

# Do Homeowners Know Their House Values and Mortgage Terms?

Brian Bucks and Karen Pence  
Federal Reserve Board of Governors

January 2006

## Abstract

To assess whether homeowners know their house values and mortgage terms, we compare the distributions of these variables in the household-reported 2001 Survey of Consumer Finances (SCF) to the distributions in lender-reported data. We also examine the share of SCF respondents who report not knowing these variables. We find that most homeowners appear to report their house values and broad mortgage terms reasonably accurately. Some adjustable-rate mortgage borrowers, though, and especially those with below-median income, appear to underestimate or not know how much their interest rates could change.

The views expressed in this paper are ours alone and not necessarily those of the Board of Governors or its staff. We thank Carolyn Aler for excellent research assistance and many generous Federal Reserve colleagues, Michael Carliner, Bill Gale, Markus Grabka, Jim Lacko, David Newhouse, Anthony Pennington-Cross, Howard Savage, Scott Susin, and participants at the AREUEA Mid-Year Conference, the Washington Statistical Society, the Society of Government Economists, and the Luxembourg Wealth Study Conference on “Construction and Usage of Comparable Microdata on Wealth” for helpful discussions and suggestions. Contact information: [brian.k.bucks@frb.gov](mailto:brian.k.bucks@frb.gov), [karen.pence@frb.gov](mailto:karen.pence@frb.gov).

## Introduction

Homeowners' understanding of their house values and mortgage terms is of interest to both researchers and policymakers. From a research perspective, many studies of housing wealth and mortgage debt are based on household-reported data. If households are uncertain of this information and systematically misreport it on surveys, the results from these studies could be misleading. From a policy perspective, if borrowers do not know or misestimate their house values, they may make consumption and saving decisions that turn out to have been inappropriate and that require adjustments in these decisions at a later date. If borrowers do not know their mortgage terms, they may be surprised by the change in their payments if interest rates rise and thus may subsequently experience financial difficulties. This last question has taken on particular policy importance in recent years with the rise in both short-term interest rates and the proportion of homeowners with adjustable-rate mortgages.<sup>1</sup>

To examine homeowners' awareness of their house values and mortgage terms, we compare the estimated rates of house price appreciation and the distributions of mortgage terms as reported by homeowners in the Survey of Consumer Finances (SCF) to the distributions of the same variables as reported by lenders in three data sources. These sources are the Office of Federal Housing Enterprise Oversight (OFHEO) house price index, which is based on mortgages held or guaranteed by Fannie Mae or Freddie Mac; the Residential Finance Survey (RFS), a survey of homeowners and lenders conducted in conjunction with the decennial Census; and data compiled by the LoanPerformance (LP) Corporation from the administrative records of large mortgage servicers. In addition, we examine the shares of SCF respondents who replied "don't know" when asked about their house values or mortgage terms. Although other researchers have studied the accuracy

---

<sup>1</sup> Data from LoanPerformance Corporation indicate that share of prime mortgages that are adjustable-rate rose from 8 percent in December 2001 to 12 percent in October 2005. Over the same period, the share of subprime mortgages that are adjustable-rate rose from 36 percent to 47 percent.

of homeowner-reported house values, our study is among the first to examine the accuracy of borrower-reported mortgage terms.

Our results suggest that household-reported surveys, and the Survey of Consumer Finances in particular, capture broad measures of housing wealth and mortgage terms reasonably accurately. Our index of house value appreciation based on household-reported data matches the aggregate OFHEO index fairly closely. In addition, almost all homeowners are able to provide a dollar amount or range when asked about their house values. The SCF distributions of mortgage maturities, types, and payments also match lender-reported distributions well. These comparisons do not necessarily indicate whether any given household in the SCF reports these values correctly, as the errors of individual households could offset each other in such a way that the distribution remains accurate. However, summary statistics from the household-reported data appear valid, and researchers studying questions such as the effect of housing wealth on consumption may be on safe ground using household-reported data.

Household-reported data do not appear, however, to depict the terms of adjustable-rate mortgages (ARMs) with the same degree of accuracy, as the borrower-reported distributions of ARM terms are quite different from the lender-reported distributions. In particular, borrowers appear to underestimate the amount by which their interest rates can change. These differences may stem partly from differences in the design and sample composition of the surveys; because adjustable-rate mortgages are complex contracts, small differences in question wording may affect how respondents interpret and answer questions. However, borrower confusion also appears to be a factor, as a sizable number of adjustable-rate borrowers report that they do not know the terms of their contracts. These results suggest that borrower-reported data remain the best choice for researchers interested in household perceptions of their mortgage terms, but that lender-reported data may be a better option for actual terms on adjustable-rate mortgages.

To address the policy implications of this borrower confusion in more detail, we examine which types of households are more likely to know the terms of their ARMs and how vulnerable these households are to an increase in interest rates. We first document that households with low income and less education are less likely to know their mortgage terms. We then simulate how household ARM payments might change if interest rates rose two percentage points for two consecutive years, for a cumulative four percentage point increase, assuming that SCF estimates represent the expected payment changes and RFS estimates represent the actual payment changes. In both datasets, we find that the broad majority of ARM borrowers might experience changes in payments of less than 5 percent of gross income under the terms of this simulation. However, lower-income households might be most likely to experience larger changes, and about 10 percent of borrowers might be surprised that their changes in payment exceed 5 percent of their income.

### **Previous studies**

Previous validation studies of household-reported housing data have used one of three approaches. The first approach, which is followed in this paper, compares homeowner estimates in the aggregate to external indexes. The second approach compares individual homeowner estimates to lender or researcher estimates of the true values. The third approach uses panel data to examine whether homeowners describe their homes and mortgages consistently over time.

An example of the first approach is DiPasquale and Somerville (1995), who use the homeowner-reported house values and transaction sales prices in the American Housing Survey (AHS) to construct aggregate measures of the changes in house prices over time. Although the homeowner-estimated house values are somewhat higher than the AHS transaction prices, the two house price series track each other and a National Association of Realtors house price series fairly closely.

Agarwal (2005), Kiel and Zabel (1999), and Goodman and Ittner (1992) follow the second approach and compare homeowner house price estimates with lender or researcher benchmark estimates derived from past transaction sales prices and house price indexes. These studies tend to find that the discrepancy between the homeowner and lender estimates is fairly small: homeowner estimates are generally 3 to 6 percent higher on average than the benchmark estimates, with an average absolute difference around 14 percent.<sup>2</sup> In addition, the discrepancy may reflect other factors than homeowner error. In particular, the benchmark estimate for each home is likely also measured with error.

Kennickell and Starr-McCluer (1997), an example of the third approach, is the only previous study to examine how accurately respondents report housing data in the SCF. They exploit the unique structure of the 1983–89 SCF panel, which contains cross-sectional information on household portfolios in 1983 and 1989 as well as retrospective questions (asked in 1989) about changes in household portfolios over the 1983–89 period. They find that only 5 percent of households in the panel provided retrospective data about home sales and purchases that was inconsistent with the cross-sectional data. They attribute this high consistency rate to the fact that home purchases and sales are “well-defined, highly salient events.” Likewise, in a study based on Dutch panel data, Alessie and Zandvliet (1993) find that housing is one of the better-measured components of wealth.

The literature on the accuracy of respondent-reported mortgage data is much more limited. In a comparison of homeowner- and lender-reported data on the 1970 Residential Finance Survey that follows the second approach, Fronczek and Koons (1976) found that most homeowners reported their mortgage payment amount accurately. However, in an application of the third approach,

---

<sup>2</sup> See the literature reviews in Agarwal (2005), Kiel and Zabel (1999), and Goodman and Ittner (1992) for summaries of earlier studies.

Leary, Newhouse, and Mihaly (2004) found that 46 percent of mortgages in the 1996–99 Survey of Income and Program Participation data had at least one mortgage term reported inconsistently over time. Thus, whereas our study of the accuracy of house values estimates builds upon a substantial literature, our study of the accuracy of mortgage terms explores questions that have received relatively little prior attention.

## **Data**

Our empirical work is based on four datasets:

The Federal Reserve’s Survey of Consumer Finances is the most comprehensive and highest quality dataset available on U.S. household wealth. The survey has been conducted every three years since 1983, with a consistent survey design since 1989. The survey design features both a standard, geographically based random sample and an over-sample of households likely to be relatively wealthy. These households are over-represented in the data in order to improve the accuracy of estimates of the types and amount of wealth concentrated among wealthy families. We use the SCF-provided nonresponse-adjusted analysis weights to make the estimates representative of the overall U.S. household population.

Data on the survey are reported by households, and missing data are imputed using multiple imputation techniques.<sup>3</sup> In 2001, the survey included 4,442 households, of which 3,162 were homeowners, 1,562 had fixed-rate mortgages, and 238 had adjustable-rate mortgages.<sup>4</sup> Aizcorbe, Kennickell, and Moore (2003) provide an overview of the 2001 data.

