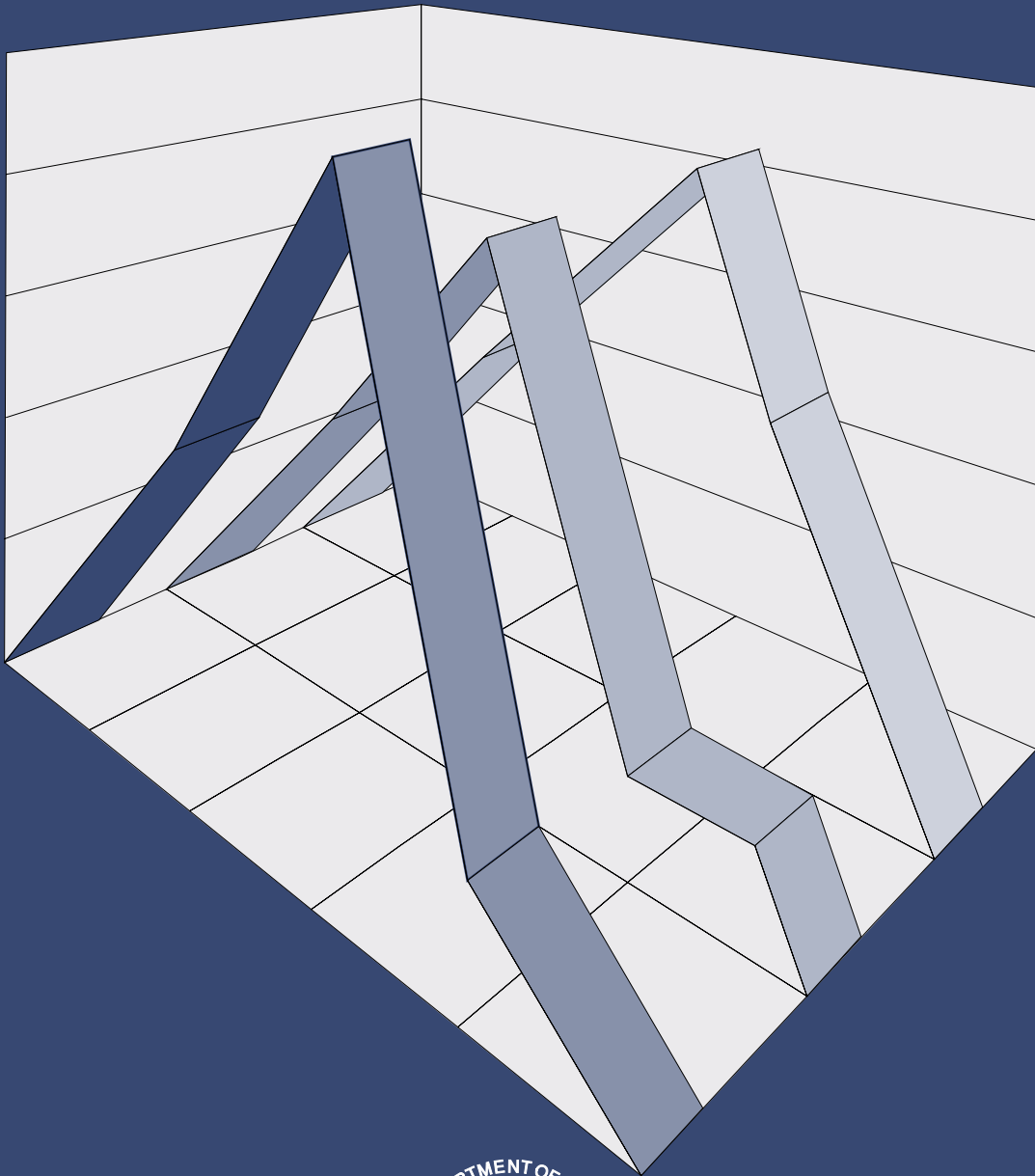




OFFICE OF POLICY DEVELOPMENT & RESEARCH

Analysis of Housing Finance Issues Using the American Housing Survey (AHS)



Analysis of Housing Finance Issues Using the American Housing Survey (AHS)

Prepared for:

U.S. Department of Housing and
Urban Development, PD&R

Prepared by:

Ken Lam
Bulbul Kaul

Abt Associates Inc.
Cambridge, MA

April 2003

Acknowledgements

This study has been made possible with the help of numerous persons and organizations.

Eric S. Belsky of the Joint Center for Housing Studies at Harvard University served as a consultant to this study. He has contributed to both the Research Design and Final Report phases of the study.

In addition to the report authors, several staff members from Abt Associates were instrumental in the study design, data collection, analysis, and final report production. In particular, Chris Herbert, Meryl Finkel and Jill Khadduri, as Technical Reviewers, have provided thoughtful and constructive comments on early versions of the report. Deborah Gruenstein and Elizabeth Burns assisted with the design phase of the study. Jessica Bonjorni and Missy Robinson provided production support throughout the project.

Various staff members from the Census Bureau and HUD's Office of Policy Development & Research, including Bill Reeder and Harold Bunce, have provided valuable feedback on early versions of this report. The report benefited greatly from their insights. We are especially grateful to the Government Technical Monitor, Sue George Neal, for her guidance, assistance, and support throughout the study.

The contents of this report are the views of the contractor and do not necessarily reflect the views or policies of the U.S. Department of Housing and Urban Development of the U.S. Government.

Table of Contents

Executive Summary	v
Overview	v
Replication Analysis.....	vi
Internal Consistency of AHS Mortgage Variables Across Surveys	vii
Reliable AHS Mortgage Variables	vii
Unreliable or Uncertain AHS Mortgage Variables	ix
Areas Requiring Further Investigation	xi
Recommendations for Modifications to the Survey	xi
Chapter One - Introduction	1
Background and Policy Context	1
Outline of the Report.....	2
Chapter Two - The AHS and Research on Mortgages	4
History and Basic Characteristics of the AHS.....	4
Mortgage-Related Variables in the AHS.....	5
Geographic Coverage of the Survey.....	10
Significant Changes to the AHS Over Time	11
Advantages and Disadvantages of the AHS for Mortgage-Related Research.....	14
AHS Mortgage Variables Chosen for the Replication Analysis	16
Chapter Three - Comparison Databases	19
Principal Sources of Comparison Databases	19
Other Comparison Databases Considered	22
Choice of Comparison Databases for the Replication Analysis	23
Chapter Four - Replication Analysis.....	26
Goals and Approach for the Replication Analysis	26
Mortgage Origination Volume	32
Loan Characteristics	45
Borrower Characteristics	68
Summary of Major Findings on AHS Mortgage Variables Tested in the Replication Analysis.....	82
Chapter Five - Internal Consistency of AHS Mortgage Variables Across Surveys.....	88
Goals and Approach	88
Consistency of Mortgage Variables Across Time	91
Summary of Major Findings	99
Chapter Six - Conclusions and Recommendations	101
Reliable AHS Mortgage Variables	101
Unreliable or Uncertain AHS Mortgage Variables	103
Potential Use of the AHS for Research on Mortgages	105

Areas Requiring Further Investigation	110
Recommendations for Modifications to the Survey	111
Appendix A - GSE Conforming Loan Limits, 1980-2000.....	A-1
Appendix B - Summary Information of Comparison Databases.....	B-1
Appendix C - Alternative Approaches to Adjusting the Mortgage Origination Volume Estimates in AHS	C-1
Appendix D - Tabulations Supporting the Analysis in Chapter Four	D-1
Appendix E - Procedures for Tracking Mortgage Records Across AHS Survey Years.....	E-1
Appendix F - Tabulations Supporting the Analysis in Chapter Five.....	F-1
Appendix G - Mortgage and Purchase Modules in the 1997 AHS Questionnaire	G-1
Appendix H - Derivation of AHS Mortgage Market Attributes Tested in the Study.....	H-1

Executive Summary

Overview

The American Housing Survey (AHS, formerly Annual Housing Survey) is the most comprehensive source of information about the characteristics and condition of the nation's housing stock. Started in 1973, the AHS national sample data were collected by the U.S. Census Bureau for the Department of Housing and Urban Development (HUD) on a nationally representative sample of housing units every year until 1981, and they have been collected every other year since then. Over the years, AHS data have been used extensively by researchers and policy analysts to answer questions about the nation's housing conditions and occupant characteristics. The longitudinal nature of the AHS also permits the analysis of dynamic changes in housing and occupancy characteristics of the housing stock.

The AHS data contain detailed questions about mortgages asked of respondents for owner-occupied units in the survey. The questions cover most basic mortgage and housing finance topics. This wealth of mortgage-related variables, combined with the occupant demographic and property location information, could be a very powerful resource for answering many housing finance research and policy questions. These micro-data have the potential to provide crucial information to support analysis of issues of interest to policy makers and the mortgage industry. The principal advantage of using the AHS for mortgage market and housing finance analysis is its detailed household, housing, loan, and geographic characteristics. In addition to use in detailed cross tabulations, these variables can also be used as micro data to conduct multiple regressions on the cross-sectional files and other loan-level statistical analyses on the longitudinal panels.

However, neither the housing research community nor HUD staff has made as much use of the mortgage variables of the AHS data as might be expected. Among the reasons for this under-utilization is the fact that the reliability of these mortgage-related variables in the AHS has not been verified. Analysis is needed to establish the extent to which limitations associated with sample size, survey design, and interview response affect the accuracy and consistency of the mortgage data in the AHS. The research presented in this study is intended to meet this need. The goals of the analysis are to determine: 1) what types of mortgage market analysis can be supported by the AHS; 2) what areas of the AHS are problematic for mortgage research; and 3) what analysis techniques or changes in the survey could potentially compensate for the problems.

The analysis of the reliability of the AHS is composed of two broad categories. First, to test the reliability of the AHS variables we replicate measures of mortgage activity from other reliable sources of data that serve as benchmarks for the AHS estimates. This analysis is referred to as the "replication analysis." Second, we use the longitudinal nature of the AHS to determine whether answers to questions on mortgages are consistent and stable over time. Findings from each of these analyses are presented in turn below. The study concludes with recommendations for the topics for which the AHS can reliably be used, subjects that are problematic given the nature of the AHS data, areas where further investigation is needed to explore the potential usefulness of the AHS, and options for improving the quality of the mortgage-related variables in the survey instrument.

Replication Analysis

The principal benchmark data sets used in the replication analysis are the FHA administrative data, data collected under the Home Mortgage Disclose Act (HMDA), the Mortgage Interest Rate Survey (MIRS), the Residential Finance Survey (RFS), the Survey of Consumer Finance (SCF) and a homebuyer survey conducted by Chicago Title & Trust Company.

Our benchmarking focuses on the AHS national sample data collected in the 1990s for all primary mortgages of owner-occupied single-family units, defined as structures with one to four units. Since many of the benchmark data sets provide information on new mortgage originations, much of the replication analysis focuses on estimates of the volume and characteristics of newly originated mortgages. Two characteristics of the AHS impose significant constraints on how this analysis can be conducted. First, since the AHS only gathers information on the year in which a mortgage is originated and not the month, analysis of originations can only be done on a calendar year basis. Second, since the AHS is conducted over the second half of the calendar year, it does not provide complete information on the year in which the survey is conducted. Given these two constraints, new originations are defined as loans made in the two complete calendar years prior to the year in which the AHS is conducted. For example, the 1997 survey data are used to study the characteristic and volume of loans originated in 1995 and 1996. An important concern about this approach is that by the time the AHS survey is conducted, some share of the new originations will have terminated and thus no longer be reported in the AHS. To account for the number of loans that have terminated prior to the interview in each cohort, we have devised a method to adjust all the AHS volume estimates based on the historical termination rates derived from an independent source.

Among the variables we have evaluated are mortgage origination volume, mortgage market segment, original loan amount, mortgage interest rate of fixed-rate loans, loan-to-value (LTV) ratio at origination, mortgage payment product type, first-time homebuyer status, borrower race/ethnicity, and income. The selection was based on the importance of these variables in mortgage market research and the availability of reliable benchmark measures.

To the extent possible, comparisons to the benchmark estimates are stratified by loan purpose (home purchase vs. refinance) and mortgage market segment (FHA, VA, FmHA/RHS, conventional conforming, and conventional jumbo).

Major findings of the replication analysis are summarized in Exhibit 0-1 at the end of the Executive Summary.

One of the findings of this study is the importance of recognizing the sampling variability of estimates obtained from the AHS. We found many cases where the small sample sizes of particular segments of the mortgage market examined resulted in very wide confidence intervals surrounding individual point estimates. Researchers using the AHS for mortgage market analysis should take sampling variability into account by reporting confidence intervals in addition to the AHS point estimates. When the data are used to support loan-level econometric analysis, we recommend that statistical software packages such as STATA[®] and SUDAAN[®] be employed to address fully the sampling errors in such analysis.

Internal Consistency of AHS Mortgage Variables Across Surveys

One of the advantages of the AHS is that it not only provides information on newly originated mortgages but also allows researchers to examine the characteristics of the entire stock of outstanding mortgages and changes in the use of mortgage finance over an owner's tenure in a unit. However, given that the AHS asks owners to report on mortgage characteristics each time the survey is conducted, there is a concern about whether the accuracy of owner's responses might degrade as the length of time since the mortgage was originated increases. In order for analysis of the characteristics of the entire stock of mortgages to be valid, homeowners residing in the sample housing units must report consistent information about their mortgages across survey years. The analysis of the internal consistency of the AHS responses is meant to investigate this concern.

The variables we have examined in this analysis are mortgage market segment, loan size, mortgage payment product type, loan term, interest rate, monthly payment amount, and first-time buyer status. We found that, for most variables, between 80 and 90 percent of the respondents report mortgage information consistently over time in the AHS surveys. As expected, the number of responses that are consistent with the base year survey tends to decrease in the later survey years regardless of the variable examined. Among the variables and survey years we examined, first-time homebuyer status and mortgage term were reported most consistently, with 90 percent of responses consistent over time. Most homeowners also provided consistent answers on questions about their loan's mortgage insurance type and payment plan type. Information on origination amount and monthly payment amount were less reliable, with consistency rates of 65 to 70 percent several years following mortgage origination. Interest rate information was by far the least reliable over time, with only 40 percent of responses consistent with initial answers eight years after origination. Of note, these last three variables all relate to the mortgage payment type: that is, whether the loan is fixed or adjustable rate. The relatively low consistency rates in these variables may reflect errors in reporting loans as fixed rate when they were actually adjustable rate.

Reliable AHS Mortgage Variables

We conclude that the following AHS variables and AHS-based estimates are reliable and can be used for mortgage market analysis, subject to the caveats described in each section:

- **Mortgage Originations for Home-Purchase Loans.** Using the adjustments and timeframes described in this report, the AHS can be used to support analyses of the total volume of recent originations of loans for home purchases.

Since only surviving loans are observed in the AHS, it is necessary to adjust volume estimates upward to account for the number of loans that terminated prior to the interview. We recommend that all AHS volume estimates should be adjusted upward using the termination rates of the market segment to which the mortgage belongs, when that is feasible. Because of the limited data availability for this study of termination rates for other market

segments, the analysis in the report used FHA historical termination rates to adjust the volume of loans in all segments of the market.¹

The AHS is best used to estimate volumes of recent originations, but not of originations that occur in the same year as the survey. Because of the timing of AHS interviews, spanning the last quarter of the year, the AHS will miss a large number of loans originated during the survey year. Therefore, information collected in the AHS should be used to examine the volume and loan characteristics of mortgages originated in the two years prior to the interview year. For example, the 1997 survey should be used to derive volume estimates for loans originated in 1995 and 1996.

- **Volume and Share by Market Segment.** Starting with 1993 originations, researchers can feel comfortable using the AHS to examine mortgage activity by market segment using the following division: FHA, VA/FmHA/RHS, and conventional. Volumes and share estimates for these separate market segments were found to be reliable. The AHS slightly undercounts the share of jumbo loans. It is possible that some borrowers in the conventional segment of the market underestimate their original mortgage amounts. An alternative explanation is that homeowners with jumbo mortgages are underrepresented in the survey as a result of non-participation. Given the fact that participation in the AHS survey commits one to repeated long interviews, more wealthy homeowners may be less willing to participate. It is also likely that, in a declining interest rate environment, those with jumbo loans are particularly quick to refinance their mortgages multiple times, which would not be captured in the AHS.

The AHS can be used to support research on mortgage activity by market segment as long as all conventional loans are grouped together.

- **Original Mortgage Amount.** Original loan amounts for mortgages in the FHA segment generally are accurately reported. There is a slight underestimate of the mortgage amounts for borrowers in the VA/FmHA and conventional conforming segments. Users should consider applying an adjustment factor to inflate the loan amounts for such mortgages. In addition, the loan amount variable provided in the public-use file is top-coded for confidentiality reasons. Since the top-coded value has been above the conventional conforming loan limit, the truncation of the loan amount primarily affects analysis of jumbo loans. Therefore, the AHS cannot be used to evaluate jumbo loan amounts.
- **Interest Rate.** Interest rates reported by borrowers in the AHS track the benchmarks for home-purchase fixed-rate mortgages in the FHA, conventional conforming, and conventional

¹ Ideally, the loan volume estimates from the AHS should be adjusted using historical termination rates from the corresponding market segments and mortgage product types, since default and prepayment speeds are likely to be different across market segments. Borrowers with conventional mortgages are more likely to refinance their mortgages than their FHA counterparts, given the same interest rate environment. On the other hand, FHA loans have higher claim rates. We expect that the termination rates for FHA loans can serve as a reasonable proxy for the other market segments because, on average, the low FHA prepayment rates should be offset by the high claim rates of FHA loans. The proxy adjustment rate is probably most problematic in a survey year such as 2001 with big refinancing boom.

jumbo segments of the market. Because of the lack of benchmark information, we have not been able to verify the reliability of interest rates for non-fixed-rate loan types reported in the AHS.

- **First-Time Homebuyer.** The AHS is generally consistent with the RFS and Chicago Title Surveys regarding the share of borrowers who are first-time homeowners. The AHS does consistently report a lower share of first-time buyers than the FHA administrative data. But the cause of this discrepancy is most likely a difference in how this information is defined and solicited in the two data sets. While AHS asks whether any household members have previously owned a home, the FHA mortgage application form inquires whether the buyer has owned a home in the past three years. The AHS definition and estimates are thus more consistent with a strict notion of first-time homebuyers.
- **Borrower Race/Ethnicity.** Racial and ethnic characteristics reported in the AHS track the benchmark database estimates reasonably well, although there is a tendency to report a slightly higher market share for African Americans and a lower share for Hispanics. These discrepancies most likely result from differences between the AHS and the administrative data of FHA and HMDA both in definitions of race/ethnicity categories and in the manner in which the race/ethnicity information is gathered.

Unreliable or Uncertain AHS Mortgage Variables

Users should avoid using the AHS (or use caution in some cases) for mortgage market research that would require the following:

- **Refinances.** Refinance originations are not captured adequately in the AHS. Across most of the survey years and all mortgage market segments, the AHS consistently undercounts the number of refinance loans originated. Surprisingly, the loan purpose questions (REFI variable) added in the 2001 survey did not provide any noticeable improvement in this underreporting problem. Researchers thus should not use the surveys to produce volume estimates of such mortgages and should not attempt other analysis for which accurate identification of the market share of refinance loans is critical.

However, the distributions of certain loan characteristics such as mortgage market segment, original mortgage amount, and interest rate among refinances that are captured in the AHS appear to be consistent with the benchmarks for most of the years. Therefore, users may be able to use the data for these loans for analysis pertaining to such loan characteristics. In particular, these records may be used as micro data to conduct regression modeling and other loan-level analyses.

The reason for this underreporting is unknown. One possible explanation is that the questionnaire does not provide a clear guidance to the homeowners regarding the distinction between lump-sum home equity loans and mortgages that are used to refinance an existing loan in order to take out housing equity (so-called cash-out refinances). As a result, some

survey respondents might have incorrectly classified their refinance mortgages as home equity loans.

The mediocre performance of the loan purpose question (REFI variable) in the 2001 survey could be because 2001 was such a big refinancing year that the proxy adjustment rates have failed to account for many of the loans that had terminated prior to the interview. In particular, some borrowers might be able to refinance their mortgages multiple times. The REFI variable may perform better in future survey years.

- **Jumbo Loans.** The AHS generally understates the share of mortgages that exceed the conforming loan limit, although it is fairly reliable at identifying conventional loans overall. Possible explanations for this underreporting are offered in the previous section titled “Volume and Share by Market Segment.”
- **Payment Product Types.** Payment product types are not identified accurately in the AHS. In general, ARMs are underreported, and the survey instrument is incapable of identifying borrowers with more sophisticated payment product types such as hybrid ARMs and two-step mortgages. The share of fixed-rate mortgages is generally over estimated, although when measured as a percent of the overall share, the magnitude of the discrepancy is smaller than for ARMs. Thus, the AHS should not be used for analyses of mortgage product choice, and analysts should use caution when including mortgages other than fixed rate mortgages in other types of analyses.
- **Loan-to-Value (LTV) Ratios.** Although average LTV ratios track well the benchmark estimates for both FHA and conventional loans, the distributions of the ratios do not track as well. Of particular note for policy makers, it appears that the AHS estimates of the share of mortgages with very high LTVs may be overstated. This may result from differences in the way house values are recorded in different databases. Discrepancies may also be caused by the extreme values of the loan amount and house value variables that were inaccurately reported in the AHS. Further investigation is needed. Thus, researchers should be cautious in their use of LTV estimates based on the AHS.
- **Borrower Income.** Borrowers’ income in AHS does not track the values reported in HMDA and the FHA data well, mainly because the AHS variable includes different components of income. The AHS measure is more complete and thus yields estimates that are consistently higher. On the other hand, when incomes captured in the Current Population Survey (CPS) and Social Security Administration data are used as benchmarks, studies by the Census Bureau have shown that the AHS respondents have a tendency to underreport incomes. Researchers need to make a judgment independent of the benchmarking to HMDA and FHA data conducted for this study on whether income as identified by the AHS is an appropriate measure of income for their mortgage market research questions.

