0.005 (0.016)	-0.005 (0.014)
Changes in labor market participation						
Change in head employment status (reference = else)						
Employed to unemployed				0.017 (0.014)	0.012 (0.014)	0.015 (0.016)
Unemployed to employed	0.006 (0.016)	-0.016 (0.016)	0.026 (0.019)			
Change in spouse employment status (reference = else)						
Employed to unemployed				0.016 (0.013)	0.000 (0.014)	0.019 (0.016)
Unemployed to employed	-0.028 (0.022)	0.018 (0.013)	-0.019 (0.024)			

	Gains (top 5%)			Losses (top 5%)		
	Absolute	Relative (to net worth)	Relative (to income)	Absolute	Relative (to net worth)	Relative (to income)
Whether change in head occupation	0.010 (0.007)	0.000 (0.006)	0.016 (0.009)	-0.002 (0.007)	-0.004 (0.007)	-0.009 (0.008)
Whether change in spouse occupation	0.000 (0.007)	-0.001 (0.007)	-0.009 (0.010)	-0.013 * (0.007)	-0.013 (0.008)	-0.020 * (0.009)
Change in head retirement status (reference = else)						
Retired to not retired	0.018 (0.024)	-0.070 * (0.030)	0.022 (0.022)			
Not retired to retired				-0.008 (0.015)	-0.009 (0.022)	0.007 (0.017)
Change in spouse retirement status (reference = no else)						
Retired to not retired	-0.054 (0.032)	-0.007 (0.035)	-0.024 (0.027)			
Not retired to retired				-0.025 (0.018)	0.038 (0.022)	-0.021 (0.023)
Changes in Health Status						
Change in head health status (reference = else)						
Worse				0.005 (0.008)	-0.003 (0.009)	0.006 (0.009)
Better	0.001 (0.008)	-0.007 (0.008)	0.000 (0.010)			
Change in spouse health status (reference = else)						
Worse				0.009 (0.009)	-0.005 (0.012)	-0.018 (0.013)
Better	-0.001 (0.010)	-0.009 (0.009)	-0.017 (0.014)			

	Gains (top 5%)			Losses (top 5%)		
	Absolute	Relative (to net worth)	Relative (to income)	Absolute	Relative (to net worth)	Relative (to income)
Change in head perceived health status (reference = else)						
Worse				-0.016	0.005	-0.004
				(0.009)	(0.010)	(0.010)
Better	-0.016	-0.004	-0.030			
	(0.011)	(0.008)	(0.017)			
Change in spouse perceived health status (reference = else)						
Worse				-0.026 *	-0.020	-0.015
				(0.012)	(0.015)	(0.017)
Better	-0.001	-0.015	0.001			
	(0.013)	(0.011)	(0.018)			
Change in health condition limiting work for head (reference = else)						
Worse				0.015	0.005	0.012
				(0.011)	(0.012)	(0.013)
Better	-0.037	0.022	-0.015			
	(0.019)	(0.014)	(0.018)			
Change in health condition limiting work for spouse (reference = else)						
Worse				0.016	-0.037	0.019
				(0.015)	(0.019)	(0.018)
Better	-0.007	-0.034	-0.013			
	(0.020)	(0.022)	(0.023)			

	Gains (top 5%)			Losses (top 5%)		
	Absolute	Relative (to net worth)	Relative (to income)	Absolute	Relative (to net worth)	Relative (to income)
Pseudo R2 (McKelvey & Zavoina)						
(1) All predictors included	0.397	0.520	0.293	0.804	0.309	0.516
(2) Only socio-econ. & demogr. changes	0.382	0.475	0.267	0.803	0.306	0.501
(3) Only imputation indicator	0.040	0.017	0.073	0.076	0.003	0.109
(4) Only change in respondent	0.005	0.005	0.017	0.000	0.003	0.001
N	6,594	4,994	4,988	6,594	4,994	4,988

Note: Statistical significance levels at * $p<.05$, ** $p<.01$, and *** $p<.001$ based on two-tailed tests. For the prediction of large absolute and relative (to income) losses, some bottom wealth quintiles had to be merged due to perfect prediction. The specification of predictors capturing directional change (e.g. health conditions worsening vs. improving) is targeted to the outcome measure (e.g. for wealth losses: health conditions worsening vs. not; for wealth gains: health conditions improving vs. not).