The Residential Finance Survey is conducted every ten years by the U.S. Census Bureau. The survey is designed to be representative of all non-farm residential properties in the United States. It included data on 16,929 properties in 2001. Households selected for the RFS sample are required by

---

<sup>3</sup> See Kennickell (1991, 1998) for more information on multiple imputation in the SCF.

<sup>4</sup> The data also include 77 balloon mortgages.

law to participate, unlike the SCF, where participation is voluntary. As might be expected given this difference in legal status, response rates are higher in the RFS than the SCF (86 percent vs. 68 percent).<sup>5</sup> U.S. Census (2005) describes the 2001 RFS data in greater detail.

The RFS collects general information on the property and the mortgage from the homeowner and detailed information on the mortgage from the lender. The lender-reported data are missing for roughly half of all mortgages. These data could be missing because the borrower did not provide information about the mortgage lender or because the mortgage had been sold and the RFS staff could not find the current servicer. In other cases, the RFS was able to find the current servicer, but the servicer did not have access to the original loan documents and thus could not report all variables. Thus loans that the originating lender did not sell—that is, kept in portfolio—are likely to be over-represented in the RFS data. The RFS does not impute missing data for most lender-reported variables, although it does impute missing data for some household-reported variables. Our tabulations of RFS variables exclude observations with missing values.

The Office of Federal Housing Enterprise Oversight house price index is a repeat-transactions house price index based on mortgages backed by single-family properties that have been held or guaranteed since 1975 by Fannie Mae and Freddie Mac.<sup>6</sup> The index is based only on conforming mortgages, which are those small enough to qualify for purchase by Fannie Mae or Freddie Mac (under \$275,000 in 2001); it also excludes government-backed Federal Housing Administration or Veterans' Administration mortgages. Our tabulations on the Residential Finance Survey indicate that about half of owner-occupied properties are captured by the OFHEO index. Appraisals from

---

<sup>5</sup> The RFS response rate is taken from U.S. Census (2005), Table 25, p. D-16. The SCF response rate is for the geographically based random sample only and is taken from Kennickell (2003) p.4.

<sup>6</sup> The OFHEO index is updated and released quarterly. The estimates in this paper are from the 2005 second quarter release and use the index value in the fourth quarter of 2001 as the baseline in computing appreciation rates. We use the RFS data released October 5, 2005 and the SCF data released December 12, 2003.

home refinancings tend to distort the index over short time periods but not over the longer periods examined in this paper.

The data from LoanPerformance Corporation consist of information collected from the administrative records of large mortgage servicers. The mortgages are originated by a wide variety of institutions and include both prime and subprime loans. All mortgages guaranteed by Fannie Mae or Freddie Mac are represented in the data. In total, the December 2001 data covered about 80 percent of U.S. home mortgages. Because LoanPerformance does not release the loan-level microdata underlying this product, the numbers reported here are based on aggregated tabulations provided by the company. All numbers shown are a weighted average of the estimates from the prime and subprime databases, with a weight of 0.88 given to the prime estimates and 0.12 to the subprime estimates.

## **Empirical Framework**

***Comparison of distributions.*** In the analyses of both house values and mortgage terms, our first step is comparing the lender- and homeowner-reported distributions of these data. In making these comparisons, we assume that the lender data represent the distribution of the *actual* mortgage characteristics and house values, and the homeowner-reported data represent the distribution of homeowner *perceptions* of these variables. We assume that the lender data are more likely to be accurate because they are drawn from administrative records.

Our comparison of distributions is a weaker test of reporting accuracy than a test that examines whether any given homeowner reports her house value and mortgage terms accurately. For example, if homeowners make offsetting errors, the distributions may match even though individual homeowners have reported data erroneously. However, if the distributions do not match, this discrepancy provides evidence of borrower uncertainty or misperception.



Of course, reasons other than borrower misperception might cause the distributions not to match. One such reason is that the surveys sample somewhat different slices of the mortgage universe. For example, by virtue of its sample design, the Survey of Consumer Finances may have more comprehensive coverage than the Residential Finance Survey of the homes and mortgages of high-wealth households. The RFS data are less likely to include mortgages that have been sold or securitized, whereas securitized mortgages are likely over-represented in the LoanPerformance data. The OFHEO index includes only mortgages purchased by Fannie Mae or Freddie Mac.

Another reason the distributions might diverge is because the surveys treat missing data differently. The SCF imputes five values for each piece of missing data, thereby allowing users to estimate the uncertainty associated with the imputation.<sup>7</sup> The RFS imputes one value for most household-reported missing data. However, the RFS does not impute values for most lender-reported variables. Consequently, we exclude from our tabulations of RFS lender-reported variables the 50 to 60 percent of observations with missing values for these variables. As a result, the lender-reported RFS data may have a greater selection problem than the SCF.<sup>8</sup>

Although both of these selection issues are potentially a problem, the fact that the RFS and LoanPerformance data match closely for many variables provides some reassurance that these issues are not severe. To address these concerns further, we note throughout our discussion where sample selection issues and differences in survey design are potential problems.

For the variables for which we have imputed values, we include them, unless otherwise noted, in our RFS and SCF estimates because the imputations may mitigate potential sample selection bias due to non-response. The imputations also provide the professional survey staff's best sense of the

---

<sup>7</sup> All counts shown in the paper are averages across the five implicates.

<sup>8</sup> More specifically, by dropping missing values, we assume that the probability of non-response does not depend on the characteristics of the household, lender, or mortgage, and so statistics based on non-missing data are unbiased estimates. Similarly, when we include imputed values, we assume that the imputation model and covariates capture systematic differences in the probability of non-response, and so the imputed values are unbiased estimates of the missing responses.

true values of these variables. Our interpretation assumes that imputations do not introduce systematic error into estimates of distributions. In the case of borrower-reported data, for instance, we assume that the distribution of a given variable, including imputed values, matches the distribution that would have been obtained if respondents who did not answer a question instead provided their best estimates.<sup>9</sup>

***Tabulation of “don’t know” responses.*** As a complement to these comparisons of distributions, we also examine the share of SCF responses that are edited or imputed, paying particular attention to the share of “don’t know” responses. These tabulations are based on the detailed edit or imputation flags provided for nearly all SCF variables. These flags indicate whether a response has been edited or imputed, as well as the reason why an imputed value was originally missing. These codes also indicate when a respondent gave a range for a dollar value, such as a home’s original purchase price.<sup>10</sup> The response rates are calculated over all households for whom the question was applicable; only the 66 percent of SCF respondents who are homeowners, for example, were asked the value of their house.

Responses are divided among six classifications:

- Original value: the original value provided by the respondent was deemed acceptable and included in the survey.
- Range: the respondent provided a range rather than an actual value, or auxiliary information used in the editing process was sufficient to bound the value. This option is available only for questions for which the response is a dollar value.

---

<sup>9</sup> This is a weaker assumption than assuming that the imputed values for a given respondent are unbiased.

<sup>10</sup> See Kennickell (1997) for further description and analysis of range responses in the SCF.

- Edited value: the respondent provided enough information in auxiliary or related fields that the correct value could be inferred by the SCF staff with a high degree of confidence.
- Missing due to editing: either a highly implausible value was set to a missing value or insufficient information was available for a given variable when the appropriate question sequence was altered in editing.
- Don't know: the respondent did not know the answer to the question. This category includes both households who did not know the answer to a particular question and households who were not asked the question because they did not know the answer to a question earlier in the sequence. Households who did not know whether their mortgages were fixed or adjustable, for example, were not asked the index to which their mortgages were tied.
- Refused: the respondent refused to answer the question. As with the "don't know" responses, this category includes both households who refused to answer a question and households who were not asked a question because they refused to answer an earlier question.

We assume that most "don't know" responses reflect genuine uncertainty on the part of respondents. Respondents may, of course, have other reasons for responding "don't know," such as privacy concerns or impatience with interview length. However, we find that the share of "don't know" rates differs dramatically across questions. As long as the share of respondents who report "don't know" for reasons other than genuine uncertainty is relatively constant across questions, higher "don't know" rates for some questions may indicate differences in respondent uncertainty.<sup>11</sup>

---

<sup>11</sup> This assumption that the share of "don't knows" generated by reasons other than genuine ignorance is comparable across questions may be particularly plausible for our variables because they are on similar topics and are asked together

The “don’t know” tabulations also help us interpret the comparisons of data distributions. Questions with high rates of “don’t know” responses may also include more guesses when respondents did answer the question, and these guesses may be less accurate. When we observe high rates of “don’t know” responses and disparities between the lender- and borrower-reported data, borrower uncertainty may be more likely to play a role in the disparity.

Although we largely interpret “don’t know” as genuine uncertainty, we do not assume that these respondents never knew the answer to a given question. Respondents who do not know the index to which their ARMs are linked, for example, may have known this information when they signed the mortgage contracts. We assume only that a respondent did not know the answer at the time of the survey.