Areas Requiring Further Investigation

- The substantial underreporting of refinance mortgages in the AHS deserves further examination. Given the longitudinal nature of the AHS survey design, linked AHS files might be used to explore whether refinances can be identified more accurately from AHS panel data. Looking at refinances from a panel setting might also shed new light on the structural cause of underreporting of such mortgages and suggest modifications that could be made to the survey. In particular, responses linked across surveys could be used to examine when home equity loans were originated to determine if this method might be able to provide a more complete accounting of refinance loans. For example, this method might be able to identify cases where a previous mortgage was retired and a new home equity loan was originated, which could be counted as a refinance.
- For longitudinal analysis, additional research is needed to examine whether there are systematic ways to delete problematic cases from the AHS files to produce a core of internally consistent records with mortgage variables that are reliable. Would the remaining sample be too small to conduct any meaningful analysis? Would deletion of such cases distort the composition of the remaining sample relative to the benchmarks?
- LTVs are another area where further study is needed before determining that the AHS is incapable of supporting LTV research. We suspect that some of the discrepancies in the LTV estimates between the AHS and other mortgage market data sets are caused by outliers in loan amounts and house values reported to the AHS. The distribution of loans by LTV category might be usable if problematic outlier cases were removed by the analyst.
- The micro data for the new round (2001) of RFS will be available to researchers in the summer of 2003. Extensive benchmarking of the 2001 AHS to the new RFS data should be undertaken to gain insight into the impact of post-1995 AHS survey changes on the accuracy of the AHS data. The RFS can be used to test mortgage market measures of recent cohorts as well as the entire stock of mortgages that were in place in 2001. It will be especially valuable if the internal version of the RFS file can be used. The RFS is particularly valuable as a point of comparison since information on mortgages is collected directly from lenders. These data elements should be quite accurate, since the lenders' information is likely supported by administrative records (rather than the homeowners' memory as in data based only on household surveys).

Recommendations for Modifications to the Survey

Based on the research conducted for this study, we have identified several potential modifications to the AHS that would make it substantially more useful for research on mortgage market issues:

- Questions should be added to identify the number of refinances that have taken place during the interval since the last survey. The home equity questions also should be changed to ask

about lines of credits rather than “home equity loans,” since there is no real distinction between lump-sum home equity loans and standard mortgages.

- In the current AHS questionnaire, for homeowners who obtained the mortgage the same year the home was purchased, the survey does not obtain the year of loan origination from the respondents; the origination year information can only be retrieved from the WHNGET variable in the “purchase” module of the AHS survey, which reports the year when the house was obtained. For other homeowners, the survey asks explicitly what year the owner obtained the mortgage and the information is stored in the YRMOR variable. HUD and the Census Bureau should consider asking *all* homeowners the month and year of mortgage origination explicitly and storing the information in one variable.
- HUD and the Census Bureau should evaluate changes to specific questions that might elicit more accurate reporting of mortgage payment product types. One of the issues to be considered is the cost-effectiveness of making such changes, given the rapid evolution of mortgage products.
- New questions should also be considered to collect information on mortgage payment status (delinquency and default) and the termination of mortgages that were previously in place.
- HUD and the Census Bureau should consider asking an explicit question on the unpaid principal balance (UPB). UPB is useful for estimating the current home equity and LTV of each homeowner. Currently, UPB is not collected in the AHS. Users can calculate an estimate using information pertaining to interest rate, original amount and mortgage term. Given the lack of details on adjustment terms for non-fixed rate loans and the instability of the interest rates and mortgage amounts reported across waves of the survey, this method is not likely to yield reliable UPB estimates. Although it is likely that some borrowers may not be able to report the UPB accurately if asked in the survey, the self-reported estimates can nonetheless provide an internal validity check against the user-calculated amounts.
- Given our findings on the instability of several key variables over time, we recommend that the “dependent interviewing” technique be extended to include mortgage insurance type, payment plan, interest rate (of fixed-rate loans), and principal payment amount. This technique is used to avoid repeating questions to the household if the answer should not have changed since the previous survey. For example, the first-time buyer status question was not asked in the 1999 survey if the same household occupied the housing unit and a valid answer had been obtained in a previous survey.

**Exhibit 0-1:
Major Findings for AHS Mortgage Market Measures Tested in the Replication Analysis**

MORTGAGE MARKET MEASURE	BENCHMARK DATA & ORIGINATION COHORTS EXAMINED	STRATIFICATION	MAJOR FINDINGS
Mortgage Origination Volume	<ul style="list-style-type: none"> • FHA: 1989-2000 for FHA • HMDA: 1993-2000 for VA/RHS, Conventional Conforming and Conventional Jumbo 	<ul style="list-style-type: none"> • Mortgage Market Segment^a • Loan Purpose^b 	<p>Home-purchase loans:</p> <ul style="list-style-type: none"> • Generally accurate for volume estimates of origination cohorts. Over the 1993 to 2000 period, the average coverage ratio was 99 percent for all loans combined (Exhibits 4-3, 4-4, and 4-5). • Coverage ratio declined for the 1998 to 2000 cohorts. Average coverage ratio was in the 83-89 percent range for all loans combined during that period (Exhibit 4-4). • Relatively low volume estimates for conventional jumbo originations (Exhibit 4-5). <p>Refinance loans:</p> <ul style="list-style-type: none"> • Low volume estimates across all mortgage market segments. Average coverage ratio was 48 percent for 1993 to 2000 (Exhibit 4-7, 4-8, and 4-9). • Especially low coverage ratio for cohorts since 1993 (1995 survey) probably because of question wording (a lack of clear distinction between cash-out refinance mortgages and lump-sum home equity loans) or coding errors in the survey (Exhibit 4-8). • Loan purpose questions added in the 2001 survey did not significantly improve the survey's ability to capture refinance loans (Exhibit 4-8).

MORTGAGE MARKET MEASURE	BENCHMARK DATA & ORIGINATION COHORTS EXAMINED	STRATIFICATION	MAJOR FINDINGS
<i>Loan Characteristics</i>			
Mortgage Market Segment	<ul style="list-style-type: none"> • RFS: 1989-1991 • HMDA: 1993-2000 • SCF: 1993-1994 	<ul style="list-style-type: none"> • Loan Purpose 	<p>Home-purchase loans:</p> <ul style="list-style-type: none"> • For loans originated in 1989-1991, slight underestimate of FHA share and overestimate of conventional conforming share (Exhibit 4-10). • Accurate market shares for loans in the FHA, VA/RHS and conventional segments, 1993-2000. Discrepancies are mostly in the 1 to 2 percent range (Exhibit 4-11). • Slight underestimate of market share of conventional jumbo loans and overestimate of conventional conforming loans for most of the cohorts (Exhibit 4-11). <p>Refinance loans:</p> <ul style="list-style-type: none"> • Small sample sizes result in wide confidence intervals that make it difficult to draw statistically reliable comparisons. • Accurate shares for loans in the 1989-1991 cohorts (Exhibit 4-10). • Underestimate of shares for conventional jumbo loans for most cohorts (Exhibit 4-12).
Origination Loan Amount	<ul style="list-style-type: none"> • FHA: 1989-2000 for FHA • HMDA: 1993-2000 for VA/RHS, Conventional Conforming 	<ul style="list-style-type: none"> • Mortgage Market Segment • Loan Purpose 	<p>Home-purchase loans:</p> <ul style="list-style-type: none"> • Accurate for loans in the FHA market segment (Exhibit 4-13). • Slight underestimate, by about 10 percent, of loan amounts in the VA/RHS segment (Exhibit 4-15). • Accurate for 1995, 1997, 1998, 1999 and 2000 conventional conforming loans. Underestimate, by about 10 percent, of those in 1993, 1994 and 1996 cohorts (Exhibit 4-15). <p>Refinance loans:</p> <ul style="list-style-type: none"> • Underestimate of loan amounts of FHA loans originated in 1995 and 1996 (Exhibit 4-14). • Small sample sizes for VA/RHS loans (Exhibit 4-16). • Accurate for 1995, 1997 and 1998 conventional conforming loans. Underestimate for 1993, 1994 and 1996 cohorts. Overestimate for 1999 and 2000 cohorts (Exhibit 4-16).

MORTGAGE MARKET MEASURE	BENCHMARK DATA & ORIGINATION COHORTS EXAMINED	STRATIFICATION	MAJOR FINDINGS
Mortgage Interest Rate (Fixed-Rate Loans)	<ul style="list-style-type: none"> FHA: 1989-2000 for FHA MIRS: 1992-2000 for Conventional Conforming and Conventional Jumbo Home Purchases 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose Payment Product Type^c 	<ul style="list-style-type: none"> Accurate averages for both home-purchase and refinance loans in segments of the market examined. Discrepancy in the averages is generally one to two tenths of a percent. Overall distribution appears to be more dispersed (Exhibits 4-17 and 4-18). Compared to MIRS, the AHS reports a higher proportion of conventional conforming purchases with relatively high interest rates, probably because subprime loans are included in the AHS but not in MIRS (Exhibits 4-19 and 4-20).
LTV at Origination	<ul style="list-style-type: none"> FHA: 1989-2000 for FHA Home Purchases MIRS: 1992-2000 for Conventional Conforming Home Purchases 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> Reasonably accurate mean LTVs by cohort for FHA home purchases. Distribution of loans by LTV categories does not track well with benchmarks (Exhibits 4-21 and 4-22). Reasonably accurate mean LTVs by cohort for conventional conforming purchases. Distribution of loans by LTV categories does not track well with benchmarks (Exhibits 4-23 and 4-24). LTVs in the other segments of the mortgage market cannot be tested because of the lack of comparable benchmarks.
Mortgage Payment Product Type	<ul style="list-style-type: none"> FHA: 1989-2000 for FHA MIRS: 1992-2000 for Conventional Conforming and Conventional Jumbo Home Purchases 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> Understates ARM shares for FHA loans (Exhibits 4-25 and 4-26). Overstates shares of loans other than fixed-rate and ARMs in the FHA segment (Exhibits 4-25 and 4-26). Understates ARM shares for conforming and jumbo conventional home purchases (Exhibits 4-27 and 4-28).
<i>Borrower Characteristics</i>			
First-time Homebuyer Status	<ul style="list-style-type: none"> Chicago Title: 1995-1999 RFS: 1989-1991 FHA: 1991-2000 for FHA 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> Estimate of first-time buyer shares for all home purchases consistent with RFS and Chicago Title data (Exhibits 4-29 and 4-30). While differences with RFS are not statistically significant, the share of first-time buyers in AHS is lower for FHA and conventional jumbo borrowers but about equal for VA/RHS and conventional conforming (Exhibit 4-30). Discrepancies between the AHS and FHA data are due to how first-time buyer status is defined. The AHS estimates are more consistent with a strict notion of first-time buyers.

MORTGAGE MARKET MEASURE	BENCHMARK DATA & ORIGINATION COHORTS EXAMINED	STRATIFICATION	MAJOR FINDINGS
Borrower Race/Ethnicity	<ul style="list-style-type: none"> FHA: 1989-2000 for FHA HMDA: 1993-2000 for VA/RHS, Conventional Conforming and Conventional Jumbo 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> Different race/ethnicity questions and method of gathering information between AHS and administrative benchmark data sets make interpretation of differences difficult. Overall, distributions are reasonably close to the benchmarks (Exhibits 4-32 to 4-40).
Borrower Income	<ul style="list-style-type: none"> HMDA: 1993-2000 for FHA, VA/RHS, Conventional Conforming and Conventional Jumbo 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> Incomes are defined differently between AHS and benchmark data, making it difficult to interpret discrepancies. Household incomes in AHS are consistently higher than in HMDA, probably because in HMDA some borrowers do not need to report all income components to lenders to qualify for loans (Exhibits 4-41 to 4-44).

Notes: ^aMortgage market segments are FHA, VA/FmHA/RHS, conventional conforming, and conventional jumbo.

^bLoan purposes are home purchase and refinance mortgages.

^cPayment product types are fixed-rate and adjustable rate mortgages.

Chapter One

Introduction

This research was conducted under Task Order 7 of HUD Contract C-OPC-18571 for an “Analysis of Housing Finance Issues Using the American Housing Survey (AHS).” The purpose of the study is to evaluate the quality of mortgage data in the AHS for analyzing issues of interest to policy makers and the mortgage industry. To test the reliability of the AHS variables we replicate measures of mortgage activity from other reliable sources of data that serve as benchmarks for the AHS estimates. We also use the longitudinal nature of the AHS to determine whether answers to questions on mortgages are consistent and stable over time. The goals of the analysis are to determine: 1) what types of mortgage market analysis can be supported by the AHS; 2) what areas of the AHS are problematic for mortgage research; and 3) what techniques or changes in the survey could potentially compensate for the problems.

Background and Policy Context

The American Housing Survey (AHS, formerly Annual Housing Survey) is the most comprehensive source of information about the characteristics and condition of the nation’s housing stock. The AHS consists of two components, a national sample and a series of metropolitan area samples. Started in 1973, the AHS national sample data were collected by the U.S. Census Bureau for the Department of Housing and Urban Development (HUD) on a nationally representative sample of housing units (about 55,000 homes) every year until 1981, and they have been collected every other year since then. The survey has contained the same basic sample of housing units since 1985, with units added to each wave of the survey to reflect additions to the housing stock. As for the Metropolitan Statistical Area (MSA) samples, surveys for 46 selected MSAs are collected every 4 to 6 years, with an average of 12 MSAs included in each survey year. Each of the metropolitan area samples covers about 3,000 housing units. The metropolitan area survey includes all of the questions from the national survey, with some additional questions related to commuting patterns and the location of previous residence.

Over the years, AHS data have been used by researchers and policy analysts to answer questions about the nation’s housing conditions (e.g., unit quality, unit and building size, repairs, structure age and neighborhood quality) and occupant characteristics (e.g., tenure choice, race, income, and housing costs). The longitudinal nature of the AHS also permits the analysis of trends for certain housing and occupancy characteristics of the housing stock.

The AHS data contain detailed questions about mortgages asked of respondents for owner-occupied units in the survey. The questions cover most basic mortgage and housing finance topics, including loan amount, mortgage product type (fixed rate versus adjustable rate), mortgage term, contract interest rate, year of origination, type of mortgage insurance, and whether the owner is a first-time or repeat home owner. This wealth of mortgage-related variables, combined with the tenant demographic and property location information, could be a very powerful resource for answering many housing finance research and policy questions. These micro-data could provide crucial information to support HUD’s regulatory and programmatic responsibilities, which include regulating

the government-sponsored-enterprises (GSEs), evaluating the role of FHA in the mortgage market, and setting performance standards for HUD programs.

However, neither the housing research community nor HUD staff has made as much use of the mortgage variables of the AHS data as might be expected. Among the reasons for this under-utilization is the fact that the reliability of these mortgage-related variables in the AHS has not been verified. Analysis is needed to establish the extent to which limitations associated with sample size, survey design, and interview response affect the accuracy and consistency of the mortgage data in the AHS. The research presented in this study is intended to meet this need.

Outline of the Report

The rest of the report is organized as follows:

- **Chapter Two** presents an overview of the AHS, a detailed description of the mortgage-related variables, changes in the survey over time that may have affected the mortgage variables, a summary of the potential advantages and disadvantages of using the AHS for mortgage-related research, and the set of variables that were tested in the replication analysis.
- **Chapter Three** describes the mortgage market data chosen to provide benchmark estimates for the AHS variables, highlighting the strengths and weaknesses of each database. The chapter provides the rationale for our choice of each data set used in the replication analysis.
- **Chapter Four** presents the findings of the replication analysis. The AHS estimates are compared to the independent benchmarks derived from the comparison databases. The variables we have examined are loan volume, loan purpose, mortgage market segment, loan amount, contract interest rate, loan-to-value (LTV) ratio, mortgage payment product type, first-time buyer status, borrower race/ethnicity, and income.
- **Chapter Five** evaluates the internal consistency and stability of the AHS mortgage data elements reported across time. Sample housing units are linked across different survey years. This investigation focuses on a few key mortgage variables.
- **Chapter Six** synthesizes the findings. It serves as a guide for HUD staff and other researchers on the potential use of the AHS variables for different types of research on mortgages and housing finance. Topics related to the reliability of AHS mortgage variables that require additional investigation are described.
- **Appendix A** documents the GSE conforming loan limits used to distinguish between conforming and jumbo conventional mortgages.
- **Appendix B** presents detailed information about the comparison databases.

- **Appendix C** discusses alternative methods that can be used to account for mortgage activity between survey waves that is not captured by the AHS.
- **Appendix D** provides additional tabulations to support the analyses presented in Chapter Four. It includes some additional estimates that may be of interest but are not discussed in the text.
- **Appendix E** describes the steps taken to link observations on mortgage activity from multiple years of AHS surveys.
- **Appendix F** presents additional tabulations to support the analyses presented in Chapter Five.
- **Appendix G** provides the text of the questions asked in the “mortgage” and “purchase” modules of the 1997 AHS.
- **Appendix H** presents a user’s guide for deriving the mortgage market attributes in the AHS that we have examined. For each mortgage market measure, it identifies the AHS variable(s), as well as the computations and selection criteria involved in the derivation. This information differs across survey years for some of the variables.

Chapter Two

The AHS and Research on Mortgages

This chapter examines the characteristics of the AHS that make it a potentially useful source of information to support analysis of issues related to residential mortgage finance. The first section provides a brief description of the history and basic characteristics of the survey. The next section describes the variables that are collected by the AHS that relate to mortgage finance. The third section describes changes over time to aspects of the survey that potentially affect the use of the data for mortgage market analysis. The fourth section summarizes the advantages and disadvantages of the AHS for mortgage related research. Specific mortgage variables to be investigated in the replication analysis are presented in the final section.

History and Basic Characteristics of the AHS

Besides the decennial census, the American Housing Survey contains the most comprehensive information available on the nation's housing stock. The AHS consists of two components, a national sample of housing units and a series of metropolitan area samples. Begun in 1973 as the Annual Housing Survey, the national AHS survey was conducted annually through 1981. Since 1983 the national survey has been conducted every other year and has been known as the American Housing Survey. The AHS is a panel survey of about 55,000 units that collects information on the same housing units in each wave. A new sample was drawn for the 1985 national survey based on the 1980 decennial census. The same basic sample of housing units has been followed since 1985, with units added to each wave of the survey to reflect additions to the housing stock. Each unit surveyed is identified by a unique control number which allows the linking of data for individual housing units from different survey years to analyze changes in housing conditions over time.

The metropolitan area surveys are conducted in 46 selected MSAs. Each area is surveyed every 4 to 6 years, with an average of 12 MSAs included in each survey year. Each of the metropolitan area samples covers 3,000 or more housing units. The metropolitan area survey includes all of the same questions from the national survey with some additional questions related to commuting patterns and the location of previous residence. Like the national survey, the metropolitan area samples include a panel of housing units surveyed each time supplemented by additional units that reflect new construction. From 1974 to 1994 the sample for the metropolitan area surveys was based on the 1970 census. All but six metropolitan area samples were redrawn in 1995 based on the 1990 census. Beginning in 1995 six of the largest metropolitan areas were covered as part of the national survey so, like the national survey, the sample for these areas is based on the 1980 census.² As a result of these changes in the metropolitan area samples, analysis that relies on the panel nature of the AHS metro surveys must either focus on the 1973 to 1994 period or the 1995 period and later.

² The metropolitan areas that are included as part of the national survey include Chicago, IL PMSA, Detroit, MI PMSA, Los Angeles-Long Beach, CA PMSA, New York-Nassau-Orange County, NY PMSA, Northern New Jersey PMSA, and Philadelphia, PA-NJ, PMSA.

The AHS covers a broad range of topics, including:

- The general characteristics of the structure (size, age, type, tenure, lot size, parking, number and type of rooms);
- The equipment used for appliances, heating and cooling;
- The quality of the unit in terms of interior and exterior conditions and reliability of equipment;
- Recent renovations to the unit;
- Neighborhood characteristics;
- All housing related costs (including rent, mortgage characteristics and costs, utilities and services, insurance, and property taxes) and the use of housing subsidies;
- Household composition including the age, sex, and race of all household members and their relationship to the respondent;
- Income for each household member; and
- The previous residence of household members and the reasons for moving.

The broad and detailed set of information covered by the survey supports a wide range of housing research. The AHS is a particularly important source of information on the cost and quality of housing over time and the housing characteristics of different demographic groups.

Mortgage-Related Variables in the AHS

In each AHS survey, a series of housing finance and mortgage-related questions is posed to respondents residing in owner-occupied housing units.³ Although many of the questions have remained the same over the years, a few have changed. For instance, questions related to home equity loans have been added since the 1995 survey, while questions concerning the existence of reverse mortgages were added in 1997 and then dropped from the 1999 survey. Appendix G provides the text of the questions asked in the “mortgage” and “purchase” modules of the 1997 AHS.⁴ Only minor changes have happened in these modules in the subsequent two survey years. The following description in this section applies to the questionnaire of survey years that we examined in this study. (A thorough discussion of relevant changes in the survey questions or methodology is presented in a later section of this chapter.)

³ These questions are grouped under the “Mortgage” module in the *Codebook for the American Housing Survey, Volume 3: 1997 SAS Files and Questionnaire*, draft, February 16, 1998.

⁴ Text of questions for the 1999 and 2001 surveys cannot be included here because the Census Bureau has not published them.

The mortgage module of the survey begins by identifying how many mortgages are on the property.⁵ There is some attempt to distinguish between regular mortgages and home equity loans. The respondent is asked explicitly how many mortgages are on the property excluding home equity loans as well as whether there is a home equity loan in place. However, no clear distinction is made between lump-sum home equity loans and mortgages that are used to refinance an existing loan in order to take out housing equity (so-called cash-out refinances). The owner is then asked whether a mortgage was obtained through a State or local government program that provides lower cost mortgages, although the survey does not link this question to a specific mortgage.

Following these introductory questions, the survey asks a series of questions on the first mortgage on the property. If there is more than one mortgage, the survey then asks the same set of questions for the second mortgage. Exhibit 2-1 identifies the variables that are gathered in this series of questions and describes questions that are used to elicit this information. As the exhibit notes, some of the questions are worded indirectly as an attempt to elicit more reliable and consistent answers. For example, the questionnaire does not ask directly whether the borrower has an adjustable rate mortgage (ARM) loan; rather, the respondents are asked whether the payments remain the same over the life of the loan and, if they do not, why they change.

Exhibit 2-1
Characteristics of First and Second Mortgages Collected in the AHS

VARIABLE	SOURCE QUESTIONS FOR THE VARIABLE
Year Mortgage Originated	The survey first asks whether the owner got the mortgage the same year the home was purchased. ⁶ If not, it then asks explicitly what year the owner obtained the mortgage.
New or Assumed Mortgage	The survey asks whether the owner got a new mortgage or assumed someone else's mortgage.
Term of Mortgage	For new mortgages, the survey asks how many years the mortgage was for when it was acquired. If the mortgage is assumed, the survey asks how many years remained on the mortgage when assumed.
Amortization Period	If the reported term is less than 15 years, the respondent is asked how many years it would take to pay off the loan, given the current monthly payments. This question is intended to elicit the amortization period used to estimate payments on balloon loans. If the loan term is 15 years or more, the amortization and term are assumed to be the same.

⁵ In the 1997 survey, the module begins by asking the respondents whether there is a reverse mortgage on the property. If there is, the survey assumes there is no other mortgage in place. If there is no reverse mortgage, questions for the mortgage module continue.

⁶ For homeowners who obtained the mortgage the same year the home was purchased, the survey does not obtain the year of loan origination from the respondents; this information can be retrieved only from the WHNGET variable in the "purchase" module of the survey, which reports the year the house was obtained. For other homeowners, the YRMOR variable contains the year of mortgage origination.

VARIABLE	SOURCE QUESTIONS FOR THE VARIABLE
Amount of Mortgage	For new mortgages, the survey asks how much was borrowed. If the mortgage is assumed, the question is how much was left to pay off when it was assumed. The respondent is asked whether the mortgage also covers other homes or apartments, farm land, or a business on the property. If so, the respondent is then asked how much of the mortgage applies to the home in the survey.
Current Interest Rate	The survey asks for the current interest rate rounded down to the nearest ¼ percent.
Current Monthly Payment	The survey asks for the current monthly payment. Separate questions are then asked to determine if the payment includes property taxes, homeowner insurance, or “anything else” in addition to principal and interest. The survey taker is instructed that “anything else” may include payments for disability or life insurance. If “anything else” is reported to be part of the monthly payment, the survey asks for the annual amount of these payments. In the 2001 survey, a question is added to determine whether the mortgage payment includes private mortgage insurance (PMI). If it does, the survey inquires how much the PMI payment is.
Government-Insured Mortgage	The survey asks if the mortgage is an FHA, VA, FmHA or some other type.
Financing Provided by Seller or Other Individual	The survey asks if the loan was provided by a bank or other organization or by an individual. If the loan was made by an individual, the survey then asks if the individual was the former owner of the property.
Mortgage Payment Features	The survey asks if the payments on the loan stay the same during the whole length of the mortgage. If not, the survey then asks how they change. Any of the following payment plan features that apply will be recorded: <ul style="list-style-type: none"> • Change in taxes or insurance, or due to decline in the principal balance; • Change based on interest rates; • Rise at fixed schedule during part of the loan term; • Rise at fixed schedule during the whole length of the loan; • Last payment is the biggest; or • Other. These payment features do not allow the AHS users to easily and unambiguously classify each loan into a mutually exclusive payment product type.

After gathering this detailed information on the first and second mortgages, the survey then inquires about the amount borrowed under all other mortgages and the monthly payments for these loans. The sum of the amounts borrowed and the total monthly payments are reported. Finally, the survey asks a series of questions about any home equity loans on the property. The same set of questions is asked for up to three home equity loans. Exhibit 2-2 summarizes the information gathered on home equity loans. Much less information is gathered about home equity loans than about mortgages. Perhaps most significantly, the survey does not ask when the home equity loan was originated.

Exhibit 2-2
Characteristics of Home Equity Loans Collected in the AHS

VARIABLE	SOURCE QUESTIONS FOR THE VARIABLE
Type of Home Equity Loan	The survey asks whether the owner is allowed to borrow against the home as often as he or she wishes up to a fixed amount (a line of credit) or if he or she obtained a one-time, lump sum payment that is repaid over a period of time.
Credit Limit	For line of credit loans, the survey asks for the total credit limit.
Lump Sum Amount	For lump sum loans, the survey asks for the amount of the lump sum.
Outstanding Balance	The survey asks for the current outstanding balance.
Current Interest Rate	The survey asks for the current interest rate rounded down to the nearest ¼ percent.
Monthly Payment	The survey asks for the amount of the last monthly payment.

In addition to the mortgage-specific variables, the AHS also contains other information relevant to the analysis of mortgages and housing finance. These include:⁷

- Amounts paid for other housing costs, including property taxes, homeowner’s insurance, homeowner association or condominium fees, and land rents;
- The year the home was purchased;
- The type of property (single-family, condominium, or mobile home);
- Sales price of the property when purchased;
- The owner-estimated current house value at the time of the interview;⁸
- Whether a down payment was used to purchase the home and, if so, the major source of the down payment: sale of previous home, sale of other investment, savings or cash on

⁷ Questions related to these data elements can be found in the “Housing Costs,” “Numbers of Units in Building,” “Rooms, Type of Building, Size, Vehicles,” and “Purchase, Inspection, Value, Year-round Use” modules in the *Codebook for the American Housing Survey, Volume 3: 1997 SAS Files and Questionnaire*, draft, February 16, 1998.

⁸ There is a sizable body of research literature testing the accuracy (against external data) and consistency (across survey years) of the owner-estimated house value in the AHS. For a review of the related research, see Katherine A. Kiel and Jeffrey E. Zabel, “The Accuracy of Owner-Provided House Values: The 1978-1991 American Housing Survey,” *Real Estate Economics*, summer 1999, v.27(2): 263-98; John L. Goodman and John B. Ittner, “The Accuracy of Home Owners’ Estimates of House Value,” *Journal of Housing Economics*, v.2(4): 339-57.

hand, borrowing other than a mortgage on the property, inheritance or gift, land where property was built was used for financing, or other; and

- Whether the owner previously owned a home or is a first-time homebuyer.

Finally, there are a number of mortgage-related characteristics that are not collected directly by the survey but can be derived by the users from the information that is collected. Exhibit 2-3 identifies the mortgage-related variables that can be derived from the AHS and indicates the process used to estimate these variables.

**Exhibit 2-3
Mortgage-Related Variables Derived from AHS Variables**

DERIVED VARIABLE	HOW IT IS DERIVED
Mortgage Market Segment	Whether the loan is government insured, conventional conforming or a jumbo can be derived from the type of mortgage (FHA, VA, FmHA/RHS, or other), the original mortgage amount, and the year of origination. ⁹
Loan Purpose	<p>Before the 2001 survey, home purchase vs. refinance can only be identified by comparing the mortgage origination year with the year of house purchase. Loans originated in years after the home was acquired are assumed to be refinancings.</p> <p>In the 2001 survey, refinances can be identified directly using the REFI variable. The survey also ask the reason for the refinance:</p> <ul style="list-style-type: none"> • Get a lower interest rate; • Increase the payment period for the mortgage; • Reduce the payment period for the mortgage; • Renew or extend a loan that has fallen due (without increasing the outstanding balance); • Receive cash or increase the outstanding balance of the loan; or • Other reason. <p>Then, the borrowers are asked how much cash they received as a result of the refinancing, and what percentage of the cash was used for additions, improvements or repairs to the house.</p>

⁹ “Conventional” refers to loans that are not insured or guaranteed by the government. “Conforming” refers to loans that are below the limits placed on residential loans that can be purchased or guaranteed by Fannie Mae and Freddie Mac. Jumbo loans are mortgages that exceed this limit. Since the loan limits are set for calendar years, the year of origination is needed to identify the appropriate limit. The loan limits are documented in Appendix A. One obstacle to accurately characterizing loans as conforming loans using the AHS is that the conforming loan limit is 50 percent higher in Hawaii and Alaska. The AHS does identify observations from the Honolulu MSA, which can be used to identify cases where this higher limit prevails. However, observations in Hawaii outside of Honolulu and in Alaska generally may not be identifiable. But since these areas represent a very small share of the U.S. mortgage market, this issue should not cause a significant distortion in the identification of the conforming loan market.

DERIVED VARIABLE	HOW IT IS DERIVED
Loan-to-Value Ratio at Origination	Ratio of original mortgage amount to purchase house price (or estimated home value if it appears to be a non-arms length transaction or a refinance loan).
Loan-to-Value Ratio Over Time	Ratio of estimated unpaid principal balance (UPB) to estimated home value. The Census Bureau staff has written a SAS computer program to calculate the estimated UPB based on ¹⁰ : <ul style="list-style-type: none"> • Original mortgage amount • Mortgage term • Interest rate • Mortgage Payment Plan • Number of years since origination Given the lack of details on adjustment terms for non-fixed rate loans, the estimated unpaid balance may be difficult to calculate accurately for these loans.
Mortgage Payment Product Type	Information collected on mortgage payment features can be consolidated by running a SAS computer program provided by the Census Bureau. ¹¹ The computer program classifies each mortgage into one the following seven mutually exclusive product types: <ul style="list-style-type: none"> • Fixed payment, self-amortizing mortgage • Adjustable rate mortgage (ARM) • Adjustable term mortgage • Graduated payment mortgage • Balloon mortgage • Combination of the above • Other
Housing Cost-to-Income Ratio	Ratio of payments for the mortgage plus property taxes, insurance, homeowner association fees, and land rents to household income.

Geographic Coverage of the Survey

Both because of concerns about spatial disparities in the availability and cost of mortgage financing and because of the need to control for market characteristics, geographic identifiers can be quite important for mortgage market research. The national AHS contains geographic identifiers for Census region (Northeast, Midwest, South, or West) and metropolitan location (central city, suburb,

¹⁰ The SAS code can be downloaded from the HUD User web site: http://www.huduser.org/datasets/ahs/1999table_recoded.txt. Another SAS program that performs the same computation for variables from the 1995 or older surveys can be obtained from Paul P. Harple of the Census Bureau.

¹¹ http://www.huduser.org/datasets/ahs/1999table_recoded.txt

or non-metro). The survey also identifies approximately 150 of the largest MSAs.¹² While the sample sizes for most of these MSAs are too small to provide reliable estimates for these markets, the MSA code can be helpful for more refined geographic analysis. For example, this variable can be used to identify the appropriate area median income to estimate household income as a percent of the area median that is used in a variety of housing programs as a determinant of program eligibility.¹³

The metropolitan area boundaries are based on the 1983 Office of Management & Budget (OMB) definition for all the AHS national sample data collected since 1984. Since MSA definitions are revised following each census, some of the MSA boundaries used in the AHS do not correspond to the most recent definitions used by most other current sources of data. One implication of this discrepancy is that the AHS definitions of metropolitan and non-metropolitan as well as central city and suburb do not always correspond to current standards.

Significant Changes to the AHS Over Time

This section describes significant changes to the AHS since the current national sample was redrawn in 1985 that may affect research on mortgages and housing finance. The section begins by discussing changes to the variables related to mortgages and then discusses changes to the survey methodology that may affect the survey results.

Changes in Information Collected

For the most part there have not been significant changes to the mortgage-related information collected in the AHS. The survey was essentially unchanged from 1985 through 1993. Since 1995 the following changes have been made:

Loan Purpose

In the 2001 survey, a series of questions were added to collect loan purpose information of the mortgage(s) and home equity loan(s) in place. Specifically, the borrower was asked explicitly whether the mortgage was a refinancing of a previous mortgage. With the addition of this question refinance loans can be identified without having to resort to a comparison between the mortgage origination year and house purchase year. However, since no guidance was provided to the homeowners regarding the distinction between “home equity loans” and a mortgage, the addition of this question did not address the issue of whether respondents are able to properly distinguish

¹² In the public-use version of the national files, the geographic areas identified on each record are metropolitan/non-metropolitan area, inside/outside central city, and urban/rural within an MSA. However, for confidentiality reasons, the following restrictions are employed with these variables: the central city portion of an MSA is identified only if the 1980 population of the central city was over 100,000; MSA is identified for urban portions of an MSA only if that area had a 1980 population of 100,000 or more; and MSA identification is not available for sample units located in the rural portions of a MSA.

¹³ Kathryn P. Nelson of HUD has developed a methodology for estimating area median income in areas outside of MSAs that relies on the Census Region and another geographic variable in the survey that identifies the long-term average degree days including heating and cooling degree days.

between lump-sum home equity loans and mortgages that are used to refinance an existing loan in order to take out housing equity (so-called cash-out refinances).

Next, additional questions were asked regarding the reason for the refinance – to get a lower interest rate, to increase the payment period for the mortgage, to reduce the payment period for the mortgage, to renew or extend a loan that has fallen due (without increasing the outstanding balance), to receive cash or increase the outstanding balance of the loan, or other reason. Borrowers were also asked how much cash they received as a result of the refinancing, and what percentage of the cash was used for additions, improvements or repairs to the house.

Private Mortgage Insurance

In the 2001 survey, homeowners were asked whether their monthly mortgage payment amount included private mortgage insurance (PMI). If it did, the survey inquired how much the payment for private mortgage insurance was for the year prior to the interview.

Home Equity Loans

In 1995 a question was added to ask owners whether they had a home equity loan on their property, but respondents are not offered guidance about how to distinguish a mortgage from a home equity loan. Furthermore, the 1995 survey did not gather any information on the characteristics of these loans and no distinction was made between regular mortgages, home equity loans and reverse mortgages. Beginning in 1997 the survey began gathering the more detailed information on home equity loans summarized in Exhibit 2-2.

Reverse Mortgages

In 1997, the survey asked borrowers a series of questions about whether they had a reverse mortgage on their property, including the amount borrowed, the interest rate, and whether the loan provided a lump sum or an annuity. These questions were dropped in the 1999 survey, but returned in the 2001 survey. They are moved to the end of the mortgage module and are asked only if the respondent reports that there is no other type of mortgage in place.

State or Local Government Low Cost Mortgage

Beginning in 1997 the survey asked whether the owner's mortgage was obtained through a State or local government program that provides lower cost mortgages.

Changes in Survey Methodology

During the 1990s there were some significant changes in the methodology used to conduct the AHS. These changes are known to have had an effect on responses to the survey. The replication analysis includes evaluations of some of the AHS mortgage variables before and after these changes to the survey methodology.

Computer Assisted Telephone Interviewing (CATI)

Large-scale experiments with CATI were conducted as part of the 1987, 1989, and 1991 AHS national surveys. The CATI experiments found that this change in methodology had a significant effect on the information collected. While the Census Bureau cannot determine definitely which

methodology provides better data, it has speculated that CATI income estimates are better, but that some other estimates are worse. In 1995 CATI was used for returning national sample households whenever possible and to the extent that CATI staff were able to handle the workload.

In 1997 the Census Bureau eliminated the paper questionnaire and instead conducted all interviews with computer assisted personal interviewing (CAPI), using laptop computers either over the phone or in person. The survey questions, including the skip instructions, were programmed into the laptop. Because of this change, the Census Bureau urges caution in comparing survey results from 1997 with prior years' findings. The computer program was based on two years of research and testing.

Dependent Interviewing

Dependent interviewing is the process by which the data from surveys conducted in previous years are used to improve the quality of the data collected in the current survey. This practice was first used in 1984 for the year built variable and then in 1987 for the square footage of the unit. Beginning in 1999, dependent interviewing was used more extensively. If the same household occupied the housing unit, some questions were modified to first ask the respondent whether the information given in the previous interview was still correct. Only when the respondent answered "no" was the question asked again. The questions affected by this change included many physical characteristics of the housing units.

Dependent interviewing was also used to verify the reported house value. If the current reported value was above or below the previous reported value by 35 percent, the interviewer was instructed to verify the owner's response.

Some questions were not repeated at all, even to verify information, if logically the answer should not have changed since the previous survey. If the same household occupied the housing unit, the answer from the prior survey was used for several questions that are important for mortgage-related research:

- The year the home was acquired and how acquired;
- The purchase price;
- Whether the owner was a first-time owner;
- The source of down payment; and
- The previous occupancy.

Finally, some questions were not asked again unless the housing unit had been added to the survey. These questions include the year the unit was built and whether it is a mobile home site placement.

Advantages and Disadvantages of the AHS for Mortgage-Related Research

The AHS can support several lines of mortgage market investigation that other databases are either unable or poorly suited to address. The principal advantages of the AHS for research on mortgages and housing finance are:

- It provides very *detailed household characteristics*, including standard demographic and income information, and some characteristics, such as first-time buyer status and length of tenure in the housing unit, that are found in few other databases. These detailed characteristics allow for in-depth analysis of the influence of household characteristics and stage in the life cycle on mortgage demand and product selection.
- It collects *information on mortgage characteristics and measures of mortgage risk* that are available from few other sources. For example, information on mortgage interest rates is collected directly in the AHS, while measures such as loan-to-value ratios (LTVs) and housing cost-to-income ratios can be computed from a combination of data elements collected.
- It provides the most *comprehensive information on property characteristics* of any data source, including the type of structure, the number of rooms, and the unit's physical quality. Physical quality and type of structure, for example, may influence the willingness of lenders to lend and the products they offer. This information allows analysts to control for them in analyses of mortgage product selection.
- The *panel nature* of the AHS supports analysis of changes in the use of mortgage financing over an owner's tenure in a housing unit.
- The AHS is a unique source of information on the *mortgage characteristics of the owner-occupied housing stock* at a given point in time.
- It provides at least some *limited measures of neighborhood quality* that might well also influence the willingness of lenders to extend credit and the terms on which that credit is extended.
- The public-use version of the AHS national sample file provides *fairly detailed geographic information* for sample data, including identification of the MSA, in addition to standard census region and metro status (central city, suburb, and non-metro) information. In addition, researchers may gain access to the internal AHS data files on-site at Census Bureau offices, which allows the identification of the census tract of each observation. Census tract information tied to detailed characteristics of mortgage borrowers and the homes and neighborhood in which they live provides unique opportunities to control for multiple determinants of mortgage supply and demand.
- It provides *comprehensive coverage of the mortgage market*, including the whole nation and all sources of mortgage credit.

- It is a *micro data set* that enables cross tabulation and regression analysis of all variables of interest.

However, there are some limitations of the AHS in terms of the nature of the survey or the information collected that would be of interest for analysis of the residential mortgage market. The principal disadvantages of the AHS are:

- The AHS provides a *relatively small sample size* for new originations. While the AHS has a large number of observations, the number of housing units with a newly originated mortgage will include only a small subset of the entire stock of housing.
- The collection of data every other year means that rapid *changes in mortgages during the period between surveys may not be captured*. For example, if a purchase occurs shortly after the survey is conducted and is then refinanced before the next survey is conducted, no information will be gathered on the purchase loan since the survey only asks about mortgages currently held on the property. As a result, the AHS only captures information on mortgages that were originated since the last survey wave *if* they have survived until the current survey wave. Over a two-year period, some newly originated mortgages will be terminated as a result of refinancing, household moves, and foreclosures. In order to avoid biases in estimates of loan originations resulting from loan terminations, one approach is to focus the analysis on originations during the calendar year the survey is conducted. However, this truncation further limits the sample size – while not necessarily removing all of the concerns about the loss of some originations, since these losses may have occurred early in the survey year.
- The AHS relies on owners to provide information on their current mortgages. In answering the mortgage-related questions, it seems likely that owners respond from memory rather than referring to their records. As a result, there are likely to be *errors in the owners' responses*.¹⁴
- The *MSA definitions are out of date* in the national survey. The MSA definitions used by the national AHS have not been updated since the sample was redrawn following the 1980 census. As a result, identification in the national AHS of metropolitan and non-metropolitan areas, as well as central city and suburban portions of MSAs does not coincide with current definitions. This discrepancy may make it difficult to compare findings by these geographic areas with other data sets that use current MSA definitions. The areas covered by the metropolitan area surveys also do not always match the current definitions of metropolitan areas.
- The AHS *does not collect any information related to mortgage termination directly*. Researchers can infer indirectly a loan termination from household moves and/or a comparison of mortgage characteristics (such as origination year and mortgage terms)

¹⁴ The Census Bureau sends a letter to surveyed households asking them to have their financial records available during the survey. Thus, our assertion about homeowners' answers are based on memory may be overstated, though errors in the responses are still of some concern.

across survey years for the same household. Since the reasons for loan termination are unknown, foreclosures cannot be distinguished from prepayments and household moves.¹⁵

Each of the issues cited above may or may not represent a significant obstacle for using the AHS for mortgage-related research. For example, while the AHS may not have a large sample of new originations, there may still be enough observations based on the two-year window between surveys to support analysis of loan originations. Similarly, it may be that owners' responses are accurate enough on average to provide reliable estimates of mortgage activity. The AHS geography may be similar enough to current OMB definitions to provide estimates of metro area mortgage markets.

In addition to the above concerns, there are also several shortcomings of the AHS that cannot be overcome. These shortcomings limit the research areas that the AHS can be used to address. These issues include:

- The AHS *does not have any information or has insufficient detail on some important borrower and loan characteristics*. The types of information that are unavailable include mortgage origination costs, other types of borrower debt, assets, and credit worthiness. An example of a variable for which the information collected may be insufficient is how frequently ARM payments are adjusted or the methods used to set the interest rate (i.e., the index rate, the margin, and the frequency of adjustment).
- The AHS *does not gather any information on the payment status of a mortgage*, so it is not possible to evaluate questions related to mortgage delinquency, default and foreclosure. This lack of default information may complicate refinance evaluations, since in reality borrowers in default will be constrained from refinancing.
- The AHS *does not provide any information on secondary mortgage market activity*. Since many owners are not aware of secondary market activity involving their mortgages, the AHS does not attempt to gather information on this aspect of the market. The AHS is thus generally not suitable for analysis of issues related to secondary market activity.

AHS Mortgage Variables Chosen for the Replication Analysis

Among the mortgage and housing finance information collected in the AHS, the following measures were chosen to be tested in the replication analysis in Chapter Four:

- Mortgage origination volume;
- Mortgage market segment;
- Loan purpose (home purchase vs. refinance);

¹⁵ It is worth reiterating that the AHS surveys follow housing units, not households. However, there are variables in the public-use file that enable us to identify whether the household residing in the sample housing unit is the same one from the last survey.

- Original mortgage amount;
- Mortgage interest rate;
- Loan-to-value (LTV) ratio at origination;
- Mortgage payment product type;
- First-time homebuyer status;
- Borrower race/ethnicity; and
- Borrower income.

The selection was based on the importance of these variables in mortgage market research and the availability of reliable benchmark measures from other independent data sources. Appendix H shows how each attribute can be derived from the AHS.

Exhibit 2-4 highlights some of the advantages of the AHS for analysis relying on these variables and identifies the primary concerns about the AHS’s reliability for each variable before we conducted the replication analysis. In the next chapter, we discuss the principal sources of comparison databases chosen for the replication analysis. The specific goals and approach to conducting the replication analysis on these AHS mortgage data elements are delineated in Chapter Four.