### **Homeowner estimates of housing values**

In the aggregate, the changes in self-reported house values in the Survey of Consumer Finances match reasonably well the changes in house values measured by the national OFHEO house price index. To show this finding, we calculate the cumulative percentage change in house price (from year of purchase to 2001) for each SCF household from its self-reported current house value and original purchase price. We then sort these percentage changes by the year of purchase and calculate the median and geometric mean of the cumulative house price changes for each year of purchase.<sup>12</sup>

---

at the same point early in the interview. Kennickell (1997) notes that the number of “don’t know” responses dropped substantially when the SCF introduced a more elaborate set of range responses in the 1995 survey. This finding suggests that “don’t know” responses for these variables had been capturing legitimate respondent uncertainty.

<sup>12</sup> Our analysis can be viewed as a special case of the OFHEO repeat-transactions methodology in which the first transaction is the initial purchase price of the house and the second “transaction,” or observation on the home’s value, is the borrower’s self-reported house value in 2001.

The median controls better for outliers in appreciation rates, whereas the geometric mean corresponds more closely to the OFHEO methodology.<sup>13</sup>

We perform the same exercise for the households in the Residential Finance Survey. The RFS does not provide a clean test of the accuracy of self-reported house price changes, because the house purchase price variable on the RFS is drawn from estimates provided by both the borrower and the lender.<sup>14</sup> However, the RFS results may be more precisely estimated, in a statistical sense, than the SCF results because of the larger RFS sample size. In both datasets, in calculating house price changes we drop farms, ranches, and mobile homes as well as houses that are not owner-occupied, were purchased in 2001, or had a purchase price of \$1 or less. We also drop houses purchased before 1991 on the RFS because the survey does not report the exact year of purchase for these houses.<sup>15</sup>

Two sample selection issues may contribute to the differences between the OFHEO and the SCF and RFS house price measures. First, the OFHEO sample includes only conforming, non-government-backed mortgages for single-family properties. As a robustness test, we limit the SCF and RFS samples to these mortgages to test whether differences are driven by these prices of these properties changing differently over time than prices in the market as a whole. Second, the OFHEO sample includes only transacting houses, whereas the SCF and RFS measure is based on all houses. Several studies have examined this potential selection issue (see, e.g., Kiel and Zabel, 1999,

---

<sup>13</sup> The geometric mean of a sequence  $\{a_i\}_{i=1}^n$  is  $\left(\prod_{i=1}^n a_i\right)^{1/n}$ . The OFHEO index is based on a regression of changes in log house prices. Thus, it models the arithmetic mean of the log difference of house prices, or equivalently the log of the geometric mean of house price growth.

<sup>14</sup> It is not possible to infer from the public RFS data whether the reported purchase price is based on data from the borrower, lender, or both. The estimated current value, however, is reported only by the borrower.

<sup>15</sup> The OFHEO index excludes farms and has only partial coverage of mobile homes. These restrictions result in a sample of 2,377 houses on the SCF and 7,421 houses on the RFS. We also drop houses purchased before 1985 on the SCF because the sample sizes are too small to calculate meaningful house price changes.

DiPasquale and Somerville, 1995, and Goodman and Ittner, 1992); our review of this literature suggests that this issue is not likely to be empirically important and we do not address it here.

The price indexes based on the medians of the household-reported price changes match the OFHEO index reasonably well, particularly for the SCF (figure 1 and table 1a), thereby suggesting that the majority of homeowners are able to report these data with some accuracy. For homes purchased in 2000, the median one-year price change was 7.8 percent in the SCF data and 8.7 percent in the RFS data, compared to a 7.5 percent change in the OFHEO index. Over longer time periods, the price indexes also match well: the cumulative 1991–2001 house price change was 54 percent in the SCF data, 42 percent in the RFS data, and 50 percent in the OFHEO index. The median change in the RFS, although still close to the OFHEO number, tends to understate it except for the most recent years, whereas the SCF median change tracks OFHEO more closely.

In both the RFS and the SCF, the house price measure based on the geometric means is higher than the measure based on the median. This result is not surprising given the skewed distribution of house prices. The SCF measure, in particular, is sometimes substantially larger than the OFHEO index (figure 2). This difference could occur because more expensive homes, which are represented in the SCF data but not in the OFHEO index, have different price dynamics than other homes; because some homeowners are over-stating the increase in their house values; or because of sampling error. The discrepancy is less pronounced in the RFS, perhaps because the sample size is larger or because the data are top-coded.<sup>16</sup> In fact, in farther-out years the RFS geometric means match the OFHEO index closely: over the 1991–2001 period, for example, the RFS measure

---

<sup>16</sup> The overall editing and imputation rates in the two surveys are comparable for the variables used in the construction of the price indexes. For the sample considered in Table 1a, 22 percent of observations in both surveys had the current house value, purchase value, or year of purchase edited or imputed. However, the majority of those in the SCF stem from imputing a value based on a respondent-reported range for the purchase price or current value. Excluding these range imputations, only 3 percent of SCF households had any of these three values edited or imputed. Many of the RFS edits stem from reconciling lender- and household-reported data, whereas a few (1 percent of our sample) stem from replacing extreme values with topcodes to reduce the possibility of identifying individual respondents. The SCF addresses this issue with imputation techniques rather than topcodes.

suggests home values increased 54 percent, compared to a 50 percent increase in the OFHEO index and a 74 increase in the SCF data.

To examine whether differences in the samples skew the results, we repeat this exercise but eliminate households with no mortgage or with a government-backed or nonconforming mortgage (table 1b). Because this restriction reduces the sample of homes purchased since 1985 in the SCF by about 66 percent and eliminates about 50 percent of homes in the RFS sample, the estimates in table 1b are more statistically variable than the earlier estimates.<sup>17</sup> However, they generally follow the same pattern as before and thereby suggest that sample selection is unlikely to explain fully the discrepancies between the OFHEO index and the survey-based numbers. For example, median house-price changes in the SCF still track the OFHEO index reasonably well, particularly for those who purchased their home within the last ten years. The SCF geometric means also generally lie above the OFHEO estimates in this more limited sample. And, as before, the estimates based on the RFS medians fall below the OFHEO values, whereas the estimates based on the geometric mean track the OFHEO index comparatively well for homes purchased in the early 1990s.

Turning now to the edit or imputation flags, the response rates to the SCF questions on house value or house purchases are quite favorable. Ninety-seven percent or more of respondents were able to provide an acceptable response or range when asked the current value of their house, its purchase price, or the year of purchase (table 2).<sup>18</sup> A remarkable 99 percent reported acceptable values for the year of home purchase. These high response rates, combined with the reasonably close correspondence between the household-reported house price changes and the OFHEO index, suggest that households are able to report their house values fairly accurately.

---

<sup>17</sup> These restrictions have a larger effect on the number of SCF observations because high-wealth households are over-represented in the survey. These households are presumably more likely to have nonconforming mortgages or no mortgage at all.

<sup>18</sup> In the case of house values, where the current house value is not known with certainty until the house is sold, a range may be as or more economically meaningful than a single value.

## Homeowner estimates of mortgage data

To examine the accuracy of household-reported mortgage data, we compare the SCF data to the lender-reported LoanPerformance and Residential Finance Survey data. As noted earlier, the LoanPerformance data are reported entirely by lenders and other institutions in the mortgage industry, whereas the Residential Finance Survey data are reported by both borrowers and lenders, with detailed mortgage data reported almost exclusively by lenders. As with the house price results, we then examine the SCF edit or imputation flags to help gauge whether differences across surveys may be due to borrower uncertainty.

In the RFS and SCF data, we limit the sample to first mortgages backed by owner-occupied primary residences and exclude mortgages backed by mobile homes, farms or ranches.<sup>19</sup> In the LoanPerformance data, seven percent of the mortgages are backed by non-owner-occupied properties, including second homes and investment properties, and less than one percent are second liens. Because the data are aggregated, we cannot exclude these mortgages, but the small number of such mortgages suggests that they do not have a substantial effect on the estimates.

To examine whether the differences between the SCF and RFS estimates are statistically significant, we bootstrap the variance of each estimate for each dataset and then sum the two variances.<sup>20</sup> To bootstrap the SCF, we use the replicates and weights provided on the SCF web site that are constructed in accordance with the sample design. We also incorporate imputation

---

<sup>19</sup> We exclude loans backed by mobile homes because many are installment loans that more closely resemble auto loans than traditional residential mortgages. This sample restriction excludes 58 mortgages on the SCF (7 of which are ARMs) and 217 mortgages on the RFS (28 of which are ARMs). We exclude farms and ranches on the SCF because the RFS survey design excludes these dwellings. We also drop 3 reverse mortgages and 111 “other” mortgages on the RFS. Although one or more reverse mortgages exist on the SCF, they cannot be identified from the information in the public data set. We include them in our tabulations so that other users can replicate our results.

<sup>20</sup> For comparison, instead of bootstrapping the variance of each estimate, and then calculating the sum, we bootstrap the variance of the difference. The two approaches provide similar estimates before accounting for imputation uncertainty. Bootstrapping the difference, however, does not easily allow for incorporation of the imputation variance.



uncertainty into the estimates, again using tools provided by the SCF.<sup>21</sup> We do not make these adjustments for the RFS because analogous tools are not available.