Another important objective of the study is to evaluate whether the mortgage information is reported consistently by homeowners across time so that the AHS can be used reliably to support research that makes extensive use of the panel structure of the survey design. Findings for such investigation are presented in Chapter Five.

**Exhibit 2-4
Summary of Advantages and Reliability Concerns for Analysis Using the Key
Mortgage Variables**

VARIABLE	POTENTIAL ADVANTAGES OF AHS	POTENTIAL PROBLEMS IN USING AHS
Mortgage Origination Volume	<ul style="list-style-type: none"> • More complete geographic and market segment coverage than most other data sources 	<ul style="list-style-type: none"> • Only outstanding loans are observed in the survey • Not all originations are captured in the interview year
Mortgage Market Segment	<ul style="list-style-type: none"> • More complete geographic and market segment coverage than most other data sources • More detailed borrower and housing characteristics than most other data sources 	<ul style="list-style-type: none"> • Unclear whether borrowers report mortgage insurance type correctly, especially over time

VARIABLE	POTENTIAL ADVANTAGES OF AHS	POTENTIAL PROBLEMS IN USING AHS
Loan Purpose (Home Purchase vs. Refinance)	<ul style="list-style-type: none"> • Largest sample of a regularly recurring survey covering loan characteristics prior to refinancing 	<ul style="list-style-type: none"> • Loan purpose is not explicitly asked in the survey • Unclear whether borrowers report year of loan origination correctly • May underestimate refinance activity in cases where purchase loan is refinanced in same year as purchase
Original Mortgage Amount	<ul style="list-style-type: none"> • More complete geographic and market segment coverage than most other data sources 	<ul style="list-style-type: none"> • Unclear whether borrowers report loan amount accurately and consistently
Mortgage Interest Rate	<ul style="list-style-type: none"> • Largest sample size of a regularly recurring survey covering interest rates across all market segments and loan purposes 	<ul style="list-style-type: none"> • Unclear whether borrowers report interest rate correctly, especially over time • Does not gather information on hybrid ARMs and adjustment factors for ARMs
LTV at Origination	<ul style="list-style-type: none"> • Unique coverage of LTV for all market segments and loan purposes 	<ul style="list-style-type: none"> • Unclear whether borrowers report loan amount and house price correctly even shortly after purchase • For refinance loans, unclear whether the owner's estimate of market value of the house is reliable
Mortgage Payment Product Type	<ul style="list-style-type: none"> • Largest sample size of a regularly recurring survey covering mortgage product types across all market segments and loan purposes 	<ul style="list-style-type: none"> • Unclear whether owner-reported mortgage payment features are complete
First-time Homebuyers	<ul style="list-style-type: none"> • Largest sample of a regularly recurring survey covering first-time homebuyers 	<ul style="list-style-type: none"> • Unclear whether borrowers' self-identification of first-time homebuyer is accurate and consistent over time
Borrower Race/Ethnicity	<ul style="list-style-type: none"> • More complete geographic and market segment coverage than most other data sources • More detailed borrower and housing characteristics than most other data sources 	<ul style="list-style-type: none"> • Unclear whether borrowers' self-identification of race/ethnicity is accurate
Borrower Income	<ul style="list-style-type: none"> • Detailed breakouts of income components • More complete geographic and market segment coverage than most other data sources 	<ul style="list-style-type: none"> • Unclear whether self-reported incomes are accurate and complete

Chapter Three

Comparison Databases

This chapter describes the comparison databases chosen for benchmarking the AHS mortgage variables. The first part of the chapter provides an overview of these data sources. The strengths and weaknesses of each database for mortgage research are examined. The second section discusses the rationale for our choice of each data set used in the replication analysis.

Principal Sources of Comparison Databases

We used seven databases to provide benchmark measures for the AHS mortgage variables. They are:

- Home Mortgage Disclosure Act (HMDA) Data
- FHA Administrative Data
- Federal Housing Finance Board's (FHFB) Mortgage Interest Rate Survey (MIRS)
- Residential Finance Survey (RFS)
- Survey of Consumer Finance (SCF)
- Chicago Title and Trust Co. Annual Survey of Recent Homebuyers
- National Association of Realtors' (NAR) Existing Home Sales

Detailed information about these databases is presented in a series of exhibits in Appendix B.

Home Mortgage Disclosure Act (HMDA) Data

The HMDA files are among the most widely used micro-data in the fields of housing finance and mortgage market research. Mandated by the Home Mortgage Disclosure Act, passed in 1975 and updated in 1989, systematic information about mortgage applications and the loan applicants is collected from most lending institutions operating in metropolitan areas. Originally, only aggregate census tract information for metropolitan areas was reported. Since 1989 information on each loan applicant has been required, including the disposition of the application; the state, county, MSA, and census tract location of the property; the income, race, and sex of each applicant; and the type and purpose of the loan. Originally HMDA reporting requirements only applied to depository institutions, but non-depository mortgage banking companies were required to report beginning in 1992. Since then the data have covered almost all mortgage loans originated in metropolitan areas.

Among the principal strengths of the HMDA data are that they provide detailed geographic information, key demographic variables, and identification of the mortgage market segment of the loan. HMDA data are widely used by the industry and research community to examine the geographic and demographic distribution of loan product types, lending volumes, and application denial rates. However, HMDA data do not cover some important mortgage characteristics, including

payment plan, loan term, interest rate, and loan-to-value (LTV) ratio. In addition, HMDA data has relatively poor coverage in non-metropolitan areas and of home equity loans.

FHA Administrative Data

FHA administrative data provide a comprehensive census of FHA-insured single-family mortgages originated since mid-1970s. Loan-level information on the universe of mortgages originated with HUD FHA insurance since the mid-1970s has been kept in several HUD relational databases.¹⁶ The FHA data provide all standard underwriting variables, including debt-to-income ratios and LTVs, as well as detailed loan characteristics including the interest rate. A variety of borrower demographic characteristics are also available. In addition to this detailed information on loan and borrower characteristics at origination, FHA data also include information on default and termination of loans over time. Thus, the data support analysis of new loan originations, the performance of mortgages over time, and the characteristics of the stock of FHA mortgages at a given point in time. The principal weakness of FHA data is that it only covers one segment of the market and so by itself it cannot be used to examine the broader mortgage market.

Federal Housing Finance Board's (FHFB) Mortgage Interest Rate Survey (MIRS)

A national sample of mortgage lenders (about 300) are asked by the FHFB each month to report the terms on all single-family (one-unit structure), fully amortized, non-farm conventional home-purchase loans closed during the last five business days of the month. A critical limitation is that the survey excludes FHA-insured and VA-guaranteed loans, multifamily loans, mobile home loans, refinancings, and product types other than ARMs and fixed-rate level-payment mortgages. In addition, loans made by financial institutions that serve the subprime market are not included in the survey.¹⁷ The MIRS data collection started in the mid-1960s. While only summary level data are available from FHFB's Internet site,¹⁸ loan-level data can also be purchased at a reasonable cost. The micro data files used in this study were provided by HUD staff. The data provide monthly information on an array of mortgage characteristics, including interest rates, loan terms, purchase house prices and LTV ratios by property type (new and previously occupied), by loan type (fixed- or adjustable-rate), and by lender type (savings associations, mortgage companies, commercial banks, and savings banks). Identifiers for state and major MSA/CMSA locations as well as a sampling weight variable are available in the data. The MIRS data are used for indexing the conforming loan limits for the two GSEs. Economic policy makers also use the data to determine benchmarks for interest rates, down payments, terms of maturity, terms of adjustable-rate mortgages (ARMs), fees and charges, and other trends in mortgage markets.¹⁹

¹⁶ Much of the information on FHA single-family loan originations, mortgage characteristics and loan performance is stored in the A43 and F42 files in FHA's Single Family Data Warehouse System (SFDWS). This information is updated monthly.

¹⁷ We thank Timothy Forsberg of the Federal Housing Finance Board for clarifying the survey universe of MIRS.

¹⁸ <http://www.fhfb.gov>

¹⁹ "Notice and Request for Comments on Monthly Survey of Rates and Terms on Conventional 1-Family Nonfarm Mortgage Loans," *Federal Register*, V.65(187), September 26, 2000.

Residential Finance Survey (RFS)

The RFS is conducted by the U.S. Bureau of the Census and is designed to provide comprehensive information about the financial, ownership, and structural characteristics of non-farm, privately owned residential properties. The RFS has been conducted as part of the Census of Housing since 1950. The latest data available are for 1991. The 1991 RFS represents about 66 million properties securing 38 million mortgages. The public-use micro data cover almost the entire set of mortgage and housing finance data elements available in the AHS.²⁰ Variables about the demographic characteristics of the occupants are also collected. A particular strength of this survey is that questionnaires are completed by the property owner or agent (in the case of multifamily rental housing) or by the holder(s) of any mortgages on the property (in the case of owner-occupied property). In particular, some of the crucial loan characteristic questions (origination year, mortgage insurance status and interest rate) are answered by mortgage holders rather than borrowers. These data elements should be quite accurate, since the lenders' information is likely supported by organizational records, rather than the homeowners' memory as in data based only on household surveys. Besides the broad coverage of mortgage and housing finance characteristics, geographic areas, and mortgage market segments, another advantage of the RFS data is its large sample size. For example, the 1991 data contain records on approximately 70,000 properties. However, like the AHS, it is a sample of the housing stock, so the sample size for new originations may be small. Other limitations of this database are its infrequent availability and limited geographic identifiers.

Survey of Consumer Finance (SCF)

The SCF is a triennial survey of U.S. families sponsored by the Board of Governors of the Federal Reserve System. The data collection started in the early 1980s, and the latest data available are for 1998. The survey is designed to provide detailed information on the families' balance sheets and their use of financial services, along with a comprehensive set of demographic characteristics. The questionnaire includes almost the same set of mortgage and housing finance data elements as the AHS. Its major strength is the detailed information on all aspects of the household's financial situation. The major disadvantage is its sample size, about 4,500 families. Therefore, the number of newly originated mortgages in this sample is quite small. Users of the data files have also reported some inconsistent tabulation results when estimates are compared across different survey years.²¹ Finally, only crude geographic identifiers (four Census regions and nine Census divisions) are available in the public-use version of the data.

Chicago Title and Trust Co. Annual Survey of Recent Homebuyers

From 1976 to 1999, the Chicago Title and Trust Co. conducted an annual survey of a sample of approximately 1,800 homebuyers in 20 major metropolitan housing markets.²² These housing

²⁰ Exceptions are whether the mortgage lender is an individual rather than a bank/organization and whether the mortgage also covers other homes, farm land, or a business on the property.

²¹ Communication with Eric Belsky of the Joint Center for Housing Studies, Harvard University.

²² The 20 major metropolitan housing markets covered by the 1999 Chicago Title survey are Atlanta, Boston, Chicago, Cleveland, Dallas/Fort Worth, Denver, Detroit, Houston, Los Angeles, Memphis (Tenn.), Twin Cities, New York City, Orange County (Calif.), Orlando (Fla.), Philadelphia, Phoenix, San Francisco, Seattle/Tacoma and Washington, D.C.

markets together account for about one-third of all home sales in the United States. Summary level data (i.e., in cross-tabulation format) were published annually in *Who's Buying Homes in America*.²³ The tabulations contain information on the mortgages, the properties, and the demographic and economic characteristics of the homebuyers. Perhaps the most frequent use of this survey is to report on the characteristics of first-time homebuyers. However, the survey is not designed to be representative of any particular geographic area or market segment. In addition, no survey has been conducted since 1999.

National Association of Realtors' (NAR) Existing Home Sales

The research division of the National Association of Realtors (NAR) regularly compiles housing statistics series related to sales volumes and prices for existing homes, using sample data collected from regional Multiple Listing Services. NAR research staff have recently (1999) re-benchmarked the home sales series using data from the 1990 Census of Housing. Data are available monthly for the nation and regions and quarterly for state and MSA levels. It is the principal source of information available on the level of sales and home prices. However, it provides a single estimate for entire market areas. Another disadvantage is that it only covers sales of existing homes.

Exhibit 3-1 at the end of this chapter provides a summary of the strengths and weaknesses of the data sets discussed.

Other Comparison Databases Considered

The study also considered the following comparison databases:

- The Census Bureau's Survey of Income and Program Participation (SIPP)
- Panel Study of Income Dynamics (PSID) conducted by the Survey Research Center at the University of Michigan
- HUD's Survey of Mortgage Lending Activity (SMLA)
- Freddie Mac's Primary Mortgage Market Survey (PMMS)
- Mortgage Information Corporation's (MIC) Loan Performance System

These are not used in the replication analysis for a number of reasons. Mortgage elements in SIPP and PSID are rarely used and their reliability is uncertain. Furthermore, the PSID data have a relatively small sample size. The SMLA data are of limited value in the replication analysis both because of its questionable coverage on some segments of the mortgage market (loans reported by commercial banks and mortgage bankers) and because of its elimination after 1997. Freddie Mac's PMMS is not useful for the purpose of this study because it only contains commitment rate information, rather than the actual contract rates at loan closing. The primary reasons for not using

²³ A PDF version of the report can be downloaded from the company's Internet site (<http://www.ctic.com/HomeSurvey/default.htm>).

the MIC data are that the data sources and loan selection criteria are unknown and that the database would be very expensive to acquire.

Choice of Comparison Databases for the Replication Analysis

The reasons for our use of each data set for the replication analysis are summarized below.

HMDA

We make extensive use of HMDA data in the replication analysis. The principal advantages of using the HMDA database for the replication analysis are that it provides coverage of loan originations in all mortgage market segments and that it is available annually.²⁴ Given HMDA's spotty coverage outside of metropolitan areas, comparisons between HMDA data and the AHS will be limited to loans made in metro areas. As will be discussed in Chapter Four, there are also differences in the AHS and HMDA in the MSA definitions used. These differences are a potential source of bias in the benchmarking. Despite these problems, HMDA is still an important comparison database given its broad coverage of mortgage market segments and geographic areas.

FHA Administrative Data

While covering only one segment of the mortgage market, FHA data provide a census of loans of this type originated since the mid-1970s and constitute a very reliable benchmark of activity in this segment. In addition, FHA data provide many mortgage characteristics that are unavailable from most other comparison data sets. Therefore, we use FHA data both to benchmark AHS volume estimates and to test the distribution of key loan and borrower characteristics of mortgages in this segment of the market.

FHFB Monthly Interest Rate Survey (MIRS)

The MIRS is a key source of detailed information on the characteristics of new conventional loan originations, so it is a valuable complement to the FHA data in the replication analysis. Its use is somewhat limited by the fact that it includes information only on home-purchase mortgages.

Residential Finance Survey (RFS)

Because the RFS gathers information directly from lenders, it provides a valuable test of the reliability of owner-reported information in the AHS for certain loan characteristics. However, there are two factors that limit the usefulness of the RFS in the replication analysis. First, it is only available for one point in time from the period studied (1991). Second, it provides only crude information on the date of loan origination which makes it difficult to match loan volume estimates to other data sources.²⁵

²⁴ The RFS also provides coverage of all market segments, but at present it provides information only on loans originated from 1989 through the first part of 1991. Another wave of the RFS conducted during 2001 will be available soon.

²⁵ For confidentiality reasons, the public-use version of the RFS allows users to only identify origination years in interval format. For example, "1989 to early 1991" is one of the intervals.

Survey of Consumer Finance (SCF)

Because it gathers information on mortgages from all market segments, we use SCF to provide benchmark measures of loan distribution by market segment. There are two factors that limit our use of the SCF in the replication analysis. First, the SCF is conducted every three years, while the AHS is conducted every other year. As a result, the only interview year when these surveys overlap during the 1990s is 1995 and we confine the comparisons between the AHS and the SCF to that year. Second, SCF has a fairly small sample size, so it does not provide very precise estimates of mortgage activity.

Chicago Title Annual Survey of Homebuyers

The Chicago Title Survey is one of the few sources of information on first-time homebuyers. Therefore, it is used to evaluate information from the AHS on this market segment.

NAR Sales Volume of Existing Homes

The NAR data series offer benchmarks for sales of existing homes. In conjunction with the sales information on new homes estimated by the Census Bureau's Survey of Construction (SOC), we use these data to benchmark the volume of home purchase activity in AHS.

Exhibit 3-1

Strengths and Weaknesses of Non-AHS Sources of Mortgage Market Data

DATA	DESCRIPTION	STRENGTHS	WEAKNESSES
Home Mortgage Disclosure Act (HMDA) Data	Annual census of mortgage loan applications by selected lenders active in metropolitan areas	<ul style="list-style-type: none"> ▪ Detailed geographic information ▪ Key demographic information ▪ Good market segmentation information 	<ul style="list-style-type: none"> ▪ Spotty coverage of lending activity in non-metro areas and of second mortgages ▪ Lacks information on key mortgage characteristics
FHA Administrative Data	Administrative data on all loan originations since mid-1970s and loan performance over time	<ul style="list-style-type: none"> ▪ Detailed information on loan, borrower and geographic location ▪ Census of FHA activity since mid-1970s ▪ Good source of information on loan performance 	<ul style="list-style-type: none"> ▪ Covers only one market segment
FHFB Monthly Interest Rate Survey (MIRS)	Monthly survey of mortgage lenders collecting detailed data on newly originated conventional home-purchase loans	<ul style="list-style-type: none"> ▪ Provides data on most important loan characteristics ▪ Timely data and long running series 	<ul style="list-style-type: none"> ▪ Does not provide demographic information ▪ Does not cover non-purchase mortgages ▪ Covers only conventional loans
Residential Finance Survey (RFS)	Large, nationally representative sample of the use of mortgage finance for the housing stock, conducted as part of Census of Housing	<ul style="list-style-type: none"> ▪ Mortgage data obtained directly from lenders ▪ Detailed information on households and mortgage characteristics ▪ Unique source of data on financing of entire housing stock, including multifamily rental housing 	<ul style="list-style-type: none"> ▪ Available only every 10 years; latest survey now 10 years old ▪ As sample survey of entire housing stock, may not have large sample of new originations
Survey of Consumer Finance (SCF)	Nationally representative survey designed to provide detailed information on household finances, including mortgages	<ul style="list-style-type: none"> ▪ Covers all segments of the mortgage market and all geographic areas ▪ Detailed information on households and mortgage characteristics 	<ul style="list-style-type: none"> ▪ Small sample size ▪ Given nature of sample and sample size, sample of new originations too small for most analysis
Chicago Title Annual Survey of Recent Homebuyers	Annual survey from 1976 to 1999 of recent homebuyers in 20 MSAs	<ul style="list-style-type: none"> ▪ Detailed information on homebuyers and their housing and mortgage characteristics ▪ One of few sources of data on first-time homebuyers 	<ul style="list-style-type: none"> ▪ Micro data are proprietary ▪ Only representative of the 20 markets surveyed ▪ Discontinued in 1999
National Assoc. of Realtors (NAR) Sales of Existing Homes	Monthly and quarterly data on sales of existing homes at MSA level	<ul style="list-style-type: none"> ▪ Covers large share of existing home market ▪ Broad geographic coverage 	<ul style="list-style-type: none"> ▪ Aggregate data with no disaggregation by property characteristics or distribution of loan types

Chapter Four

Replication Analysis

This chapter presents our findings for the replication analysis through which we test the accuracy of the mortgage measures reported in the AHS by comparing them to external benchmark estimates from other mortgage market databases. The first section discusses the general goals and approach for the replication analysis. Results regarding the reliability of the AHS for deriving volume estimates for mortgage originations are presented in the second section. The third section examines the distribution of key mortgage characteristics of the loans captured in the AHS and compares those distributions to estimates derived from other databases. The reliability of borrower characteristics in the AHS is examined in the last section of this chapter.

Goals and Approach for the Replication Analysis

The goal of the replication analysis is to evaluate the quality of key data elements in the AHS for conducting mortgage market research. The basic approach is to examine the AHS estimates and see whether they replicate mortgage market measures from external data sources that are thought to provide accurate measures of market activity. Unless indicated otherwise, our analysis focuses exclusively on benchmarking the AHS estimates for all primary mortgages of owner-occupied single-family units, defined as structures with one to four units.

Potential discrepancies in estimates could arise from a variety of factors. These may include:

- **Sampling errors.** These reflect the fact that the AHS or other survey databases used in the comparison are based on a sample of housing units, properties or households and therefore estimates may vary from the actual values from a complete census.
- **Measurement errors.** These include all the nonsampling errors (other than sample coverage and nonresponses) that occur during the data collection process in the AHS survey. Some of the possible sources of measurement error are:
 - Incorrect answers from the survey respondents;
 - Questionnaire design, content, and wording; and
 - Interview mode.
- **Differences in definitions, timing, and coverage** between the AHS and other data sources.
- **Data processing errors.** Potential errors could occur when the raw AHS data are processed and when data elements in the public-use file are created by the Census Bureau.

In the remainder of this section, we discuss how some of the above factors have informed the selection of comparison databases and time periods of the AHS data used in the study.

In general, administrative data (such as the FHA administrative records) are preferable for use in the replication analysis since there should not be any sampling and measurement errors associated with such databases. Survey data (such as RFS) that do not rely on the respondents' memory should have fewer measurement errors. Also, larger surveys are assumed to have fewer sampling errors, when compared to surveys with small sample sizes (such as SCF).

To account for the possibility that differences in estimates are due to sampling errors, we present 95-percent confidence intervals for most of the AHS measures in addition to the point estimates. If the comparison database is based on survey data, we also present confidence intervals for the benchmark estimates.²⁶ In conducting the replication analysis, as long as the benchmark estimate falls within the two end-points of the AHS confidence interval, any discrepancies are considered to be not statistically significant.

Developing common time frames for the comparisons of the AHS and benchmark estimates from other data sources represents a challenge. The first thing to note is that the AHS interviews are conducted over a period of time between July and December of the survey year. Some mortgage activity from the second half of the survey year, therefore, is not captured in the AHS. As a result, using the AHS to estimate origination volume for the calendar year in which the survey is conducted will understate the actual volume. Second, records in most of our comparison data sets, including FHA, HMDA, and MIRS files, are designed to capture information on all mortgage *originations* in a given time interval. Information collected in the AHS, on the other hand, represents a snapshot of the mortgage *stock* as of the time of the interview. In other words, mortgages observed in the survey are only surviving loans (i.e., loans still outstanding at the time of the interview).