In the analysis below, we first examine terms and features applicable to all mortgages and then focus on features specific to adjustable-rate mortgages.

***Terms and features common to all first mortgages.*** The borrower- and lender-reported distributions match well for many mortgage terms (table 3). For example, all three datasets agree that roughly 85 percent of first mortgages were fixed-rate in 2001, slightly more than 10 percent were adjustable-rate, and the rest were balloon.<sup>22</sup> The distribution of amortization periods is also consistent: about a quarter of all mortgages had an amortization period of 15 years or less, between 64 and 70 percent had a 26–30 year amortization period, and the rest were scattered across different categories. Likewise, annual mortgage payments align closely throughout the distribution, with a median of \$8,520 in the RFS data and \$8,400 in the SCF data.<sup>23</sup> This finding is consistent with the Froncznek and Koons (1976) result that borrowers report mortgage payment amounts fairly accurately. As further evidence that borrowers are fairly familiar with these mortgage terms, the SCF edit or imputation flags indicate that over 96 percent of borrowers provided accurate values or ranges for these variables (table 2). The three datasets are not benchmarked against each other, so the close correspondence does not result from the method used to construct the sample weights.

The distributions of interest rates are quite similar in the SCF and RFS even though 9 percent of SCF respondents did not know this value; the median is reported as 7.5 percent in the RFS and

---

<sup>21</sup> Thus, the total estimated variance for each SCF estimate is  $(6/5) * (\textit{imputation variance}) + (\textit{sampling variance})$ , where *imputation variance* is the variance of the SCF point estimates across the five implicates and *sampling variance* is the bootstrapped variance estimate. See Kennickell (2000) for more information on variance estimation procedures in the SCF and Montalto and Sung (1996) for a discussion of calculating point estimates and variances in the SCF.

<sup>22</sup> Adjustable-rate mortgages may be a bit more prevalent in the RFS because lenders are more likely to hold these mortgages on their books rather than securitize them.

<sup>23</sup> The RFS and SCF measures are not completely comparable because the SCF estimate includes homeowner insurance payments for some households. Further, the mortgage payment variable on the RFS combines reports from both lenders and borrowers.

7.25 percent in the SCF.<sup>24</sup> The weighted average coupon in the SCF, which is the average interest rate weighted by the outstanding loan balance, matches the aggregate estimate calculated by the Bureau of Economic Analysis (BEA) almost exactly: 7.35 in the SCF, compared to 7.36 in the BEA.<sup>25</sup>

Other variables do not match as closely across datasets. Although the distributions of the “year of loan origination” variable are quite similar in the RFS and the SCF, the LoanPerformance data show more loans originated in 2001. Timing may explain part of the difference: some SCF and RFS households may have taken out mortgages in 2001 after completing the survey.<sup>26</sup> Although we attempt to address this issue by dropping mortgages originated in the fourth quarter of 2001 from the LP estimate, timing issues may still bias the comparison. Because we cannot access the microdata underlying the LP data, we cannot explore what additional special factors might account for this discrepancy.

The share of mortgages guaranteed by the Federal Housing Administration (FHA) or carrying private mortgage insurance (PMI) is substantially higher in the SCF than the other data sets. In the lender-reported data, about 10 percent of mortgages are FHA-guaranteed, compared to 23 percent on the SCF. The percent of mortgages with PMI is 16 percent in the LP data, 12 percent in the RFS, and 23 percent in the SCF. The edit or imputation flags indicate that 4 and 9 percent of SCF respondents did not know the answers to the government guarantee or PMI questions, respectively, or reported a value deemed inaccurate by the SCF staff.

Private mortgage insurance coverage may be higher in the LP data than the RFS because mortgages securitized by Fannie Mae and Freddie Mac—which require PMI for mortgages with

---

<sup>24</sup> The SCF and RFS estimates do not include upfront fees or points. We exclude imputed interest rates from the RFS calculation because some of the imputed values seem implausible.

<sup>25</sup> We use the 2001 third quarter number from the BEA, available at <http://www.bea.gov/bea/dn/mortfax.xls>.

<sup>26</sup> SCF respondents were interviewed from June to December 2001. RFS respondents were interviewed from April 2001 to January 2002.

loan-to-values (LTVs) over 80 percent—are overrepresented in the LP data. The high coverage rates in the SCF, however, may stem at least partly from borrowers confusing these products with other types of insurance. For example, although PMI is not typically required of borrowers with LTVs below 80 percent, 81 percent of SCF borrowers who report PMI also report a current LTV below 80 percent.<sup>27</sup>

***Terms and features specific to adjustable-rate mortgages.*** Adjustable-rate mortgage terms match well across the datasets for features related to a household’s regular monthly payment (table 4). The RFS and SCF data agree, for example, that the interest rate can change monthly on about 15 percent of ARMs and annually on 50 to 60 percent of ARMs.<sup>28</sup> The three datasets also agree that only a small minority of ARMs have negative amortization features.<sup>29</sup> The SCF edit or imputation flags indicate that about 82 percent of ARM borrowers gave answers deemed accurate for how often the rate can change, and 94 percent gave an answer to the negative amortization question that was deemed accurate. Although these reporting rates are lower than those for broad mortgage terms, they are high relative to the rates for other adjustable-rate mortgage terms.

The distributions do not match well for the terms that govern the change in ARM payments. For example, SCF respondents report lower caps on per-period interest rate changes than RFS respondents. The modal SCF respondent believes that her interest rate can increase, at most, one percentage point per period; the modal RFS respondent reports a limit of two percentage points per

---

<sup>27</sup> Some of these borrowers may be reporting PMI coverage that has been cancelled.

<sup>28</sup> Collapsing the SCF responses to these categories required several assumptions, but the results are robust to minor permutations of these assumptions.

<sup>29</sup> This result should be interpreted with caution because the SCF and RFS questions differ substantially. The SCF asks, “When the interest rate on your mortgage changes, does the size of your monthly payment also change?” The RFS asks, “Does this mortgage allow negative amortization?” However, the RFS only asks this question if the respondent answers “yes” to the question, “Can the regular principal and interest payments change during the life of the mortgage other than through a change in the interest rate?” Only 33 observations responded “yes.” We run the tabulation over this universe only; the results would likely differ if the RFS asked the question for all adjustable-rate mortgages.

period. The estimates of the lifetime cap on interest rate changes are also lower in the SCF.<sup>30</sup> Fifty-seven percent of SCF respondents believe that the most their interest rate can increase over the life of the loan is less than five percentage points. In contrast, only 6 percent of RFS respondents report lifetime caps in this category, with 51 percent indicating caps of five to six percentage points. Only 2 percent of SCF respondents, but 25 percent of RFS respondents, report a lifetime cap greater than 12 percentage points or no cap at all.<sup>31</sup>

Response rates on the SCF are particularly poor for these questions. Thirty-five percent of ARM borrowers did not know the value of the per-period cap on interest rate changes.<sup>32</sup> Similarly, 44 percent of respondents (not shown in table 2) did not know the values of one or both of the two variables used to calculate the lifetime interest cap. Specifically, 41 percent of respondents did not know the maximum interest rate that could be charged over the life of the loan, and 20 percent did not know the interest rate at origination.

Further discrepancies are apparent when respondents are asked what indexes their ARMs are tied to and whether their ARMs are convertible to fixed-rate mortgages. The LoanPerformance and RFS data agree that approximately two-thirds of adjustable-rate mortgages are linked to the rates on U.S. Treasury bills, with the rest linked to a bank cost-of-funds index or the London InterBank Offered Rate (LIBOR). However, only 25 percent of SCF adjustable-rate borrowers report that their mortgage is linked to any of these rates; the modal answer is the prime rate, with 48 percent of respondents, and 20 percent report implausible indexes or colloquial terms (the Consumer Price

---

<sup>30</sup> This question also differs substantially on the SCF and RFS. The SCF asks, “What is the highest the rate can go over the course of the loan?” and “What was the interest rate on this mortgage when you first got it?” We construct the lifetime cap as the difference between these two rates. The RFS asks, “What are the caps on the interest rate change over the life of the mortgage?”

<sup>31</sup> Some ARM borrowers may be unaware of the caps on interest rate changes because negative amortization features of the mortgage mask the interest rate changes. However, the RFS and SCF distributions remain quite different even when mortgages with negative amortization features are excluded.

<sup>32</sup> Of this 35 percent, 33.5 percentage points came from households who did not know the value of the cap, and 1.5 percentage points came from households who did not know if their mortgages were adjustable (1.2 percentage points) or if they had a mortgage (0.3 percentage points), and thus were never asked about the cap. We assume that these households would not know the value of the cap if asked.

Index, the “going” rate, or the Federal Funds rate). Likewise, only 9 percent of adjustable-rate mortgages are reported as convertible to a fixed-rate mortgage on the RFS, yet nearly half of SCF adjustable-rate borrowers indicate that their mortgages are convertible. Perhaps many SCF borrowers consider their mortgages “convertible” if they can refinance without paying prepayment penalties. Reporting rates are also low for these questions: 28 percent of borrowers did not know which indexes their mortgages were tied to, and 16 percent did not know if their mortgages were convertible.