Because of concerns about loan turnover resulting from property sales, refinancings, or foreclosures, only loans originated in the two years prior to the AHS interview are examined in the replication analysis.²⁷ For example, we use the 1997 survey data to study the characteristics and volume of loans originated in 1995 and 1996. Because of concerns about missing loans originated in the second half of the year, mortgage activity that happened in the interview year is thus studied using data collected in the following survey.²⁸

²⁶ In particular, the Survey module of the STATA[®] statistical package is used to estimate the confidence intervals taking into account the sampling design used in the survey. Since the Primary Sampling Unit (PSU) information is not available from the public-use version of the AHS, we assume each sampling housing unit represents a unique PSU in calculating the confidence intervals. This assumption is likely to make the confidence intervals smaller than they should be.

²⁷ HUD internal analysis has suggested that the undercounting of loan originations in the survey year due to the timing of the interview is more serious than the underestimates for the prior or second prior years due to loan terminations that happened before the interview.

²⁸ There is one exception. Whenever data from the 1991 RFS are used as benchmarks, they are compared to data collected in the 1991 AHS for mortgage activity that happened between 1989 and 1991. There are two reasons for this exception. First, interviews for both surveys were conducted during calendar year 1991. Second, for confidentiality reasons, the public-use version of the RFS file allows users to only identify origination years in interval format and "1989 to 1991" is one of the interval periods.

Since it is assumed that the findings of this study will be of most use in informing analyses of the AHS surveys from recent years and in the future, the replication analysis focuses on the AHS national surveys conducted in the 1990s. Exhibit 4-1 illustrates the correspondence between the AHS survey years and calendar years of mortgage activity examined in the study. Using all six surveys over the decade allows us to investigate the impact of changes in survey methodology and questions in the late 1990s on the reliability of the AHS mortgage variables.²⁹

**Exhibit 4-1:
Correspondence between AHS Survey Years and Calendar Years of Mortgage Activity Examined**

AHS SURVEY YEAR	CALENDAR YEARS OF MORTGAGE ACTIVITY EXAMINED
1991	1989 1990
1993	1991 1992
1995	1993 1994
1997	1995 1996
1999	1997 1998
2001	1999 2000

To account for the number of loans that have terminated prior to the interview in each origination cohort, we have devised a method to adjust all the AHS volume estimates based on the historical termination rates for FHA loans.³⁰ A detailed description of the adjustment method as well as other

²⁹ For the 1999 survey, we used the public-use file released on November 14, 2002. This version corrects an error in the WHNGET variable associated with the previous release of the data. This variable is used to identify the origination year for mortgages that were obtained at the same times as home purchase. The error had caused the 1999 reported answer to the question of when the home was acquired to be overwritten by the value from 1997 survey. The result of this error was an overestimate of loan volumes in 1997 and an underestimate of loan volumes in 1998.

³⁰ Ideally, the volume estimates should be adjusted using historical termination rates from the corresponding market segments and mortgage product types, since default and prepayment speeds are likely to be different across market segments. For example, borrowers with conventional mortgages are more likely to refinance their mortgages than their FHA counterparts, given the same interest rate environment; on the other hand, FHA loans have higher claim rates. However, termination rates are not readily available for the non-FHA segments of the market. We expect that the termination rates for FHA loans can serve as a reasonable proxy for the other market segments because on average the low FHA prepayment rates should be offset by the high claim rates of FHA loans. An alternative approach is to use prepayment (termination) rates of securities issued by the GSEs (Ginnie Mae, Fannie Mae, and Freddie Mac). These data can be purchased from a number of Wall Street firms. But this method may be problematic because GSE securities represent pools of mortgages and their prepayment and claim speeds could potentially be different from their loan-level counterparts.

alternative approaches to study recent mortgage market activity in the AHS can be found in Appendix C.³¹

The replication analysis focuses on examining the reliability of national estimates of mortgage activity in the AHS. Nonetheless, one of the principal comparison data sets, HMDA, is known to have poor coverage of mortgages made in the non-metro areas. Any HMDA tabulations and the corresponding AHS comparison are, therefore, restricted to metro areas of the nation. There is a potential bias in this comparison. As mentioned in Chapter Two, the metro area boundaries in the AHS are based on the 1980 Census, while the HMDA records use MSA definitions from the 1990 Census. This implies that the fast growing outlying counties of metro areas are not included in the AHS estimates but are contained in the HMDA files. Given the fact that metro areas have grown in size since 1980, it is likely that the AHS understates the volume of mortgage activity relative to the HMDA data. At the same time, HMDA's coverage for originations in the metro areas is known to be incomplete because some lenders and loan originations are not required to report. The net impact of these two biases on the related replication analyses is unknown.

The study originally also intended to evaluate the reliability of the AHS metropolitan area surveys, focusing on mortgage origination volumes. However, our preliminary tabulations indicated that the relatively small sample sizes of these surveys prevent us from drawing any statistically meaningful comparisons with the benchmark databases. This is especially the case for examining the volume of loan cohorts by loan purpose and mortgage market segment. Given the small sample sizes and large standard errors, we have not performed any replication analysis on those AHS files.

Exhibit 4-2 on the following page provides a crosswalk between the specific AHS mortgage variables that we intend to test in the replication analysis and the corresponding benchmark data set(s) used. It also outlines the cohorts of mortgage market activity examined for each database and any stratification necessary when the evaluation is performed. The last column summarizes whether any geographic or loan type restrictions are imposed as a result of the nature of the benchmarking data.

Appendix H shows how each of the mortgage market measures can be derived from the AHS.

³¹ This adjustment is crucial. According to Exhibit C-2 of Appendix C, the adjustment factor varies substantially across origination cohorts. The average annual adjustment rate is 15.04 percent over the 1989-2000 period. For three of the origination cohorts, the adjustment rates are over 20 percent.

Exhibit 4-2:

Crosswalk Between Benchmark Data Sources and AHS Mortgage Market Measures Tested in the Replication Analysis

MORTGAGE MARKET MEASURE	BENCHMARK DATA & ORIGINATION COHORTS EXAMINED	STRATIFICATION	COMMENTS
Mortgage Origination Volume	<ul style="list-style-type: none"> FHA: 1989-2000 HMDA: 1993-2000 	<ul style="list-style-type: none"> Mortgage Market Segment^a Loan Purpose^b 	<ul style="list-style-type: none"> FHA data are used to benchmark FHA loans originated nationally. HMDA data are used to benchmark VA/FmHA/RHS, conventional conforming and jumbo loans made in metro areas.
<i>Loan Characteristics</i>			
Mortgage Market Segment	<ul style="list-style-type: none"> RFS: 1989-1991 HMDA: 1993-2000 SCF: 1993-1994 	<ul style="list-style-type: none"> Loan Purpose 	<ul style="list-style-type: none"> HMDA analyses are restricted to loans made in metro areas. RFS and SCF are used to benchmark loans originated nationally.
Origination Loan Amount	<ul style="list-style-type: none"> FHA: 1989-2000 HMDA: 1993-2000 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> FHA data are used to benchmark FHA loans originated nationally. HMDA data are used to benchmark VA/FmHA/RHS, conventional conforming and jumbo loans made in metro areas.
Mortgage Interest Rate	<ul style="list-style-type: none"> FHA: 1989-2000 MIRS: 1992-2000 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose Payment Product Type^c 	<ul style="list-style-type: none"> FHA data are used to benchmark FHA loans. MIRS data are used to benchmark conventional conforming home-purchase loans of non-mobile home 1-unit structures.
LTV at Origination	<ul style="list-style-type: none"> FHA: 1989-2000 MIRS: 1992-2000 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> FHA data are used to benchmark FHA loans. MIRS data are used to benchmark conventional conforming home-purchase loans of non-mobile home 1-unit structures.
Mortgage Payment Product Type	<ul style="list-style-type: none"> FHA: 1989-2000 MIRS: 1992-2000 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> FHA data are used to benchmark FHA loans. MIRS data are used to benchmark conventional conforming home-purchase loans of non-mobile home 1-unit structures.

MORTGAGE MARKET MEASURE	BENCHMARK DATA & ORIGINATION COHORTS EXAMINED	STRATIFICATION	COMMENTS
<i>Borrower Characteristics</i>			
First-time Homebuyer Status (Home Purchase Originations Only)	<ul style="list-style-type: none"> Chicago Title: 1995-1999 RFS: 1989-1991 FHA: 1991-2000 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> Chicago Title data are used to benchmark aggregate shares of home purchases made in twenty metro areas. RFS data are used to benchmark aggregate share as well as shares by market segment of purchases made nationally. FHA data are used to benchmark FHA loans originated nationally.
Borrower Race/Ethnicity	<ul style="list-style-type: none"> FHA: 1989-2000 HMDA: 1993-2000 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> FHA data are used to benchmark borrowers of FHA loans originated nationally. HMDA data are used to benchmark borrowers of VA/FmHA/RHS, conventional conforming and jumbo loans made in metro areas.
Borrower Income	<ul style="list-style-type: none"> HMDA: 1993-2000 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> HMDA analyses are restricted to loans made in metro areas.

Notes: ^aMortgage market segments are FHA, VA/FmHA/RHS, conventional conforming, and conventional jumbo.

^bLoan purposes are home purchase and refinance mortgages.

^cPayment product types are fixed-rate and adjustable rate mortgages.

Mortgage Origination Volume

This section examines the accuracy of the AHS data in capturing mortgage origination volumes over the last decade overall and by mortgage market segment and loan purpose. All AHS volume estimates used in the comparisons are adjusted by the termination-rate method documented in Appendix C. One challenge in this analysis is that a portion of homeowners did not provide an answer to the mortgage insurance status question in AHS. They account for 2 to 3 percent or less of the total originations in a given year, except for 1993 and 1994, for which they represent about 5 percent of the originations. To account for these mortgages, we have further adjusted the loan volume estimates for each market segment (separately for home purchases and refinances) by these proportions, assuming borrowers who did not answer this question are proportionally distributed across the four market segments according to the market share of each segment.

Our analyses focus exclusively on benchmarking the *number* of mortgages originated in a given year against independent volume estimates derived from external sources. The alternative approach of examining loan volume in terms of *dollars* (i.e., total origination amount) is inferior since it would require the assumption that the AHS respondents reported the principal amount of the mortgage accurately as well as the loan origination information. The reliability of loan amounts in the AHS will be evaluated separately in a later section of this chapter.

Two comparison databases are selected to provide benchmark estimates. While covering only one segment of the mortgage market, FHA data provide a census of loans of this type originated since the mid-1970s and so should be a very reliable benchmark of activity in this segment. Given its broad coverage of mortgage market segments and the availability of a long time-series, HMDA data are a logical choice for benchmarking the AHS volume estimates of originations in the other three segments of the market.³² The 1991 RFS is not used in this particular evaluation because the survey was conducted approximately four months earlier than the 1991 AHS, which causes the volume estimates of recent originations from the two surveys to be different.

Our comparisons indicate that the AHS data are successful in capturing home-purchase mortgage originations for most cohorts over the decade. The adjusted volume estimates are on average very close to the benchmarks for loans in the FHA, VA/FmHA/RHS, and conventional conforming segments of the mortgage market. Coverage for originations in the conventional jumbo segment is less satisfactory in relative terms. Across all market segments, there appears to be a slight decline in coverage rate for the last three cohorts of loans (1998, 1999, and 2000).

Refinance originations were not captured adequately in the AHS for most of the cohorts examined, although reporting appears to be complete for some of the years. This is true across all market segments. The coverage for such mortgages has been especially poor in the 1995 survey. We suspect this may be the result of the questionnaire redesign in 1995 to add questions about home equity loans which created confusion among some respondents between cash-out refinance mortgages and lump-sum home equity loans.

³² For any comparisons that involve the AHS, HMDA and the FHA administrative data in this study, no attempt is made to exclude manufactured homes from the three data sets.

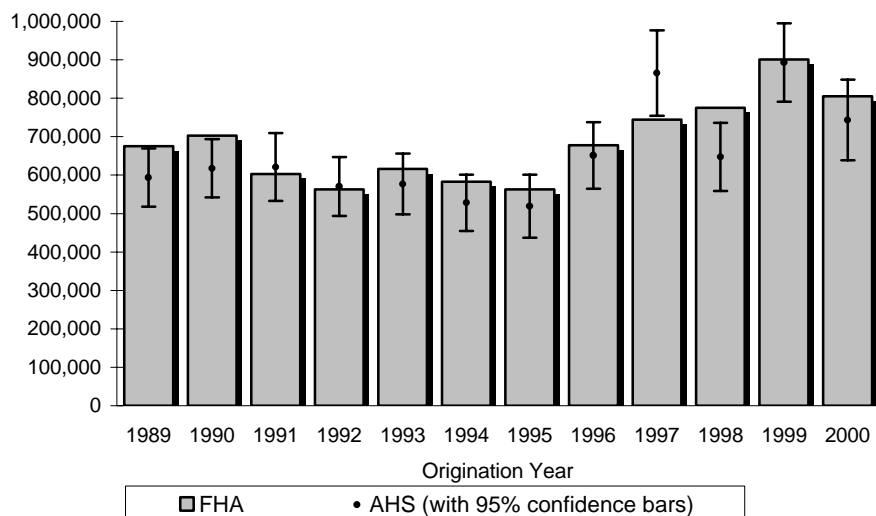
Except for the 2001 survey, loan purpose (home purchase vs. refinance) determination in the AHS is based on a comparison between loan origination year and house acquisition year. Loan purpose information was collected explicitly in the 2001 AHS; the public-use file includes a variable called REFI (refinance) that allows to us separate refinances from home purchases. The 2001 data, therefore, provide us with an opportunity to identify the 1999 and 2000 originations using both methodologies. Our analysis indicates that for home purchases the two methods yield volume estimates that are very similar. For refinance originations, surprisingly, the old methodology appears to outperform the REFI variable by generating volume estimates that are slightly closer to the benchmarks. This is especially true for capturing refinances in the conventional conforming segment of the market. A possible explanation is that 2001 was such a big refinancing year that the proxy adjustment rates have failed to account for many of the loans that had terminated in the two years prior to the interview. The REFI variable may perform better in future survey years.

The following sections provide more detail on the accuracy of the AHS for estimating mortgage origination volumes.

Home-Purchase Mortgages

Exhibit 4-3 compares the origination volume of home-purchase mortgages with FHA insurance estimated from the AHS and from FHA administrative data.

**Exhibit 4-3:
Number of FHA-Insured Mortgages
Home-Purchase Originations
AHS vs. FHA**



This analysis covers twelve cohorts of originations, from 1989 to 2000.³³ Unless stated otherwise, loan purpose determination for the 1999 and 2000 cohorts is based on the REFI (refinance) variable provided in the public-use file. The exhibit indicates that, overall, the point estimates for AHS mortgage volumes have tracked well with estimates calculated from FHA data. Estimates for some of the years line up with the benchmarks better than others. This is especially the case if we make use of the confidence interval estimates in the comparison, which account for the sampling variability of the AHS. Except for originations in the 1989, 1990, 1997, and 1998 cohorts, volume estimates for other years based on the FHA database all fall within the 95-percent confidence interval of the AHS estimate, indicating that any discrepancies in the volume estimates between the two data sources are not statistically significant. One noticeable anomaly is that the volume estimate for the 1998 cohort is considerably lower than the FHA benchmark, while the estimate for 1997 is substantially higher than any other years.

To further quantify how well the AHS has captured the mortgage volume information, we compute the coverage rate for each origination cohort using the point estimates. The coverage rate is defined as the volume derived from the AHS divided by the benchmark estimate. A coverage rate greater than 100 percent would mean that the AHS has identified more loan originations than the comparison database. Coverage rates for loans in the FHA market segment are presented in the first panel of Exhibit 4-4.

Exhibit 4-4:
Coverage of Home-Purchase Mortgage Originations in the AHS Surveys

MORTGAGE MARKET SEGMENT	ORIGINATION YEAR	AHS SURVEY YEAR	AHS COVERAGE ^C
FHA ^a	1989	1991	88%
	1990	1991	88%
	1991	1993	103%
	1992	1993	101%
	1993	1995	94%
	1994	1995	90%
	1995	1997	92%
	1996	1997	96%
	1997	1999	116%
	1998	1999	84%
	1999	2001	99% (99%)
	2000	2001	92% (93%)
	Average 1989-2000		95%
VA/RHS ^b	1993	1995	111%
	1994	1995	129%
	1995	1997	111%
	1996	1997	95%
	1997	1999	111%

³³ The FHA Single-Family Warehouse Data we obtained from HUD were organized by fiscal year. For the purpose of this study, we have reorganized the records by calendar year and excluded home improvement loans from the FHA database.

MORTGAGE MARKET SEGMENT	ORIGINATION YEAR	AHS SURVEY YEAR	AHS COVERAGE ^C
	1998	1999	112%
	1999	2001	116% (116%)
	2000	2001	120% (121%)
	Average 1993-2000		113%
Conventional Conforming ^b	1993	1995	120%
	1994	1995	108%
	1995	1997	115%
	1996	1997	97%
	1997	1999	114%
	1998	1999	89%
	1999	2001	79% (78%)
	2000	2001	83% (82%)
	Average 1993-2000		100%
Conventional Jumbo ^b	1993	1995	78%
	1994	1995	73%
	1995	1997	75%
	1996	1997	87%
	1997	1999	58%
	1998	1999	77%
	1999	2001	72% (71%)
	2000	2001	53% (58%)
	Average 1993-2000		72%
All Conventional ^b	1993	1995	116%
	1994	1995	105%
	1995	1997	111%
	1996	1997	96%
	1997	1999	107%
	1998	1999	88%
	1999	2001	79% (77%)
	2000	2001	79% (79%)
	Average 1993-2000		98%
All Segments ^b	1993	1995	113%
	1994	1995	107%
	1995	1997	108%
	1996	1997	97%
	1997	1999	110%
	1998	1999	89%
	1999	2001	86% (85%)
	2000	2001	83% (83%)
	Average 1993-2000		99%

Sources: Authors' tabulations of the 1991-2001 AHS, 1993-2000 HMDA, and 1989-2000 FHA administrative data.

Notes: Coverage ratio is defined as the mortgage originations in AHS (adjusted) divided by the benchmark estimates.

^aBased on data from the AHS and FHA administrative records for the whole nation.

^bBased on data from the AHS and HMDA files for loans made in metro areas.

^cLoan purpose (purchase vs. refinance) determination for the 1999 and 2000 AHS estimates is based on the REFI variable. Loan purpose estimates in parentheses are based on a comparison between mortgage origination year and house acquisition year as with other survey years.

We can see that the AHS has been fairly accurate in capturing home-purchase FHA loans, with coverage rates averaging at 95 percent for the period examined. On average, there is a tendency for the AHS to slightly understate the origination volume in this segment of the market. Coverage shoots up to 116 percent in 1997 and drops to 84 percent for loans originated in 1998. For the 1999 and 2000 cohorts, we present volume estimates using both the REFI variable and the old method of identifying loan purpose (by comparing the mortgage origination year with house acquisition year). Estimates based on the old method are shown in parentheses. They are by and large very close to the ones derived from the REFI variable.

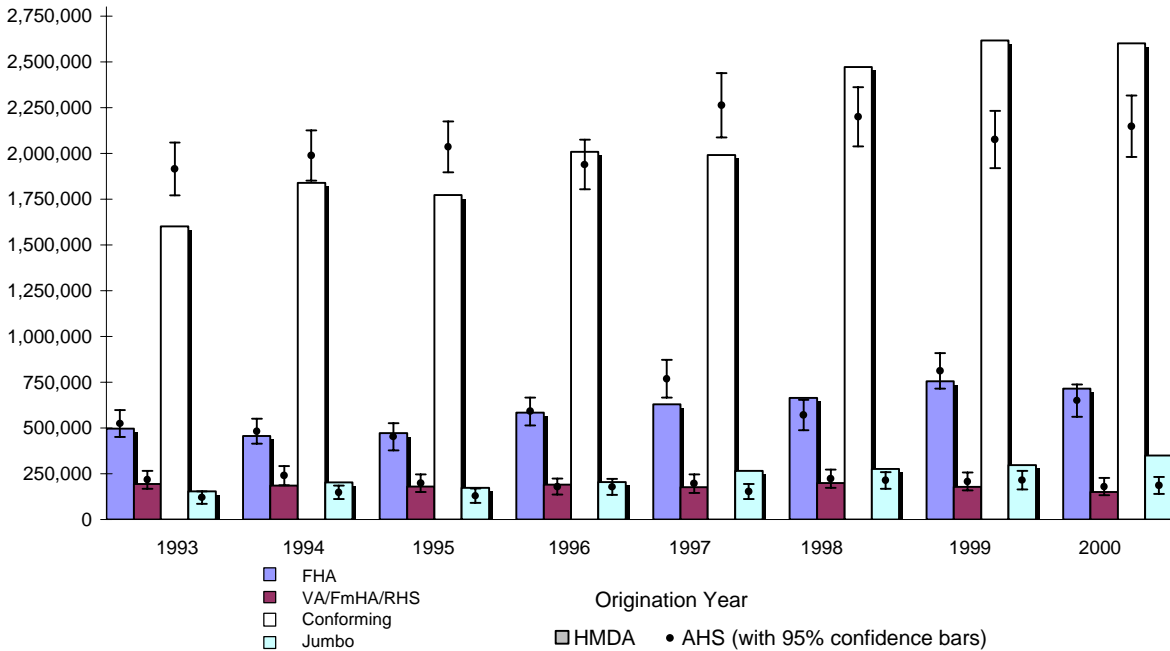
One potential contributing factor to the overall discrepancies is that mortgage origination date might be defined differently in the two databases. In the FHA administrative data, we defined origination cohorts according to the mortgage amortization date (first payment date) reported in files. This definition might not coincide with the mortgage origination year information provided by the borrowers in the AHS surveys. It is likely that most homeowners in the AHS considered closing date as the loan origination date.

Origination volumes of home-purchase mortgages in other segments of the market are benchmarked using the HMDA data. As noted above, one constraint for using the HMDA files is that comparisons that involve HMDA estimates need to be restricted to loans made in metro areas. In addition, given changes in lender reporting requirements, HMDA coverage of the market is fairly complete only since 1993. For the purpose of this study, loans with guarantees from VA and FmHA/RHS are grouped into a single market segment.³⁴ Results are presented in Exhibits 4-4 and 4-5. In general, findings are consistent with those for the FHA market segment. For VA/FmHA/RHS purchases, the volume estimates from AHS track fairly well with the HMDA estimates for most of the origination years examined, although there is a tendency for the AHS to overestimate mortgage volumes with coverage rates averaging 113 percent. Focusing on the point estimates, there appears to be a slight tendency for the AHS to overstate the origination volume of loans in this market segment. However, when the sampling variability of the AHS estimates is taken into account, HMDA benchmark estimates for the eight cohorts are all located within the 95-percent confidence interval of the AHS volume estimates. This implies that any discrepancies in the point estimates are not statistically significant at the 95-percent significance level.