***Demographic subgroups.*** Although ARM borrowers overall may be unaware of their loan terms, some types of borrowers may be more knowledgeable. To examine this possibility, we calculate the share of different subgroups of ARM borrowers who did not know the per-period cap, lifetime cap, or index of their mortgages (table 5). As shown in the last column of the table, these estimates are based on a small number of SCF respondents, so many differences are statistically imprecisely estimated.

We first examine typical demographic subgroups: age, income, education, and race or ethnicity. Age does not appear to be a major factor: younger households are somewhat more likely than older households not to know their ARM indexes, but are slightly less likely not to know their lifetime interest rate caps. High-income borrowers, however, consistently demonstrate more knowledge of their loan terms than middle- and low-income borrowers. For example, 40 percent of borrowers with income less than \$50,000—corresponding roughly to the bottom half of the income distribution of ARM borrowers—do not know the per-period caps on their interest rate changes. In contrast, 13 percent of borrowers with income exceeding \$150,000—roughly the top income decile of ARM borrowers—are unaware of their per-period caps. Higher-income households are also more knowledgeable about their lifetime caps on interest rate changes and their ARM indexes. These differences across income groups are consistent with studies that suggest that consumer

disclosure information benefits middle- and high-income borrowers more than low-income borrowers (Day, 1976).

Reported knowledge of mortgage terms also differs across education and race/ethnicity groups. Twenty-seven percent of borrowers with a college education, compared to 42 percent of borrowers without a college education, do not know their per-period cap. This gap is consistent with studies that indicate that borrowers with less education have lower levels of financial literacy (Lusardi and Mitchell, 2006). Similarly, 26 percent of white borrowers and 59 percent of minority borrowers are unaware of their per-period cap. As with the income variables, these differences are also apparent for the lifetime cap and ARM index and are generally statistically significant.

We next examine the frequency of “don’t know” responses among other groups who may be expected to be more knowledgeable about their mortgage terms. These groups include those who took out a mortgage recently, consulted loan documents during the interview, or expressed above-average interest in the survey. We also report the results for groups for whom the mortgage terms may have more significant financial consequences—those who are unlikely to move soon, who face high and presumably subprime mortgage interest rates, or who have high mortgage payments relative to income. Although intuitively we expect these two broad sets of households to be more aware of their loan terms, in practice they are not.<sup>33</sup> For the most part, the differences between these groups and other households are statistically insignificant and are often small.

To examine the extent to which these differences across groups reflect correlated demographic factors, we estimate multivariate regression versions of table 5, shown in table 6. We find that income remains a large and significant factor even after we control for other variables. In particular, the regression estimates suggest that households with income less than \$50,000 are 28 percentage points more likely than high-income households to be unaware of the per-period cap on their

---

<sup>33</sup> Other studies, in fact, find that subprime borrowers are less knowledgeable about mortgages than prime borrowers. See, for example, Courchane, Surette, and Zorn (2004) and Lax, Manti, Raca, and Zorn (2004).

interest rate change, almost the same as the 27 percentage point difference implied by the estimates in table 5. Likewise, the regression indicates that lower-income households are 31 percentage points less likely than high-income households to know the lifetime cap on interest rate changes, compared to a 32 percentage point difference in table 5.

Education and race/ethnicity are also factors in the regressions, although the effects are smaller than the table 5 estimates and are only statistically significant for one of the three mortgage terms. Borrowers without a college education, for example, are 27 percentage points more likely than borrowers with a college education to be unaware of the indexes to which their mortgages are tied. Minority borrowers are 30 percentage points less likely than white borrowers to know the per-period cap on their interest rate change. The regression results also suggest, somewhat unintuitively, that borrowers who are more likely to stay in their homes over the next 2 years are less likely to know their lifetime caps and ARM index.

***Simulated effect of an increase in ARM interest rates.*** If borrowers underestimate the caps on possible interest rates changes, they are likely to underestimate the amount by which their monthly payments could rise if the underlying indexes increase substantially. To provide a rough estimate of the possible difference between the distributions of actual and perceived payment increases, we simulate the increase in monthly payments that would occur, according to the RFS and SCF data, if the ARM index increased 2 percentage points for two consecutive years, for a cumulative 4 percentage point increase. This 4 percentage point rise is a bit larger than the approximately 3-1/4 percentage point rise in the U.S. one-year constant-maturity Treasury—the most common ARM index—from mid-2003 to the end of 2005.<sup>34</sup> We assume that the ARM index is the only factor that changes. Borrowers do not respond to the increase, for example, by moving

---

<sup>34</sup> Four percentage point increases in the U.S. one-year Treasury were seen most recently in the 1970s and early 1980s.

or by refinancing their mortgages. If we allowed for such responses, the rise in interest rates might be associated with a smaller increase in household mortgage payments.

In considering the results of this simulation, three caveats should be kept in mind. First, the simulations illustrate the possible effects of a rate increase in 2001 and not necessarily at any other point in time. For example, in 2005 the payments on many ARMs were governed by “option” or “hybrid” features that were largely unknown in 2001.<sup>35</sup> Second, because the most commonly cited cap on the per-period interest rate change is 1 percentage point in the SCF and 2 percentage points in the RFS, the parameters chosen in this simulation likely maximize the difference between the SCF and RFS payment-increase distributions. If the ARM index increased only one percentage point a year, the distributions of implied payment increases in the SCF and RFS would almost coincide.

Third, the prevalence of missing values in both data sets raises concerns about the robustness of our estimates. For the RFS, we only keep in the simulation observations with no missing values for any variable used in the simulation. This restriction drops the RFS sample of ARMs from 1,321 to 320. For the SCF, we include all ARMs in the simulation, even though 69 percent (153 out of 223) of these mortgages have at least one necessary variable imputed. Thus, for households who did not report particular loan terms, we assume that the imputed values represent the values the households would report if pressed to do so. Although this assumption may be questionable, we believe it is the best way to mitigate sample selection problems. In practice, running the simulation and excluding SCF households with imputed values has only a minor effect on the results.

Keeping these caveats in mind, we find that even under this somewhat extreme interest rate scenario, most ARM borrowers experience fairly small payment increases in this simulation. The RFS data suggest that 68 percent of borrowers will experience changes in payments less than 5

---

<sup>35</sup> In an option ARM, borrowers have the choice each month whether to make a full payment that includes both interest and principal or to make a smaller payment. In a hybrid ARM, the interest rate is fixed for an initial period, usually 3 to 10 years, and then becomes a standard ARM.



percent of gross income (table 7), whereas the SCF data suggest that 79 percent of borrowers anticipate changes of this size. Within this category, the RFS data show slightly larger payment changes for borrowers than the SCF data. The median change in payments for the sample as a whole, for example, is 3 percent of gross income in the RFS and 1.8 percent in the SCF.

A small share of borrowers, however, experience larger increases. The RFS predicts that 23 percent of borrowers would experience an increase between 5 and 10 percent of income, whereas 15 percent of SCF borrowers anticipate an increase this large. Likewise, the RFS indicates that 9 percent of borrowers would have an increase over 10 percent of income, whereas 6 percent of SCF borrowers anticipate increases of this size. In total, the difference between these distributions suggests that 11 percent of SCF ARM borrowers would not have anticipated that the change in their mortgage payments exceeds 5 percent of income.

The simulation also suggests that borrowers in the lower half of the income distribution of ARM borrowers are more likely to experience large increases in payments relative to income. Sixteen percent of these borrowers, according to the RFS, and 12 percent, according to the SCF, might experience payment increases larger than 10 percent of their incomes under these circumstances. In contrast, only a few borrowers in the top half of the distribution experience payment changes of this magnitude.

Although lower-income ARM borrowers appear more vulnerable than higher-income borrowers to payment changes, a fair number of these borrowers can draw upon home equity or other forms of wealth to help smooth through these shocks. For example, of the lower-income households with projected payment changes exceeding 5 percent of income, about half have mortgages with current loan-to-value ratios below 80 percent. In addition, the difference between the RFS and SCF shares of ARM borrowers in each payment change category does not differ

markedly across income groups, a fact that suggests that lower-income households are roughly as likely as other households to anticipate these changes.

Two results from this simulation may seem puzzling: first, that so many ARM borrowers experience relatively small changes in payments, and second, that lower-income borrowers are not more likely than higher-income borrowers to be surprised by their changes in payments. A partial answer to the first puzzle may be that many borrowers already have fairly small mortgage payments relative to income: the median payment-to-income ratio on both the RFS and the SCF is around 12 percent. In addition, if the remaining maturity on a mortgage is short, either because the original maturity was short or because the household has been paying down the mortgage for many years, the payment is less sensitive to interest rate fluctuations. Our data suggest that about a quarter of ARMs in our simulation have a remaining maturity of 15 years or less.

The answer to the second puzzle may be that although lower-income ARM borrowers are less likely to know their loan terms, those lower-income borrowers who think they know their loan terms—and thus report data to the SCF—are about as accurate as borrowers with more income. Of course, this result depends on our assumption that the values imputed for borrowers who responded “don’t know” are reasonable proxies for their expectations. If this assumption is incorrect, the share of lower-income borrowers who are surprised may be higher.<sup>36</sup>

## **Discussion**

Aggregate estimates of gains in housing wealth and of broad mortgage terms from household-reported data match estimates from lender-reported data quite well. Given the differences in design and methodology across data sets, this close correspondence is somewhat remarkable. The evidence

---

<sup>36</sup> Sample sizes in some of the cells of table 7 are also quite small.

presented in this paper suggests that policymakers and researchers can generally have confidence in homeowners' knowledge of these variables.

Lender- and household-reported data do not match well, however, for many of the terms that govern how the payment on an adjustable-rate mortgage changes. Differences in how questions are asked or which types of mortgages are represented in the data may underlie some of these discrepancies. In particular, the fact that lender-reported data are missing for about half of RFS mortgages raises the concern that the RFS data are not representative. Because the LoanPerformance and RFS data match closely for most variables, though, we conclude that sample selection does not seem to be a major issue.

Instead, the SCF edit or imputation flags suggest that borrower confusion underlies some of the discrepancy. Borrowers with less income or education seem especially likely not to know their mortgage terms. This confusion may be a concern for policymakers as well as researchers, because borrowers who do not recall their mortgage terms also may not understand fully the inherent risks. In addition, borrowers with a limited understanding of mortgage terms may take out subprime mortgages when they could have qualified for less-costly mortgages (Lax, Manti, Raca, and Zorn, 2004).

For some adjustable-rate mortgage terms, this borrower confusion may be of little practical importance. Lack of knowledge of ARM indexes, for example, may not matter much because indexes tend to move together over the long run. For borrowers who plan to move in the near future, or who have enough wealth to weather market fluctuations, awareness of other ARM terms may also be less important.

For other borrowers, this confusion might have more serious consequences. Our simulation suggests, for example, that lower-income borrowers might be more vulnerable to an increase in interest rates. Borrowers who do not monitor financial markets and changes in their ARM index

may also be surprised by the change in their payments. Overall, though, the simulation suggests that only a small fraction of ARM borrowers, regardless of their knowledge of mortgage terms, will face large unexpected increases in payments in response to a sustained rise in short-term interest rates.

## REFERENCES

- Agarwal, Sumit, "Why do Homeowners Misestimate their House Value?," unpublished manuscript, Bank of America, 2005.
- Aizcorbe, Ana M., Arthur B. Kennickell, and Kevin B. Moore, "Recent Changes in U.S. Family Finances: Evidence from the 1998 and 2001 Survey of Consumer Finances," *Federal Reserve Bulletin*, January 2003.
- Alessie, Rob, and Christine Zandvliet, "Measurement of Household Saving Obtained from First-Differencing Wealth Estimates," VSB-CentER Savings Project Progress Report 13, 1993.
- Courchane, Marsha J., Brian J. Surette and Peter M. Zorn, "Subprime Borrowers: Mortgage Transitions and Outcomes," *Journal of Real Estate Finance and Economics*, 29(4):365-92, 2004.
- Day, George S., "Assessing the Effects of Information Disclosure Requirements," *Journal of Marketing*, 40(2):42-52, April 1976.
- DiPasquale, Denise, and C. Tsurriel Somerville, "Do House Price Indices Based on Transacting Units Represent the Entire Stock? Evidence from the American Housing Survey," *Journal of Housing Economics*, 4: 195-229, 1995.
- Fronczek, Peter, and David Koons, "Study of Homeowners' and Lenders' Responses for Monthly Mortgage Payments, Yearly Real Estate Taxes, and Yearly Property Insurance Payments," unpublished manuscript, Department of the Census, September 1976.
- Goodman, John L. Jr., and John B. Ittner, "The Accuracy of Home Owners' Estimates of House Value," *Journal of Housing Economics*, 2(4):339-57, December 1992.
- Kennickell, Arthur B., "Imputation of the 1989 Survey of Consumer Finances: Stochastic Relaxation and Multiple Imputation," October 1991. Available at <http://www.federalreserve.gov/pubs/oss/oss2/method.html>
- Kennickell, Arthur B., "Using Range Techniques with CAPI in the 1995 Survey of Consumer Finances," January 1997. Available at <http://www.federalreserve.gov/pubs/oss/oss2/method.html>
- Kennickell, Arthur B., "Multiple Imputation in the Survey of Consumer Finances," September 1998. Available at <http://www.federalreserve.gov/pubs/oss/oss2/method.html>.
- Kennickell, Arthur B., "Revisions to the Variance Estimation Procedure for the SCF," October 2000. Available at <http://www.federalreserve.gov/pubs/oss/oss2/method.html>.
- Kennickell, Arthur B., "Reordering the Darkness: Application of Effort and Unit Nonresponse in the Survey of Consumer Finances," August 2003. Available at <http://www.federalreserve.gov/pubs/oss/oss2/method.html>.
- Kennickell, Arthur B., and Martha Starr-McCluer, "Retrospective Reporting of Household Wealth: Evidence from the 1983-1989 Survey of Consumer Finances," *Journal of Business and Economic Statistics*, 15(4):452-63, October 1997.
- Kiel, Katherine A., and Jeffrey E. Zabel, "The Accuracy of Owner-Provided House Values: The 1978-1991 American Housing Survey," *Real Estate Economics*, 27(2): 263-98, 1999.
- Lax, Howard, Michael Manti, Paul Raca, and Peter Zorn, "Subprime Lending: An Investigation of Economic Efficiency," *Housing Policy Debate*, 15(3):533-71, 2004.

- Leary, Jesse B., David Newhouse, and Kata Mihaly, “The Dynamics and Wealth Effects of High-Rate Loans,” unpublished manuscript, Federal Trade Commission, 2004.
- Lusardi, Annamaria, and Olivia S. Mitchell, “Financial Literacy and Planning: Implications for Retirement Wellbeing,” Pension Research Council Working Paper 2006-1, 2006.
- Montalto, Catherine Phillips, and Jaimie Sung, “Multiple imputation in the 1992 Survey of Consumer Finances,” *Financial Counseling and Planning*, 7: 133–46, 1996.
- U.S. Census Bureau, Census 2000 Special Reports, CENSR-27, *Residential Finance Survey: 2001*, 2005. Available at <http://www.huduser.org/datasets/rfs.html>.