Once again, the use of REFI variable in the 2001 survey in determining loan purpose yields volume estimates that are indistinguishable from the ones based on the old method (by comparing the mortgage origination year with house acquisition year). This is true for identifying home purchases in the other segments of the mortgage market.

³⁴ In 1995, the Farmers Home Administration (FmHA) was reorganized. Since this reorganization, the Rural Housing Service (RHS) has handled the mortgage functions formerly done by the FmHA.

**Exhibit 4-5:
Number of Mortgages by Origination Year and Market Segment
Home-Purchases in Metropolitan Areas
AHS vs. HMDA**



Conventional conforming mortgages represent the lion’s share of mortgage originations in each year. Coverage in the AHS for such originations appears to be more adequate for the earlier cohorts, 1993 to 1997. The coverage ratio drops below 90 percent for the last three years (1998, 1999, and 2000), ranging from 79 percent to 89 percent. Across the eight cohorts we have examined, the average coverage ratio in the AHS is 100 percent. In four of the eight cohorts, both end points of the confidence intervals of the AHS volume estimates are higher than the HMDA benchmarks. This could simply reflect the incomplete reporting of HMDA itself, rather than indicate inflated volume estimates from the AHS for this market segment. Analysts have pointed out that HMDA reporting has improved and become more complete in recent years. This may explain the drop in the AHS coverage ratio in the last three years we have examined. Another possible explanation could be because 2001 was such a big refinancing year that the proxy adjustment rate has failed to account for some of the terminated home-purchases that were originated in the two years prior to the interview.

Volume estimates for conventional jumbo purchases in the AHS are consistently lower than the HMDA benchmarks for all the cohorts, an indication that the AHS has been relatively unsuccessful in identifying such mortgages. Coverage is especially poor for the 2000 cohort. For that year, the AHS enumeration only reflects 53 percent of the originations reported in the HMDA data. Excluding the year 2000 estimate, coverage ratios are quite consistent over the decade, ranging from 72 percent to 87 percent. Across the eight cohorts, the average coverage ratio is 73 percent, implying that the AHS is able to capture only about three quarters of the originations of such mortgages.

Causes for the understatement of jumbo loan volumes in the AHS are unclear. It may be the result of homeowners understating the original loan amount in the conventional segment of the market, as will be discussed in the next section of this chapter. Recall that the distinction between conforming and jumbo conventional loans is based on the original loan amounts. An alternative explanation is that homeowners with jumbo mortgages are underrepresented in the AHS as a result of non-participation in the survey. Given the fact that participation in the AHS survey commits one to repeated long interviews, more wealthy homeowners may be less willing to participate. But such an undersampling problem should already be accounted for in the sampling weight variable provided in the public-use version of the AHS files.³⁵ It is also possible that, in a declining interest rate environment, borrowers with jumbo loans are particularly quick to refinance multiple times, which would not be accounted for in the proxy adjustment rates we used. This would explain the undercount of such mortgages in the AHS. A case in point are volume estimates derived from a big refinancing survey year such as 2001.

The lower two panels of Exhibit 4-4 present the AHS coverage for all conventional loans, jumbo and conforming conventional combined, as well as loans of all segments combined, using data from HMDA alone. Movements of the coverage ratio are fairly similar to those for the conventional conforming loans alone, since they account for the vast majority of originations in the mortgage market.

It is important to note that the observed coverage ratios of more than 100 percent for home-purchase loans in the conventional and VA/FmHA/RHS market segments do not necessarily mean that the AHS has overestimated the volume for such mortgages. Instead, the issue may be that our benchmark database – the HMDA files – does not represent a complete census of loans originated. Even though for those market segments we have limited our analysis to loans made in metro areas where the enumeration is known to be more complete, it is still unlikely that the HMDA data have captured all the originations in those areas. This is especially true for some of the earlier cohorts we examined. There are two reasons why that is the case. First, reporting for certain mortgages and lenders are not required under the HMDA statutes.³⁶ Second, some lenders do not comply with HMDA. Two recent studies, Scheessele (1998) and Berkovec and Zorn (1996), have carefully documented the coverage of the HMDA data.³⁷ HMDA was found to have an average coverage rate of around 90 percent for FHA loans over the period 1993-1996, while its coverage of loans sold to the two GSEs has ranged from 76 to 93 percent for the same period. Another concern with HMDA as a comparison database was that

³⁵ The Census Bureau assumes that the housing units missed are similar in some ways to the units interviewed in the AHS. The “noninterview adjustment” on the sampling weights takes into account factors such as geographic location, structure type, tenure, number of units, number of rooms, and occupancy status. For details, see Appendix B of *Current Housing Reports, Series H150/01, American Housing Survey for the United States: 2001*, U.S. Census Bureau, 2002.

³⁶ For example, lenders are required to report only if their dollar volume of home purchase or refinance mortgages exceeds 10 percent of their total loan originations. Thus, there may be lenders that originate a large volume of home mortgages, but do not have to report because the dollar volume of non-mortgage loans (such as personal or business loans) exceeds 90 percent of their total loan originations.

³⁷ Randall M. Scheessele “HMDA Coverage of the Mortgage Market,” HUD PD&R Housing Finance Working Paper No. 7, July 1998. Jim Berkovec and Peter Zorn, “How Complete is HMDA?: HMDA Coverage of Freddie Mac Purchases.” *Journal of Real Estate Research*, 1996: 39-55.

the broader definition of metro areas used compared to the AHS would result in higher estimates of mortgage activity. This bias does not appear to be evident in our findings.

Another potential source of bias is that the AHS volume estimates for conventional loans are adjusted using historical termination rates from FHA loans. In general, FHA loans have substantially higher claim rates than conventional loans, although their prepayment rates are lower. Our adjustment factors might have inflated the volume estimates too much for loans in the non-FHA segments of the market.

Given the problems with the HMDA data and adjustment rates, the FHA data should provide the most reliable benchmark of loan volumes reported in the AHS. The fact that the AHS provides reasonable estimates of overall volumes of FHA purchase mortgages suggests that the AHS is fairly accurate at capturing purchase mortgage originations.

Finally, the general decline in AHS coverage ratios in the last three cohorts deserves further investigation. This drop in coverage happens in both the FHA and conventional conforming segments of the mortgage market. Another anomaly is that the AHS volume estimate for 1997 is noticeably higher than the FHA benchmark, while the estimate for 1998 is considerably lower than for any other years in that market segment. This pattern resembles the discrepancies we found in the earlier release of the 1999 AHS public-use data – coverage ratio for FHA loans shoots up to 110 percent (compared to 116 percent using the latest data) in 1997 and drops to 57 percent (compared to 84 percent using the latest data) in 1998. According to the Census Bureau, the previous discrepancies most likely resulted from a coding error with the house acquisition year variable (WHNGET) when the 1999 interview records were first processed.³⁸ The WHNGET variable identifies the year that the survey respondent purchased/obtained/received his or her house. This is a crucial data element in many AHS mortgage analyses, since it is used to identify the origination year for active mortgages that were obtained at the same time of the home purchase. The similarity between our previous and current estimates makes us question whether the coding error has been completely eliminated in the latest release (dated November 14, 2002) of the 1999 data.

To shed additional light on these anomalies, we examine whether the AHS has captured home purchase activity, not simply mortgage activity, adequately over the decade. For independent benchmarks, we make use of the annual sales volume estimate of existing single-family homes published by the National Association of Realtors (NAR) and the sales of new homes estimated from the Survey of Construction (SOC) data. These estimates are published periodically in the *U.S. Housing Market Conditions* reports by the Office of Policy Development & Research, U.S. Department of Housing and Urban Development. In Exhibit 4-6, these benchmark estimates are compared to the total volume of homes bought/obtained/received (regardless of mortgage originations) derived from the AHS. It is clear that the AHS estimates track the benchmark figures very well until 1998. The coverage ratio ranges from 98 percent to 104 percent. In the last three years, coverage drops to the 84-85 percent range. This comparison appears to confirm our previous findings from the FHA and HMDA data that the AHS understates mortgage origination volumes in 1998, 1999 and 2000.

³⁸ Barbara T. Williams of the Census Bureau provided assistance on this issue.

These lower coverage ratios could be the result of the failure of the proxy adjustment rates to account for loans that have terminated prior to the interview, because 2001 was such a big refinancing year. Many borrowers were able to refinance their mortgages multiple times in a declining interest rate environment.

**Exhibit 4-6:
Comparison of Total Home Acquisitions Between AHS, NAR and SOC Data (In Thousands)**

AHS SURVEY YEAR	YEAR	AHS HOMES ACQUISITIONS	NEW AND EXISTING HOME SALES BENCHMARK	AHS COVERAGE
1991	1989	3,971	3,996	99%
1991	1990	3,805	3,745	102%
1993	1991	3,815	3,729	102%
1993	1992	4,218	4,130	102%
1995	1993	4,358	4,468	98%
1995	1994	4,823	4,616	104%
1997	1995	4,531	4,479	101%
1997	1996	4,999	4,953	101%
1999	1997	5,170	5,185	100%
1999	1998	4,916	5,856	84%
2001	1999	5,190	6,085	85%
2001	2000	5,151	6,029	85%

Sources: Authors' tabulation of 1991-2001 AHS; Survey of Construction; National Association of Realtors data.

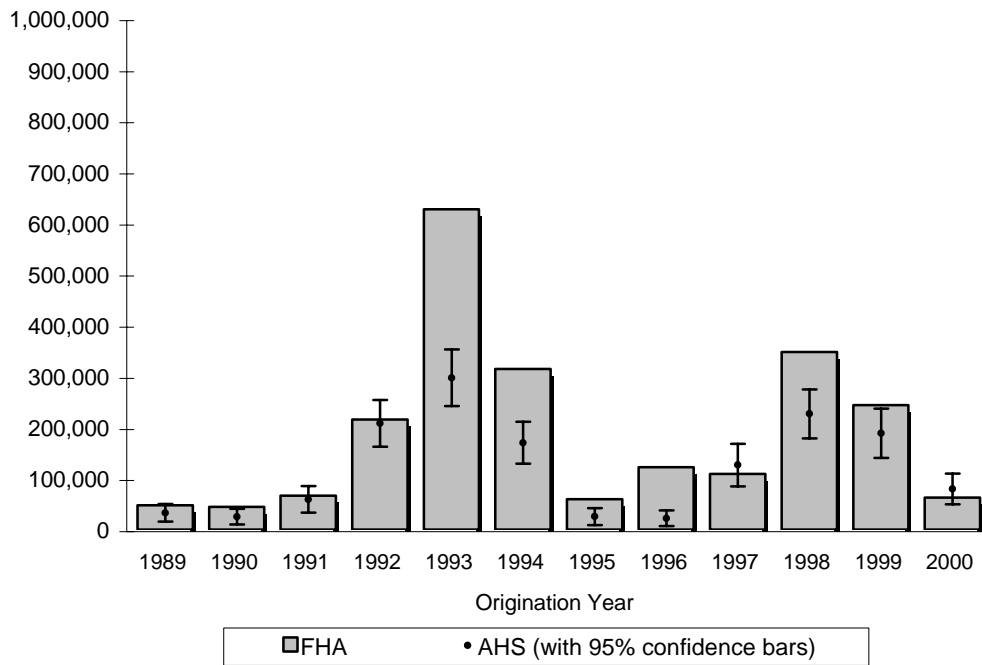
Notes: Tabulations from the AHS are based on the WHNGET variable and include mobile homes.

Refinance Mortgages

Before the 2001 survey, respondents were not asked explicitly whether there was a refinance loan in place at the time of the AHS interview. Rather, AHS users infer which loans were refinancings when the year of origination is not the same as the year of house purchase. In the 2001 survey, the borrower was asked explicitly whether the mortgage was a refinancing of a previous mortgage (i.e., the REFI variable); in other words, refinance loans can be identified in the data without comparing the mortgage origination year and house purchase year. Key issues of the replication analysis are whether the pre-2001 survey design permits reliable estimates of refinance volumes for loans in different segments of the mortgage market, and whether the change in the 2001 survey resulted in more accurate volume estimates.

Exhibit 4-7 compares the volume of refinance originations with FHA insurance for the period from 1989 to 2000, using the AHS and FHA administrative data. Coverage ratios are reported in the first panel of Exhibit 4-8. Estimates for the 1999 and 2000 cohorts are based on the REFI (refinance) variable. For comparison purposes, numbers in parentheses represent refinances identified using the old methodology.

**Exhibit 4-7:
Number of FHA-Insured Mortgages
Refinance Originations
AHS vs. FHA**



**Exhibit 4-8:
Coverage of Refinance Mortgage Originations in the AHS Surveys**

MORTGAGE MARKET SEGMENT	ORIGINATION YEAR	AHS SURVEY YEAR	AHS COVERAGE ^C
FHA ^a	1989	1991	69%
	1990	1991	57%
	1991	1993	90%
	1992	1993	96%
	1993	1995	47%
	1994	1995	54%
	1995	1997	43%
	1996	1997	18%
	1997	1999	116%
	1998	1999	65%
	1999	2001	77% (76%)
2000	2001	126% (119%)	
	Average 1993-2000		72%
VA/RHS ^b	1993	1995	58%
	1994	1995	56%
	1995	1997	20%
	1996	1997	5%

MORTGAGE MARKET SEGMENT	ORIGINATION YEAR	AHS SURVEY YEAR	AHS COVERAGE ^C
	1997	1999	96%
	1998	1999	54%
	1999	2001	96% (95%)
	2000	2001	392% (373%)
	Average 1993-2000		97%
Conventional Conforming ^b	1993	1995	64%
	1994	1995	85%
	1995	1997	22%
	1996	1997	14%
	1997	1999	66%
	1998	1999	42%
	1999	2001	44% (53%)
	2000	2001	39% (51%)
Average 1993-2000		47%	
Conventional Jumbo ^b	1993	1995	40%
	1994	1995	42%
	1995	1997	10%
	1996	1997	17%
	1997	1999	36%
	1998	1999	45%
	1999	2001	37% (40%)
	2000	2001	58% (51%)
Average 1993-2000		36%	
All Conventional ^b	1993	1995	62%
	1994	1995	81%
	1995	1997	21%
	1996	1997	15%
	1997	1999	63%
	1998	1999	42%
	1999	2001	43% (52%)
	2000	2001	40% (51%)
Average 1993-2000		46%	
All Segments ^b	1993	1995	62%
	1994	1995	82%
	1995	1997	22%
	1996	1997	15%
	1997	1999	66%
	1998	1999	44%
	1999	2001	48% (56%)
	2000	2001	45% (55%)
Average 1993-2000		48%	

Sources: Authors' tabulations of the 1991-2001 AHS, 1993-2000 HMDA, and 1989-2000 FHA administrative data.

Notes: Coverage ratio is defined as the mortgage originations in AHS (adjusted) divided by the benchmark estimates.

^aBased on data from the AHS and FHA administrative records for the whole nation.

^bBased on data from the AHS and HMDA files for loans made in metro areas.

^cLoan purpose (purchase vs. refinance) determination for the 1999 and 2000 AHS estimates is based on the REFI variable. Loan purpose estimates in parentheses are based on a comparison between mortgage origination year and house acquisition year as with other survey years.

Volume estimates from both data sources appropriately reflect the refinancing booms in 1993 and 1998, when mortgage interest rates were relatively low. However, the Exhibits clearly show that the ability of AHS to capture FHA refinances has been inconsistent across survey years. Coverage ratios vary widely. It appears that the reporting is relatively complete for the 1991 (90 percent) and 1992 (96 percent) cohorts. Volume estimates for those years are based on the 1993 survey data. The AHS reports a higher origination volume than the FHA data for loans in 1997 and 2000. The FHA benchmarks all fall within the 95-percent confidence interval of the AHS estimates for those four years. Coverage is inadequate for the other years. The underreporting became substantially worse in the 1995 and 1997 surveys, which are used to estimate volumes for the 1993 to 1996 cohorts of loans. For example, as shown in Exhibit 4-8, relative to the volume information calculated from the FHA data, the AHS captures only 47 percent and 54 percent of the originations in 1993 and 1994. For loans in the 1996 cohort, less than one-fifth (18 percent) have been captured in the AHS. Overall, the survey has done a poor job in capturing FHA refinances, with coverage rates averaging 72 percent over the decade.

Compared to the old method of identifying refinances, the new design in the 2001 survey (REFI variable) appears to yield volume estimates that are slightly closer to the FHA benchmarks.

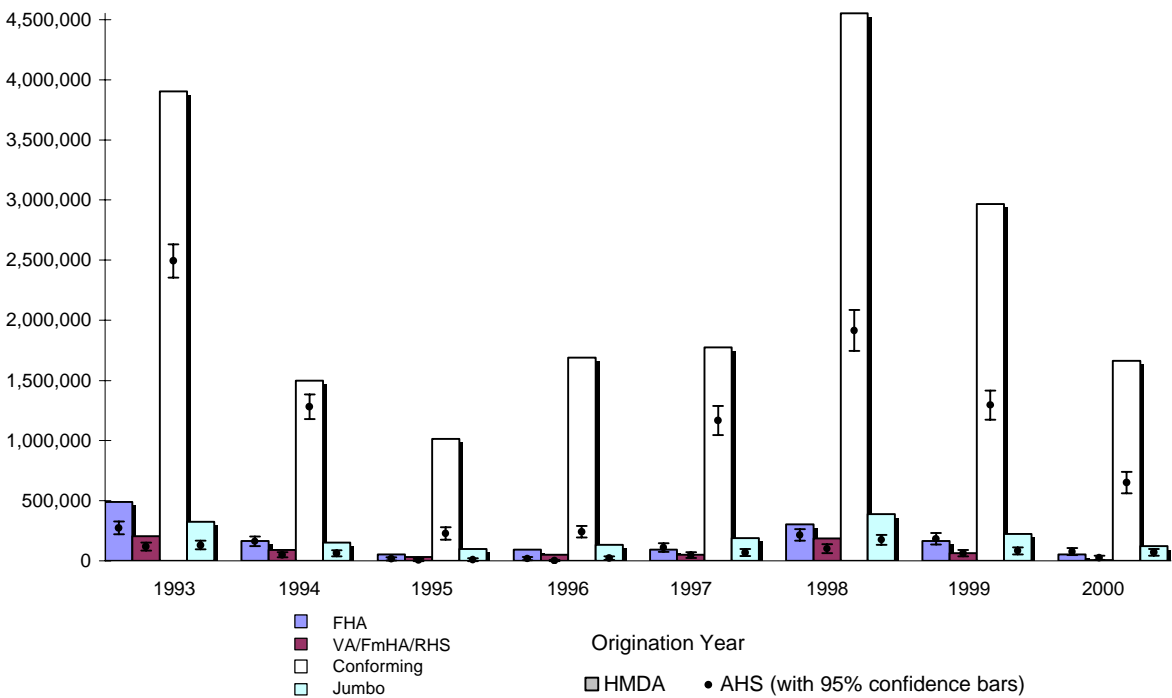
Data from the HMDA files are used to benchmark the volumes of refinances originated in the non-FHA segments of the market. Once again, our analyses are restricted to loans made in metro areas over the 1993 to 2000 period. Comparisons between the AHS estimates and HMDA benchmarks by market segment are presented in Exhibit 4-9. Exhibit 4-8 reports the coverage ratios. They indicate that the AHS has done a poor job of reporting refinance originations in those market segments. Focusing on conventional conforming refinances, the average coverage ratio for is 47 percent across the eight cohorts we have examined, implying that only around half of such originations are captured. Reporting is especially incomplete for loans originated in 1995 and 1997. Coverage rates show that only a quarter of the originations are reported in the AHS for those years. Reporting appears to be relatively complete for loans originated in 1994. The survey captures 85 percent of the originations in that year. Surprisingly, there appears to be no noticeable improvement in terms of coverage for this market segment as a result of the design change in the 2001 survey (questions added for identifying loan purpose). As shown in parentheses in Exhibit 4-8, the old methodology in determining loan purpose is able to yield volume estimates that are closer to the benchmarks. One possible explanation is that 2001 was such a big refinancing year that the proxy adjustment rates have failed to account for many of the loans that had terminated prior to the interview. The REFI variable may perform better in future survey years.

The reporting in AHS for refinances in the conventional jumbo segment of the market is even worse, with coverage ratio averaging just 36 percent for the years examined. In particular, coverage ratios are lowest for 1995 (10 percent) and 1996 (17 percent).

Benchmarking the volume estimate of refinances in the VA/FmHA/RHS market segment is challenging, given the relatively small number of such mortgages originated each year. Volume estimates derived from the AHS data are all associated with very wide confidence intervals, as shown in Exhibit D-5-9 of Appendix D. This implies that it is difficult to draw any statistically valid

comparisons between the AHS estimate and HMDA benchmark for this market segment. Focusing on the point estimates, reporting in AHS for the 1997 cohort appears to be very complete (96 percent). For loans originated in 2000, however, the AHS reports almost four times the number of VA/RHS refinances captured in the HMDA data. This anomaly most likely results from an underreporting of such mortgages in HMDA. The number of originations plummeted unexpectedly from 65,775 in 1999 to 6,243 in 2000.

**Exhibit 4-9:
Number of Mortgages by Origination Year and Market Segment
Refinances in Metropolitan Areas
AHS vs. HMDA**



Given the fact that the HMDA data itself is known to capture less than 100 percent of the mortgage activity in these market segments, it is quite likely that the *actual* coverage ratios for these refinances in the AHS are even worse than the comparison with the HMDA volumes has suggested here.

The fact that the AHS has not been able to capture many of the refinance originations over the decade is puzzling and deserves further examination. This undercounting problem appears to be prevalent across all mortgage market segments. Evidence offered by the FHA administrative data, as discussed above, has indicated that the share of refinances captured by AHS has varied substantially over time

and has deteriorated markedly since the 1995 survey (i.e., for refinances in the 1993-1998 cohorts).³⁹ One possible explanation may be the introduction of the set of home equity loan questions in the 1995 survey, which were then expanded in the 1997 survey. The questions do not provide a clear guidance to the homeowners regarding the distinction between lump-sum home equity loans and mortgages that are used to refinance an existing loan in order to take out housing equity (so-called cash-out refinances). As a result, some survey respondents might have incorrectly classified their refinance mortgages as home equity loans. To test this hypothesis, one would want to compare the total originations of home equity loans and refinances combined in a given year in the AHS to the volume of refinances reported in the benchmark databases. However, this comparison is not possible, since the survey does not ask for the origination year for home equity loans.