FIGURE 1: HOUSEHOLD-REPORTED HOUSING APPRECIATION, 1991–2000  
 MEDIAN COMPARISON

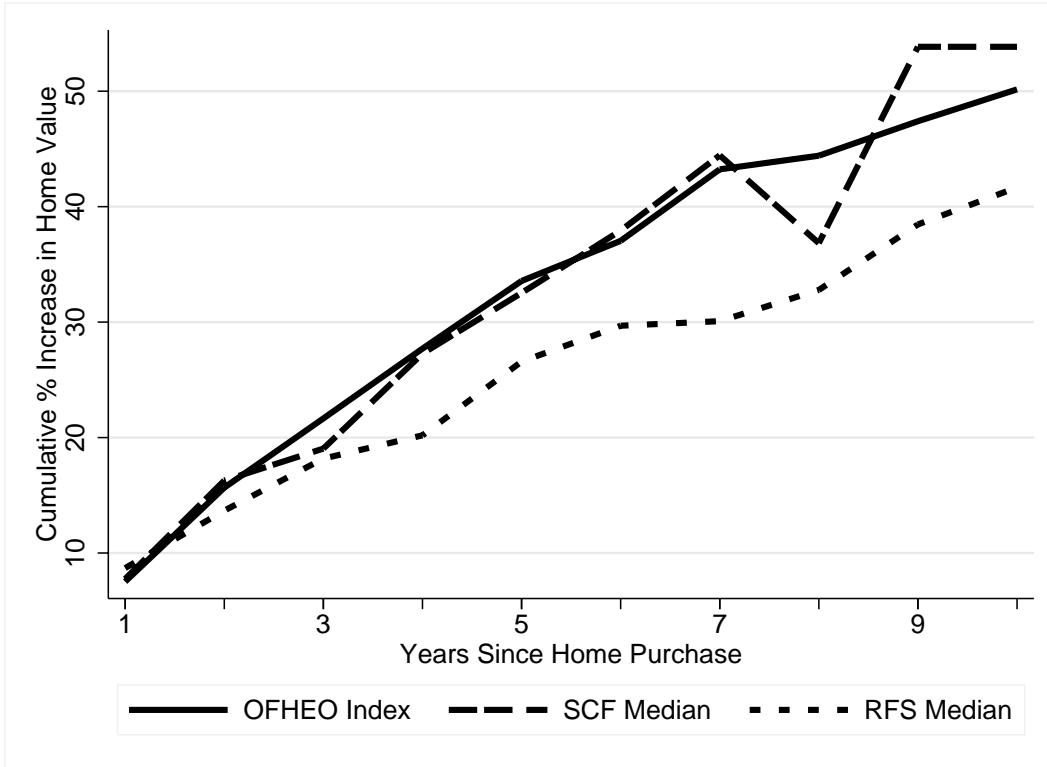


FIGURE 2: HOUSEHOLD-REPORTED HOUSING APPRECIATION, 1991–2000  
 GEOMETRIC MEAN COMPARISON

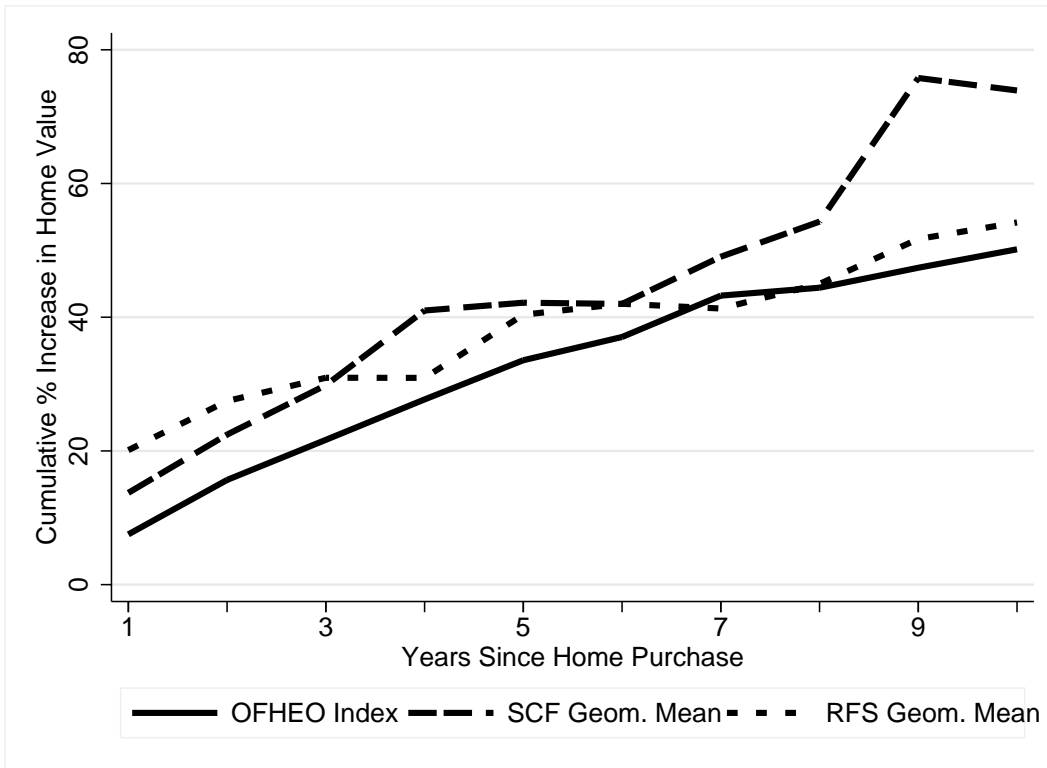


TABLE 1A. COMPARISON OF HOUSE VALUE APPRECIATION ACROSS DATA SOURCES  
ALL HOUSES

Year Purchased	OFHEO Index (Percent)	2001 Survey of Consumer Finances			2001 Residential Finance Survey		
		Count	Median	Geom. Mean	Count	Median	Geom. Mean
2000	7.5	197	7.8	13.7	1020	8.7	20.1
1999	15.7	208	16.3	22.4	1174	13.7	27.5
1998	21.7	169	19.0	29.8	1108	18.2	30.9
1997	27.7	139	27.3	41.0	805	20.2	30.9
1996	33.6	135	32.5	42.2	739	26.6	40.4
1995	37.0	142	37.9	42.0	649	29.7	42.0
1994	43.2	110	44.4	49.0	576	30.1	41.3
1993	44.4	118	36.8	54.3	504	32.8	45.0
1992	47.4	84	53.8	75.8	450	38.5	51.7
1991	50.2	105	53.8	73.9	396	41.7	54.2
1990	54.0	107	57.9	75.9			
1989	54.4	96	54.2	71.9			
1988	63.7	83	84.8	93.4			
1987	73.7	86	75.7	83.8			
1986	85.4	80	104.5	125.3			
1985	100.5	82	83.3	111.2			

**NOTES:** The Residential Finance Survey does not provide the exact year of purchase for houses purchased before 1991. RFS and SCF estimates are weighted with sample weights. Samples include owner-occupied primary residences excluding mobile homes, farms, ranches, and houses with a purchase price of \$1 or less.



TABLE 1B. COMPARISON OF HOUSE VALUE APPRECIATION ACROSS DATA SOURCES  
OVER HOUSES ELIGIBLE FOR INCLUSION IN THE OFHEO INDEX

Year Purchased	OFHEO Index (Percent)	2001 Survey of Consumer Finances		2001 Residential Finance Survey		
		Median	Geom. Mean	Count	Median	Geom. Mean
2000	7.5	7.8	13.9	495	8.3	19.9
1999	15.7	13.6	13.2	559	13.6	26.0
1998	21.7	14.3	23.3	558	15.8	27.4
1997	27.7	30.0	37.5	428	22.2	29.9
1996	33.6	30.0	38.8	382	26.8	41.8
1995	37.0	23.8	29.9	342	29.6	40.5
1994	43.2	51.5	55.0	300	29.6	41.0
1993	44.4	42.9	67.4	263	35.1	47.8
1992	47.4	56.3	64.1	239	36.0	55.0
1991	50.2	56.8	52.8	193	42.9	52.5
1990	54.0	76.5	102.4			
1989	54.4	68.5	76.5			
1988	63.7	103.3	111.0			
1987	73.7	67.4	81.1			
1986	85.4	100.0	124.4			
1985	100.5	87.5	129.9			

**NOTES:** The Residential Finance Survey does not provide the exact year of purchase for houses purchased before 1991. Samples include unattached, mortgaged primary residences (excluding mobile homes, farms and ranches) that are not financed by government-insured loans or by loans that exceed the conforming loan limits. These restrictions rely in part on variables that are unavailable on the public SCF data. Counts are suppressed for the SCF estimates for disclosure protection reasons. RFS and SCF estimates are weighted with sample weights.

TABLE 2. REPORTING RATES FOR HOUSING AND MORTGAGE CHARACTERISTICS: 2001 SCF

Variable (Percent applicable)	Original Value	Range	Edited Value	Don't Know	Refused	Missing Due to Editing
<i>House Value</i>						
Current value of home (66.1%)	88.0	9.7	0.1	0.9	0.8	0.6
Purchase price of home (66.1%)	90.7	6.3	0.1	1.0	1.3	0.6
Year home purchased (66.1%)	99.3		0.1	0.5	0.1	<0.1
<i>Mortgage Terms and Features</i>						
Adjustable rate (43.0%)	98.8			0.8	0.2	0.2
Amortization period (43.0%)	98.4		0.1	0.9	0.3	0.3
Amount of regular payment (42.1%)	90.7	5.5	0.3	0.8	2.2	0.6
Annual interest rate (43.0%)	89.8		0.6	9.0	0.3	0.3
Year mortgage obtained (43.0%)	97.8		0.1	0.9	0.4	0.7
Government guarantor (12.8%)	95.7			3.8	0.2	0.3
Private mortgage insurance (30.0%)	91.0			5.6	0.2	3.2
<i>Adjustable Rate Mortgage Terms and Features</i>						
Frequency rate can change (4.9%)	82.3		0.2	16.7	0.5	0.2
Negative amortization allowed (4.9%)	93.6			4.9	0.5	1.0
Maximum rate can rise at once (4.9%)	62.1		0.7	35.1	1.1	1.0
Maximum rate can be charged (4.9%)	55.6			41.1	2.2	1.0
Original interest rate (4.9%)	75.7			20.2	2.2	1.9
On what index does it depend (3.8%)	69.1			28.3	0.4	2.2
Convertible mortgage (4.9%)	83.0		< 0.1	15.5	0.5	1.0

Number of obs: 4,161

**NOTES:** Estimates are weighted with sample weights. Mobile homes, farms, and ranches are excluded. Reporting rates are calculated over the sample of observations for which the question was applicable.