Coverage ratios are especially low for refinances in the 1995 and 1996 loan cohorts, which are derived from the 1997 AHS. This is true across all segments of the market. While volume estimates for the other cohort years are low relative to the benchmarks, the underreporting for these two years is dramatic. One possible explanation is that the expansion of home equity loan questions in the 1997 survey resulted in more cash-out refinance loans being reported as lump-sum home equity loans. However, this explanation would not account for the rebound in refinance volumes found by the 1999 survey. Another change in the 1997 survey is that a series of new questions related to reverse annuity mortgages and Home Equity Conversion Mortgages (HECMs) was inserted in the survey before the questions on any other mortgage types. This ordering of questions might have degraded the information collected on other, more common mortgage types. (The reverse mortgage questions were dropped in the 1999 survey but were added back to the end of the mortgage module in the 2001 survey.) A third explanation would be that there was some type of coding error in the processing of survey responses. While there is no clear explanation, given the consistent problems with refinance loan volumes, it does seem clear that the lack of a clear distinction between regular mortgages and home equity loans is a significant contributing factor.

Loan Characteristics

This section presents the results of replication analysis on mortgage characteristics estimated from the AHS. The approach we have taken represents a departure from the previous section. Rather than focusing on the number of loans for each origination cohort, the focal point of our analysis here is the distribution of loan characteristics reported in the data. In other words, we examine the shares of mortgages associated with different characteristics for loans captured in the AHS. This approach provides a test of whether the AHS accurately reports loan characteristics even if it does not accurately report loan volumes. These share estimates are compared to those derived from other mortgage market databases by origination cohort.⁴⁰ For origination amount, LTV and interest rate,

³⁹ This pre- and post-1993 comparison of coverage ratios cannot be made among cohorts of loans associated with other non-FHA market segments because reliable and complete HMDA data are not available for the pre-1993 era.

⁴⁰ We report both the point estimates and confidence intervals for the AHS share estimates and examine whether the benchmark estimates fall within the AHS confidence intervals. Alternatively, to assess the impact of sampling variability in such comparisons, one could conduct Chi-Square statistical tests.

which are continuous variables, we also present the distribution (mean, median, 25th, 75th percentile, and standard deviation) of the numeric values in Appendix D.

Because only surviving loans are observed in the survey, a challenge in this evaluation is that the differential prepayment speeds associated with different borrowers and loan types may lead to biased estimates with respect to some of the mortgage characteristics. If there are discrepancies between the estimates reported in AHS and the comparison data set, they cannot be easily allocated between measurement errors and biases resulting from differential prepayment rates.

The variables we have examined are:

- Mortgage market segment;
- Original mortgage amount;
- Mortgage interest rate (fixed-rate mortgages only);
- Loan-to-value (LTV) ratio at origination; and
- Mortgage payment plan.

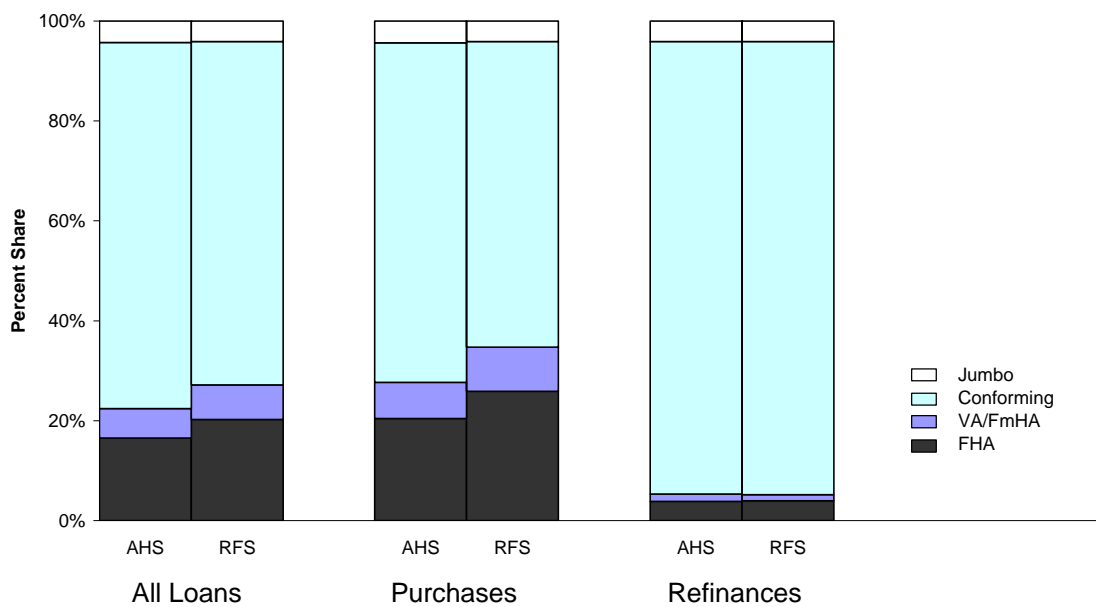
In light of our findings that the AHS does a poor job of capturing refinance loans, our analysis of the loan characteristics is performed on home-purchase and refinance loans separately. Whenever sample sizes are sufficient, tabulations are also stratified by mortgage market segment.

Mortgage Market Segment

For the purpose of this study, the four mutually exclusive mortgage market segments are FHA, VA/FmHA/RHS, conventional conforming, and conventional jumbo. Evaluating the reliability of the distribution of this information collected in AHS requires a mortgage market database that covers loans originated from all these four market segments. Of the available comparison databases, the RFS is best suited for this purpose since it covers all market segments, is presumed to provide accurate information, and has a large enough sample size to provide fairly precise estimates of market shares. However, as has been noted, the RFS suffers from only crude reporting of the year of origination and from only being available for a single year more than a decade ago. Given the limitations of the RFS, we use the HMDA data as the primary comparison database for mortgage market segments, given its coverage of all mortgage market segments and its availability for multiple years. Because of its incomplete coverage for loans originated by financial institutions in non-metro areas, all comparisons that involve HMDA are limited to loans made in metro areas in both data sets. Additional benchmark estimates are provided by the 1995 SCF data, which also contain mortgages from all market segments. However, given its small sample size, consequently lower statistical reliability, and its availability for only a single year, the usefulness of the SCF for comparison to the AHS is limited.

We begin our discussion with the analyses using the RFS. Exhibit 4-10 compares the share of loans by market segment between the AHS and RFS data for mortgages originated between 1989 and 1991.⁴¹

**Exhibit 4-10:
Shares of Loans by Mortgage Market Segment - AHS vs. RFS
1989-1991 Origination Cohort**



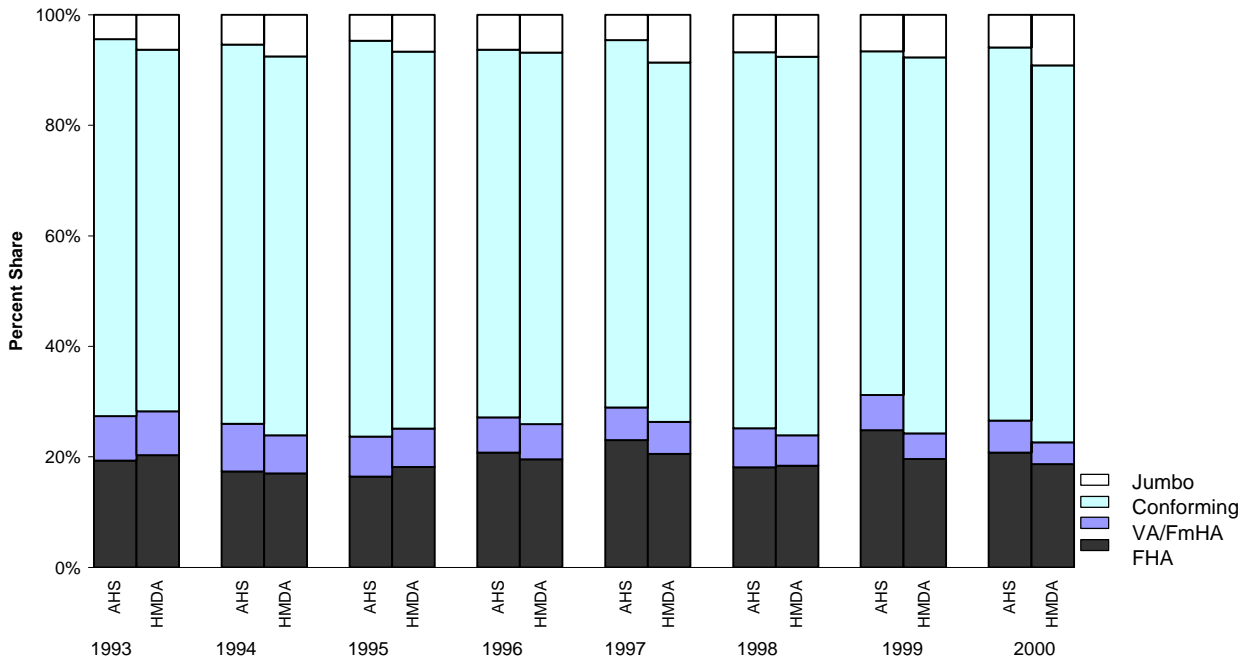
The RFS is particularly valuable as a point of comparison since information on mortgages (in this case, mortgage insurance status) is collected directly from lenders. These data elements should be quite accurate, since the lenders' information is likely supported by administrative records (rather than the homeowners' memory, as in data based only on household surveys.) In addition, the RFS covers a time period for which reliable HMDA data are not available. Exhibit 4-10 shows that among the loans identified in the two data sources, the AHS contains a slightly smaller share of loans with FHA insurance relative to the RFS file. At the same time, there is a larger proportion of conventional loans in the AHS. This indicates that some of the survey respondents might have misclassified their FHA mortgages as conventional products. The higher share in the AHS could also stem from an over inflation of conventional loan volume estimates, because we used the FHA termination rate as an adjustment factor. Further examination of the 95-percent confidence interval estimates from the two data sets, provided in Exhibit D-10 of Appendix D, reveals that the underestimate of FHA loans (and overestimate of conventional loans) in terms of shares is statistically significant, since the interval estimates from the two data sets do not overlap. On the other hand, the market shares in the two data sources for VA/FmHA/RHS and conventional jumbo loans are close. These relationships hold for the entire cohort of 1989-1991 originations as well as for just home-purchase loans. For refinance loans originated in that time period, the share estimates of each market segment are almost identical

⁴¹ For confidentiality reasons, the public-use version of the RFS file only identifies mortgage origination year information in interval format.

between the AHS and RFS. Consistent with our expectation, the vast majority of refinance loans are in the conventional market segment.

The HMDA data are next used to benchmark the AHS estimates of market segment shares for loans in the 1993 to 2000 cohorts. Exhibits 4-11 and 4-12 present the analysis, separately for home-purchase and refinance loans.

**Exhibit 4-11:
Shares of Loans by Mortgage Market Segment - AHS vs. HMDA
1993 to 2000 Origination Cohorts, Home-Purchase Loans**



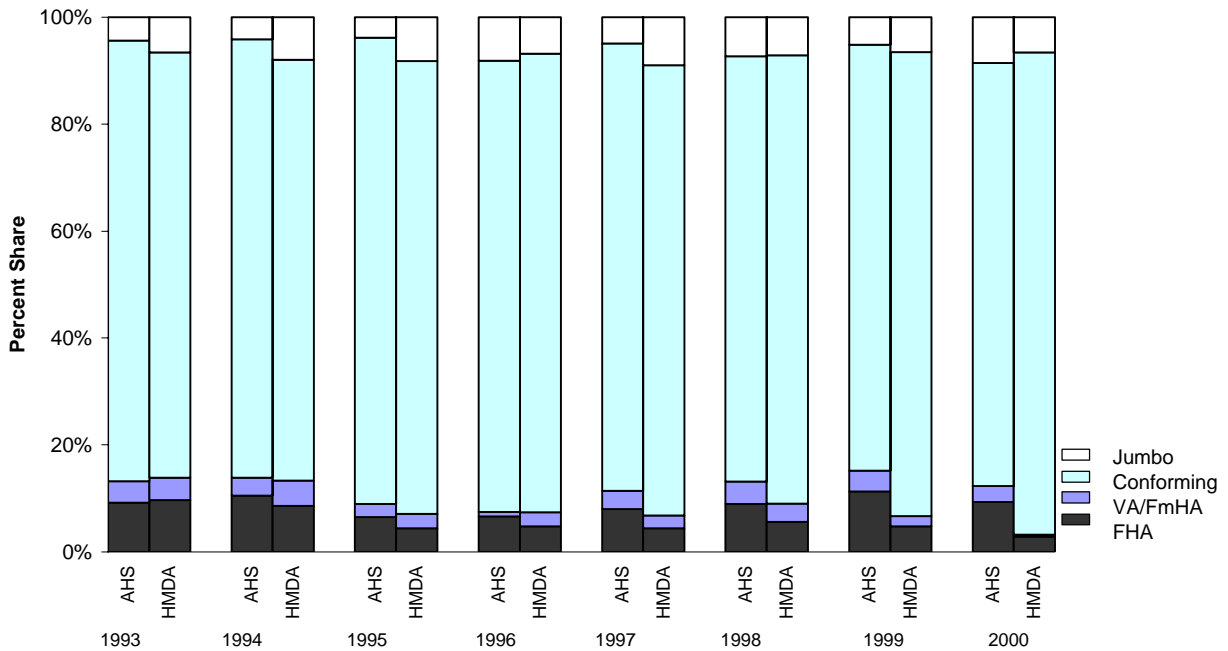
For home purchase originations, the distribution of loans across the four market segments appears to be fairly stable over the 1993-2000 period, according to the HMDA data. The market share estimates derived from the AHS match remarkably well with the HMDA benchmarks. FHA loans have an average market share of 20 percent over this period in the AHS, compared to an estimate of 19 percent from HMDA. Both HMDA and AHS indicate that, on average, 7 percent of the annual originations are in the VA/FmHA segment of the market. While share estimates for some of the years are different between the two data sets, the differences are not statistically significant at the 95 percent level for vast majority of the years examined. Confidence interval estimates are reported in Exhibit D-11-12 of Appendix D.

Over the same time period, conventional conforming loans have an average market share of 68 percent in the AHS, compared to an estimate of 67 percent in HMDA. Among the years we have examined, the AHS appears to slightly overstate the share of such originations in three years, while the share is understated in one year. None of these discrepancies between the two data sets is statistically significant when sampling variability is taken into account.

It is, however, noteworthy that the AHS contains a smaller share of jumbo loans in every year when compared to the HMDA benchmarks (although the differences are not statistically significant at the 95-percent confidence level in 1993, 1996, 1998 and 1999). Over the 1993-2000 period, the average market share is 5 percent in the AHS, compared to 7 percent reported in the HMDA data. This underestimate of jumbo shares in AHS may in part be due to underreported loan amounts in the conventional market, given that in three of the years the AHS share of conventional conforming loans exceeds the share in HMDA to compensate for the undercounting of jumbo loans. An alternative explanation is that homeowners with jumbo mortgages are underrepresented in the survey as a result of non-participation. Given the fact that participation in the AHS survey commits one to repeated long interviews, more wealthy homeowners may be less willing to participate. It is also possible that, in a declining interest rate environment, those with jumbo loans are particularly quick to refinance their mortgages multiple times, which would not be captured in the AHS.

We now turn our attention to the comparison for refinance loans, as shown in Exhibit 4-12.

Exhibit 4-12:
Shares of Loans by Mortgage Market Segment - AHS vs. HMDA
1993 to 2000 Origination Cohorts, Refinance Loans



Because of the relatively small number of refinance loans captured in the AHS, many of the share estimates are associated with rather wide confidence intervals. This makes it difficult to draw definite conclusions from the estimates. For example, the AHS indicates that 7 percent of refinance loans originated in 1995 had FHA insurance, compared to an estimate of 4 percent from HMDA. While the difference between these two estimates is large (the AHS share is nearly twice as large as the HMDA share), the 95-percent confidence interval for the AHS estimate is very wide (from 1 to 12 percent)

and contains the HMDA benchmark. Thus, the difference in the two estimates is not statistically significant. By themselves, the wide confidence bounds on these estimates are an indication that the AHS is not a particularly reliable measure of market shares for refinance loans.

Focusing on just the point estimates provided by the AHS, there is a general tendency for the AHS to overstate the share of refinance loans that are FHA insured and to understate the share of conventional jumbo loans. Over the 1993-2000 period, the average market share for FHA refinances in the AHS is 8.8 percent, compared to 5.6 percent in HMDA. Jumbo loans have an average market share in the AHS of 5.8 percent, compared to 7.3 percent in HMDA. Overall, the estimates for conventional conforming and VA/FmHA loans are fairly close in the two data sets. Conventional conforming loans average 82.3 percent in the AHS compared to 84.2 percent in HMDA, while VA/FmHA loans average 3.1 percent in AHS compared to 2.8 percent in HMDA.

We have also compared the market share estimates from AHS and SCF for the 1993 and 1994 cohorts, separately for home purchases and refinances. Results are shown in Exhibits D-11-12a and D-11-12b of Appendix D. When compared to the SCF, the AHS appears to understate the share of FHA loans and overstate the share of conventional conforming mortgages for both home purchase and refinance loans. This is at odds with our finding derived from the HMDA data for these two origination cohorts where the AHS share estimates appear to match well with the HMDA benchmarks.⁴² We believe that the HMDA benchmarks should be more accurate than estimates based on the SCF because of the small sample size of the SCF data. In addition, given our finding in the previous section (Mortgage Origination Volume) that the AHS volume estimates for FHA home-purchases are very close to the benchmarks derived from the FHA administrative data, it is unlikely that the AHS has understated the share of FHA loans as suggested by the SCF.

Original Mortgage Amount

In evaluating the original mortgage amounts reported in the AHS, the primary goal is to check whether the distribution of the variable mirrors that calculated from other mortgage market databases. Tabulations are done separately for loans in different mortgage market segments and also stratified by loan purpose.

Although not fully documented in the latest version of the AHS codebook, the original mortgage amount variable provided in the public-use version of the AHS file is top-coded for confidentiality reasons. For all the survey years examined in this study, the variable appears to be top-coded at \$350,000, except in the 1995 survey, for which it is capped at \$375,000. This implies that, in the AHS, any distribution of the mortgage amount information calculated for loans in the jumbo market segment will not be accurate. (For completeness, we include estimates of such loans in the Exhibits of Appendix D.)

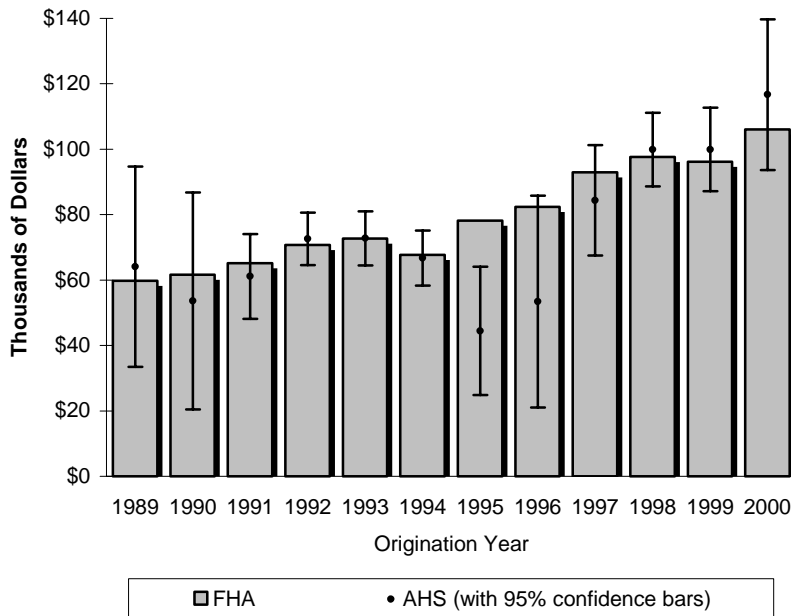
⁴² To make the SCF and HMDA analyses comparable, one would want to limit the SCF records to loans made in metro areas. Unfortunately, the public-use version of the SCF file does not allow us to do so. Even if such analysis were possible, it seems unlikely that this would account for the inconsistent finding between HMDA and SCF.

For FHA-insured mortgages, we use benchmark estimates derived from the FHA administrative database, given the thoroughness of its coverage of this type of originations. Again, HMDA data are used to benchmark estimates for the other three mortgage segments because of HMDA's broad coverage and availability of information for recent originations.

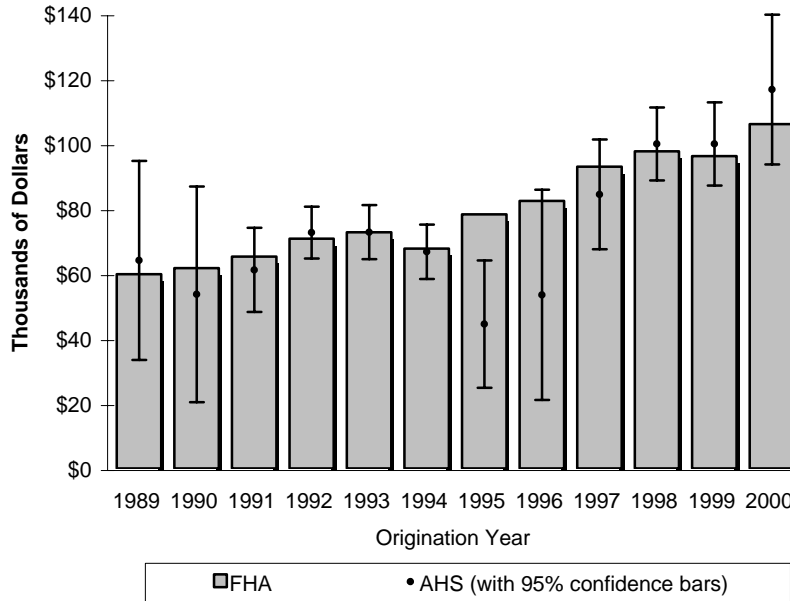
Comparisons for loans in the FHA market segment are presented in Exhibits 4-13 and 4-14. The two charts display the average mortgage amount for each origination cohort of loans separately for home-purchase and refinance loans, using the AHS and FHA administrative data. For home-purchases, according to the FHA data, the average loan amount rose steadily, from \$64,000 to \$105,000, over the decade. The AHS averages track extremely well with the FHA averages, which lie consistently within the 95-percent confidence interval estimates from the AHS. From 1989 to 1992 there was a tendency for the AHS estimate to be slightly higher the FHA records, while from 1993 to 1998 there was a tendency for the AHS to be slightly lower than FHA (with the exception of 1995). Over the entire period the discrepancy is never greater than \$4,000. Estimates for the 1999 and 2000 cohorts of loans are virtually identical between the two data sets.