TABLE 3. COMPARISON OF MORTGAGE TERMS ACROSS DATA SETS  
ALL FIRST MORTGAGES

	LoanPerformance (December 2001)	Residential Finance Survey (2001)	Survey of Consumer Finances (2001)
<i>Mortgage type</i>			
Fixed	86	83	87***
Adjustable	11	13	11**
Balloon	3	5	2***
<i>Amortization period</i>			
1-15	23	25	27**
16-20	4	4	5**
21-25	1	1	2***
26-30	70	69	64***
30+	2	1	1
<i>Annual mortgage payment</i>			
10 <sup>th</sup> percentile		3,840	3,840
25 <sup>th</sup> percentile		5,760	5,760
Median		8,520	8,400
75 <sup>th</sup> percentile		13,200	12,400
90 <sup>th</sup> percentile		18,000	19,200
<i>Interest rate (percentage points)</i>			
10 <sup>th</sup> percentile		6.50	6.38**
25 <sup>th</sup> percentile		6.88	6.88
Median		7.50	7.25***
75 <sup>th</sup> percentile		8.25	8.00***
90 <sup>th</sup> percentile		9.88	9.00***
<i>Year of loan origination</i>			
2001	20	12	11
2000	9	11	11
1999	14	16	16
1998	19	16	14**
1997	6	7	8
Earlier	31	38	40**
<i>Government guaranteed?</i>			
FHA	10	11	23***
VA or other	3	7	7
Conventional	87	81	70***
<i>Private mortgage insurance</i>			
Yes	16	12	23***
No	84	88	77***

**NOTES:** Loans backed by mobile homes, farms or ranches are excluded from the SCF and RFS tabulations. RFS and SCF estimates are weighted with sampling weights.

Differences in proportions relative to the RFS are statistically different from zero at the \*\*\* 1 percent level or \*\* 5 percent level. Standard errors are bootstrapped with 999 replicates; those for the SCF are drawn in accordance with the sample design and adjusted for imputation uncertainty.

TABLE 4. COMPARISON OF MORTGAGE TERMS ACROSS DATA SETS  
ADJUSTABLE-RATE FIRST MORTGAGES

	LoanPerformance (December 2001)	Residential Finance Survey (2001)	Survey of Consumer Finances (2001)
<i>Frequency with which interest rate can change</i>			
Monthly		15	15
Quarterly or every six months		10	15
Annually		60	52**
3 or 5 years		12	14
Other		3	5
<i>Negative amortization allowed</i>			
Yes	9	16	13
No	91	84	87
<i>Caps on interest rate changes per period</i>			
Less than 1 ppt.		1	12***
1 ppt.		0.5	28***
Between 1 & 2 ppts.		5	3
2 ppts.		47	23***
Between 2 & 9 ppts.		18	15
9 ppts +		10	11
No caps		18	7***
<i>Caps on lifetime interest rate changes</i>			
Less than 5 ppts.		6	57***
5 ppts.		25	6***
5.01–6 ppts.		26	17**
6.01–11.99 ppts.		9	18**
12 ppts.		8	0.3***
GT 12 ppts.		13	2***
No caps		12	0***
<i>ARM Index</i>			
Treasury bills	67	64	14***
Cost of Funds Index	19	15	4***
LIBOR or CD	14		7
Prime			48
Consumer Price Index			10
“Going” rate			5
Federal funds rate			5
<i>Convertible to a fixed-rate mortgage?</i>			
Yes		9	47***
No		91	53***

**NOTES:** Loans backed by mobile homes, farms or ranches are excluded from the SCF and RFS tabulations. RFS and SCF estimates are weighted with sampling weights.

Differences in proportions relative to the RFS are statistically different from zero at the \*\*\* 1 percent level or \*\* 5 percent level. Standard errors are bootstrapped with 999 replicates; those for the SCF are drawn in accordance with the sample design and adjusted for imputation uncertainty.

TABLE 5. SHARE OF “DON’T KNOW” RESPONSES BY SUBGROUP

	Per-period Cap	Lifetime Cap	Index	N
<i>All ARM borrowers</i>	35	41	28	244
<i>Age of mortgage borrower</i>				
25–44	33	37	33**	104
45–64	33	41	19	114
<i>Income of mortgage borrower</i>				
Less than \$50K	40***	53*** ††	40*** †	69
\$50K–\$150K	35***	34	23**	71
More than \$150K	13	21	8	104
<i>Education</i>				
College education	27**	34*	14***	158
Less than college	42	47	43	86
<i>Race or ethnicity</i>				
White, non-Hispanic	26***	37**	24*	204
Nonwhite or Hispanic	59	52	38	40
<i>Year of mortgage origination</i>				
2000 or 2001	46**	47	39*	67
1999 or earlier	31	39	25	177
<i>Documents consulted during interview</i>				
Loan documents	44	47	23	33
Other relevant documents	38	59	22	15
Other or no documents	33	38	30	196
<i>Interest in the interview</i>				
Very high or above average	35	40	27	162
Average	30*	40	28	67
Below average or very low	55	53	38	14
<i>Chance stay at residence next 2 years</i>				
30 percent or less	41	22*** †	22	30
40–70 percent	31	46	23	33
80 percent or higher	35	45	31	180
<i>Subprime</i>				
No	37	42	25*	219
Yes	23	36	45	24
<i>Mortgage debt service ratio</i>				
15 percent or less	32	40	26	141
15–28 percent	35	35	27	61
Greater than 28 percent	41	51	36	42

**NOTES:** Data are from the 2001 Survey of Consumer Finances. Estimates are weighted with sampling weights. Sample size in final column is the unweighted count of observations to which the per-period cap and lifetime cap questions applied; counts for the ARM index are roughly 80% lower (varying between 69-91%) since this question is asked only of those ARM borrowers who additionally said their interest rate depended on another rate (203 observations total). “Subprime” is defined as having a mortgage interest rate more than 3 percentage points above the annual rate on the comparable 30-year Treasury. “Other relevant documents” includes account statements, real estate records, and other miscellaneous personal documents.

Differences relative to final category within a group are statistically different from zero at the \*\*\* 1 percent level; \*\* 5 percent level; or \* 10 percent level. For variables with three categories, differences relative to the middle category are statistically different from zero at the †† 1 percent level; † 5 percent level; or † 10 percent level. Standard errors are bootstrapped with 999 replicates in accordance with the SCF sample design and are adjusted for imputation uncertainty.

TABLE 6. OLS REGRESSION ESTIMATES OF “DON’T KNOW” RATES

	Per-period Cap	Lifetime Cap	Index
Age 25–44	-.05 (.07)	-.03 (.07)	.21*** (.07)
Older than 65 or younger than 25	.04 (.12)	.10 (.13)	.21 (.14)
Income <\$50K	.28*** (.10)	.31*** (.12)	.21 (.13)
Income \$50K–\$150K	.24*** (.07)	.14 (.10)	.05 (.09)
No college education	.05 (.08)	.06 (.08)	.27*** (.07)
Minority	.30*** (.07)	.12 (.08)	.11 (.08)
Mortgage originated in 2000 or 2001	.17** (.07)	.06 (.09)	.07 (.08)
Consulted loan documents during interview	.12 (.08)	.10 (.11)	-.01 (.08)
Consulted selected other documents	.04 (.17)	.17 (.19)	-.16 (.14)
Interest in interview very high or above average	-.07 (.15)	-.03 (.14)	.19 (.16)
Average interest in interview	-.18 (.16)	-.05 (.14)	.10 (.17)
Probability of staying at residence next 2 years 80 percent or greater	-.03 (.12)	.23** (.11)	.17** (.08)
Probability of staying 40–70 percent	-.14 (.14)	.24* (.14)	.01 (.15)
Subprime loan	-.22** (.10)	-.18 (.13)	.10* (.13)
Subprime missing	.02 (.05)	.06 (.15)	-.05 (.11)
Mortgage debt service ratio >28 percent	-.03 (.11)	-.06 (.10)	.04 (.12)
Mortgage debt service ratio 15–28 percent	.03 (.09)	-.08 (.09)	.01 (.12)
Constant	.14 (.22)	.01 (.20)	-.42** (.20)
Adjusted R <sup>2</sup>	.18	.13	.21
N	244	244	203

**NOTES:** Data are from the 2001 Survey of Consumer Finances. Estimates are weighted with sampling weights. Adjusted R-squared based on regression over all 5 SCF implicates. Standard errors are bootstrapped with 999 replicates and are adjusted for imputation uncertainty.

\*\*\* Statistically significant at the 1 percent level; \*\* 5 percent level; \* 10 percent level.

TABLE 7. SIMULATED EFFECT OF A FOUR PERCENTAGE-POINT INCREASE  
IN THE ARM INDEX ON ARM PAYMENTS

	Change in payment as a percent of gross income		
	Less than 5 percent	5-10 percent	Over 10 percent
<u>All ARM borrowers</u>			
RFS	68	23	9
SCF	79	15	6
<u>By position in income distribution</u>			
<i>Bottom half</i>			
RFS	60	24	16
SCF	70	18	12
<i>50<sup>th</sup>-90<sup>th</sup> percentiles</i>			
RFS	77	20	3
SCF	91	9	< 1
<i>Top decile</i>			
RFS	68	28	4
SCF	78	22	< 1
<b>Memo</b>			
<u>Median change as percent of income</u>			
RFS	3.0		
SCF	1.8		

**NOTES:** Simulation assumes that the ARM index increases 2 percentage points for 2 consecutive years. Estimates are weighted with sampling weights. Income ranges refer to position in the income distribution of ARM borrowers.