Additional comparisons of other distribution statistics (the 25th, 50th, and 75th percentiles) provided in Exhibit D-13-14 of the Appendix also confirm that the AHS variables have in fact captured the loan amount information very accurately for mortgages in this segment of the market.

Exhibit 4-14:
Average Original Mortgage Amount of FHA-Insured Loans
AHS vs. FHA
Refinances Only



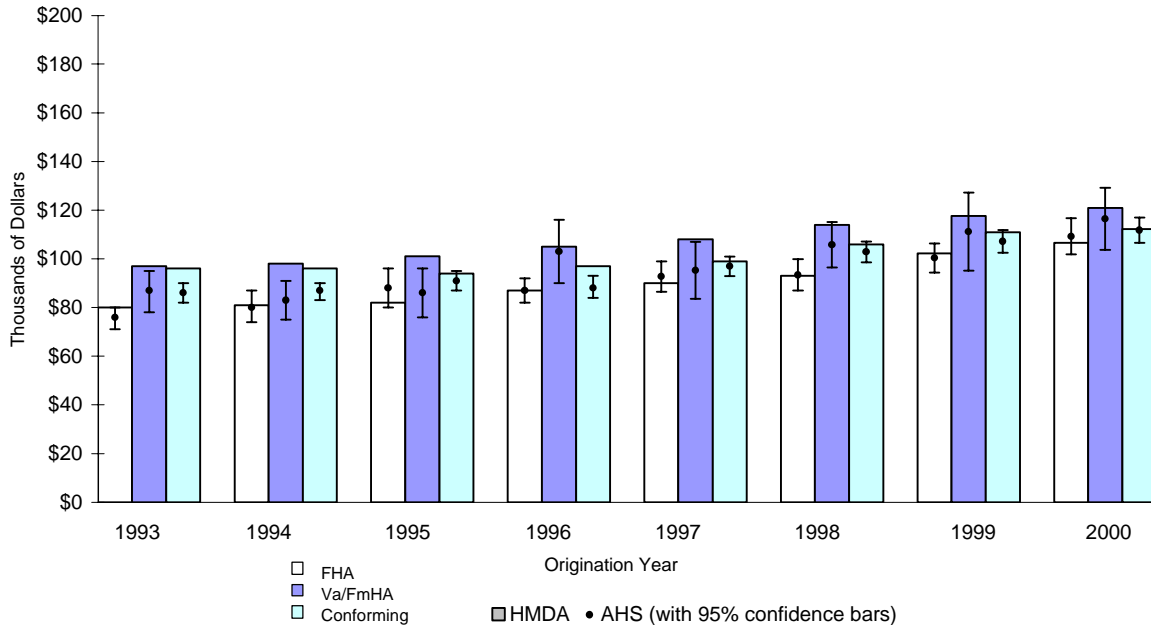
**Exhibit 4-14:
Average Original Mortgage Amount of FHA-Insured Loans
AHS vs. FHA
Refinances Only**



Test results are mixed for the FHA refinance loans. The first thing to note is that, because of the small number of such loans captured in the AHS, the 95-percent confidence interval estimated for the average loan amount variable is very wide, especially for data from the 1991 survey (1989 and 1990 originations) and 1997 survey (1995 and 1996 originations). Therefore, readers should exercise caution when interpreting the findings that are based on comparisons of the point estimates with external benchmarks. As Exhibit 4-14 shows, it appears that AHS has seriously underestimated the origination amount of such loans in the 1995 and 1996 cohorts. For the other origination cohorts, the point estimates for the mean loan amount are reasonably close to the FHA benchmarks. Excluding these two cohorts, over the entire period the average loan amount in the AHS is \$79,170 compared to \$79,340 in the FHA data.

Replication analyses on loan amounts in the other market segments are based on HMDA data. Exhibits 4-15 and 4-16 present the average loan amounts derived from the AHS and HMDA for loans made in metro areas between 1993 and 2000, stratified by mortgage market segment and loan purpose. The pattern between the AHS and HMDA estimates for FHA average loan amount is quite similar to the comparison between AHS and FHA. Overall, the AHS averages are quite close to the HMDA averages, and the AHS average falls short of the HMDA average in 1993 and exceeds the HMDA average in 1995, just as was true of the AHS-FHA comparison. This consistency of findings on FHA loan amounts suggests that HMDA may be fairly accurate in reporting loan amounts. With regard to the other market segments, the comparisons show that AHS homeowners tend to under-report loan amount of originations in both the VA/FmHA and conventional conforming segments.

**Exhibit 4-15:
Average Original Mortgage Amount by Mortgage Segment
AHS vs. HMDA
Loans Originated in Metro Areas, Home Purchase Originations Only**



Focusing on home-purchases in the conventional conforming market, further examinations reveal that the AHS estimates track relatively well with the HMDA benchmarks for loans in the 1995, 1997, 1998, 1999 and 2000 cohorts, for which the AHS interval estimates contain the HMDA average loan amounts. For loans in the 1993, 1994, and 1996 cohorts, however, the AHS has provided mean estimates that are slightly lower (by approximately 10 percent) than the HMDA benchmarks. The differences are statistically significant at the 95-percent level.

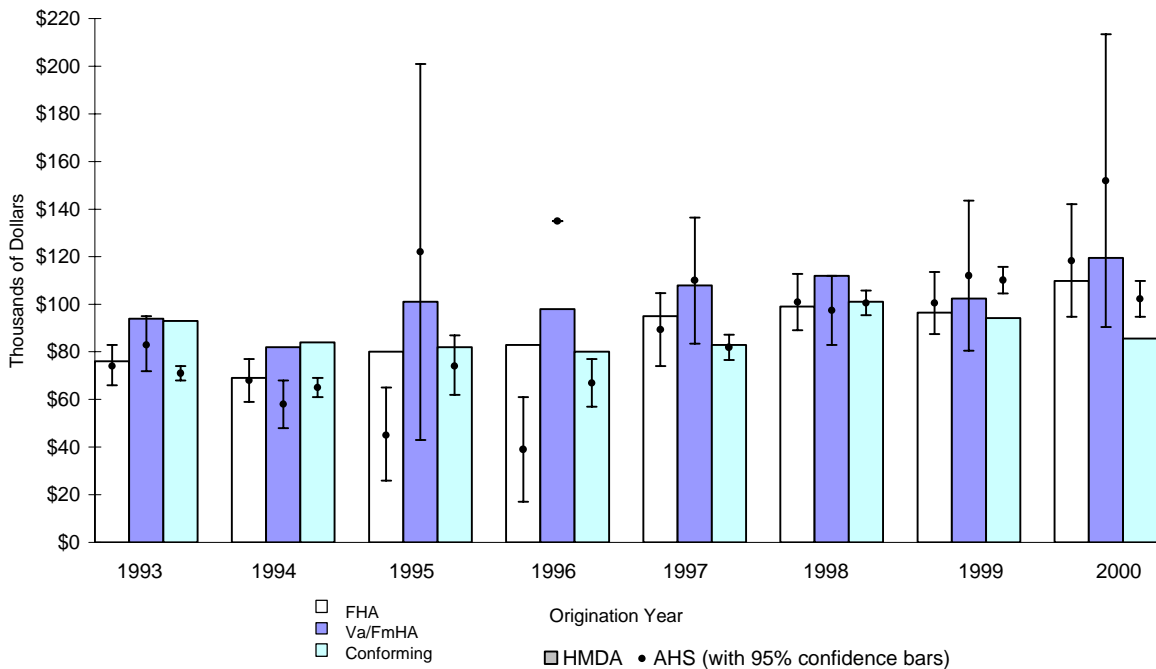
A similar pattern is observed for home-purchase loans in the VA/FmHA segment, as shown in Exhibit 4-15. In particular, the AHS appears to be more accurate in capturing loan amount information for loans in the 1996 to 2000 cohorts. For the early years, the average loan amount reported in the AHS is about 10 percent lower than the HMDA estimate.

Comparisons based on additional distribution statistics (25th, 50th and 75th percentiles of the loan amount) confirm our findings based on the average loan amounts. The AHS borrowers in the non-FHA segments of the market tend to report loan amounts that are lower than the HMDA benchmarks. (See Exhibit D-15-16 in Appendix D.) Reasons for this discrepancy are unknown. It is possible that the differences result from errors in homeowner's recall of the loan amount. There are a number of alternative explanations. One hypothesis is that the non-reporters in HMDA are mostly smaller lenders and they tend to originate relatively small size loans. This could lead to an inflated loan size distribution reported in HMDA. The higher loan amount observed in the AHS could also be an indication that a portion of the FHA loans, which in general have a smaller mortgage amount at origination, have been mistakenly identified by the AHS respondents as conventional loans. This is

less likely, however, given our finding in the previous section that the AHS has done a fairly accurate job in capturing the volumes and distribution of mortgages by market segment.

We now turn our discussion to the results for refinances, as shown in Exhibit 4-16. In general, the AHS does not appear to be very accurate in reporting average loan amounts for refinance loans. However, average loan amount estimates appear to be more accurate in 1997 and 1998 than in other years. For the earlier cohorts (1993-1996), it appears that the AHS has on average understated the loan amount of such mortgages, whereas in 1999 and 2000 the loan amount were overstated. Analysis for loan amounts of VA/FmHA refinances is problematic, given the small number of such loans captured in the AHS.⁴³

Exhibit 4-16:
Average Original Mortgage Amount by Mortgage Segment,
AHS vs. HMDA,
Loans Originated in Metro Areas, Refinance Originations Only



Mortgage Interest Rate

The AHS collects information on the current interest rate on mortgages. For many of our comparison databases, however, interest rates are captured at the time of loan origination. This presents a problem for obtaining benchmark estimates for the interest rate of adjustable-rate mortgages (ARMs). In addition, given the lack of information in the AHS about the timing of adjustments to ARMs, as

⁴³ Our tabulations using the AHS file indicate that there are just three sampled housing units with VA/FmHA refinances originated in 1995. Only one sampled housing unit is associated with such loan type in 1996, which explains why there is no confidence interval around the corresponding point estimate.

well as the index and margin used to make these adjustments, it is impossible for us to make these adjustments and check the reliability of the interest rate information of such loans reported in the AHS against any external data sources. Therefore, our replication analysis for this variable has focused exclusively on loans that were classified as fixed-rate mortgages in the survey. Another challenge for this evaluation is that any discrepancies we found in the interest rate distribution in the AHS relative to other mortgage market databases may either reflect errors in the owners' recall of the interest rate information or errors in the reported type of mortgage product (i.e., fixed rate vs. ARM). It is not possible to distinguish between these types of errors in our analyses.

The comparison databases selected for this evaluation are the FHA administrative data and the MIRS. Each covers a major mortgage market segment and provides a relatively long time-series for comparison. In addition, the MIRS micro data have previously been used for other interest rate studies.⁴⁴ Other available databases that provide information on interest rates include the RFS and the SCF. These data sets are not used as comparisons with the AHS, because the RFS data are dated and the SCF does not have an adequate sample size to support analysis by market segment and payment product type. HMDA, which has been used extensively for other comparisons, does not provide interest rate information.⁴⁵

Comparisons of average interest rate information between the AHS and FHA data are shown in Exhibits 4-17 and 4-18, separately for cohorts of home-purchase and refinance fixed-rate loans originated between 1989 and 2000. Both data sources confirm the interest rate dips in 1993 and 1998, which fueled the two refinancing booms of the last decade. Overall, average interest rates in the AHS are quite close to the averages reported in FHA's administrative data. Over the entire period from 1989 to 2000, the average interest rate in the AHS for purchase mortgages is only 0.09 percentage points below the FHA-reported average, while on refinance loans the AHS is only 0.15 percentage points below the FHA-reported average. Among purchase mortgages the discrepancy between the AHS and FHA data is never more than 0.3 percentage points, while among refinance loans it is also within this range except for 1991, 1997, and 2000. These findings suggest that the AHS variable is very accurate.

⁴⁴ David Torregrosa, "Interest Rate Differentials Between Jumbo and Conforming Mortgages, 1995-2000." Congressional Budget Office (CBO) paper, 2001.

⁴⁵ Recently, the Federal Reserve Board issued a proposal suggesting that interest rate information and a designation for a "high cost loan" (as defined in the Home Ownership and Equity Protection Act of 1994) be collected in the future HMDA data. The Federal Reserve Board has accepted the proposed changes.

Exhibit 4-17:
Average Interest Rate for
Fixed-Rate FHA Insured Mortgages, Home-Purchases,
AHS vs. FHA

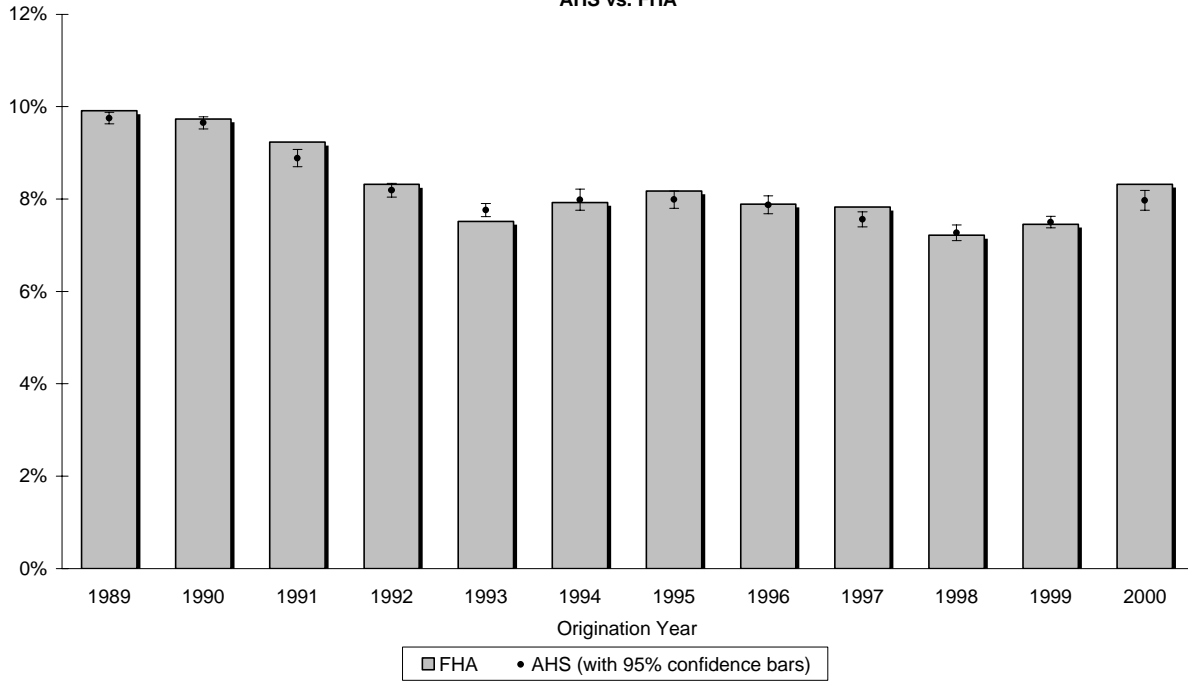
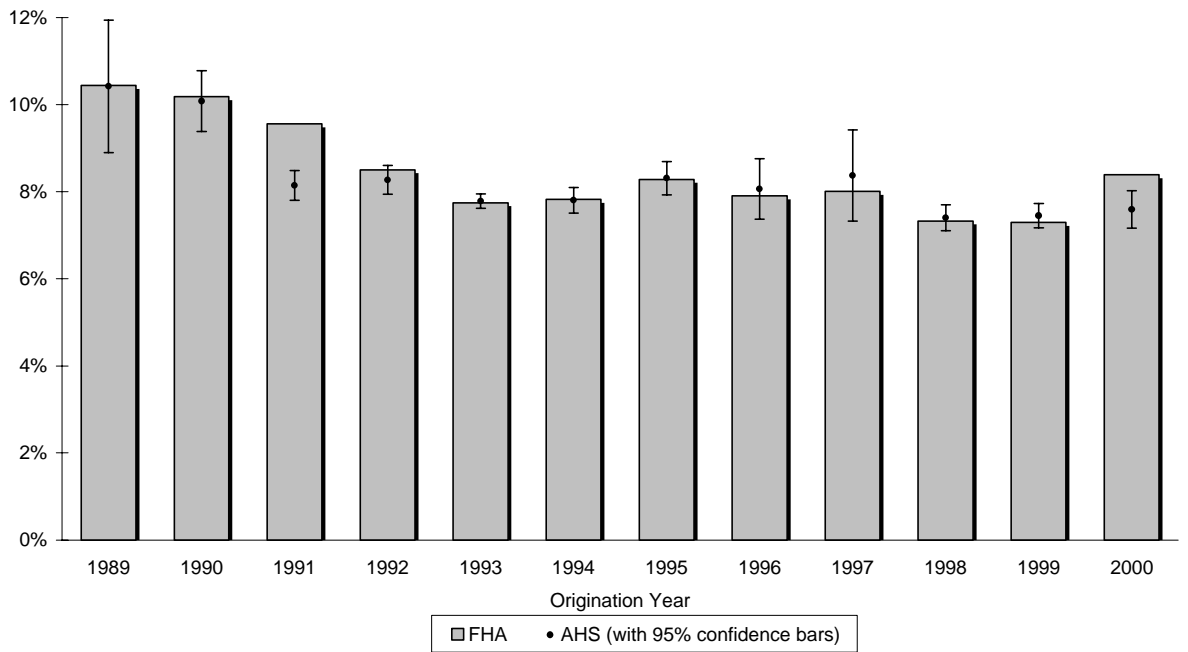


Exhibit 4-18:
Average Interest Rate for
Fixed-Rate FHA Insured Mortgages, Refinances
AHS vs. FHA



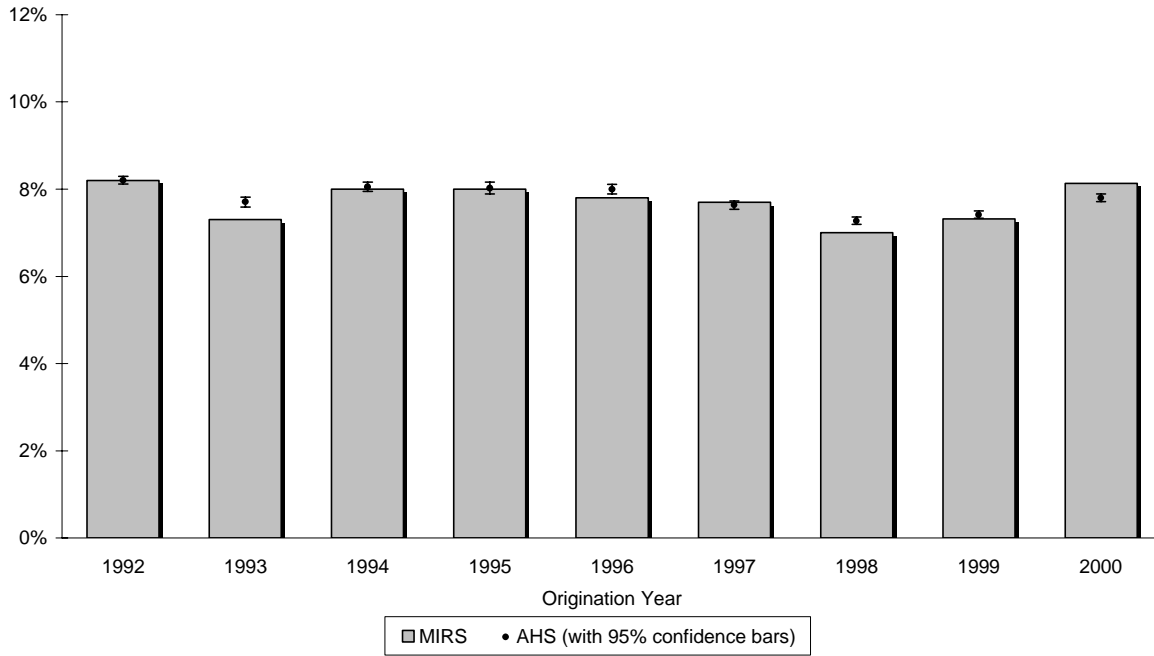
A few small discrepancies are worth mentioning nonetheless. Focusing on home purchases, the AHS appears to slightly overstate the interest rates for loans in the 1993 cohort, while the rates for the 1991, 1997 and 2000 cohorts seem to be understated. Exhibit D-17-18 in Appendix D provides additional statistics that characterize the distribution of the interest rate variable more comprehensively in the two data sets. They suggest that the interest rate distribution in the AHS is more dispersed than in the FHA benchmark data. Interest rate estimates for the 25th, 50th and 75th percentile, nevertheless, are very close between the two data sets, indicating that once again overall the AHS has done an excellent job in capturing the interest rate information for FHA loans. In particular, a surprising and very interesting finding is that while refinancings in this market segment are covered poorly in the AHS, the interest rate distribution of such mortgages when they are captured in the surveys appears to match reasonably well with the benchmark estimates. This suggests the under-reporting of refinance originations is randomly distributed with respect to interest rates of loans.

The MIRS data offer an opportunity to test the reliability of AHS interest rate information for loans originated in the conventional home-purchase segment of the market. Analyses are performed for conforming and jumbo loan types separately. We have excluded from the AHS data any mortgages for mobile homes and for housing units with more than one unit in structure because the sampling universe for the MIRS data does not include those loans. Exhibits 4-19 and 4-20 display the comparisons of mean interest rates computed from the two databases for cohorts of conforming and jumbo purchases made between 1992 and 2000. In general, the comparison suggests that interest rate reporting in AHS for these market segments has been accurate. As shown in the two Exhibits, the average interest rate estimates track very well between the two data sets for most of the cohorts examined. Among conforming purchases the discrepancy in average interest rates between the AHS and MIRS is never more than 0.4 percentage points, while among jumbo purchases it is also within this range except for the 2000 cohort.

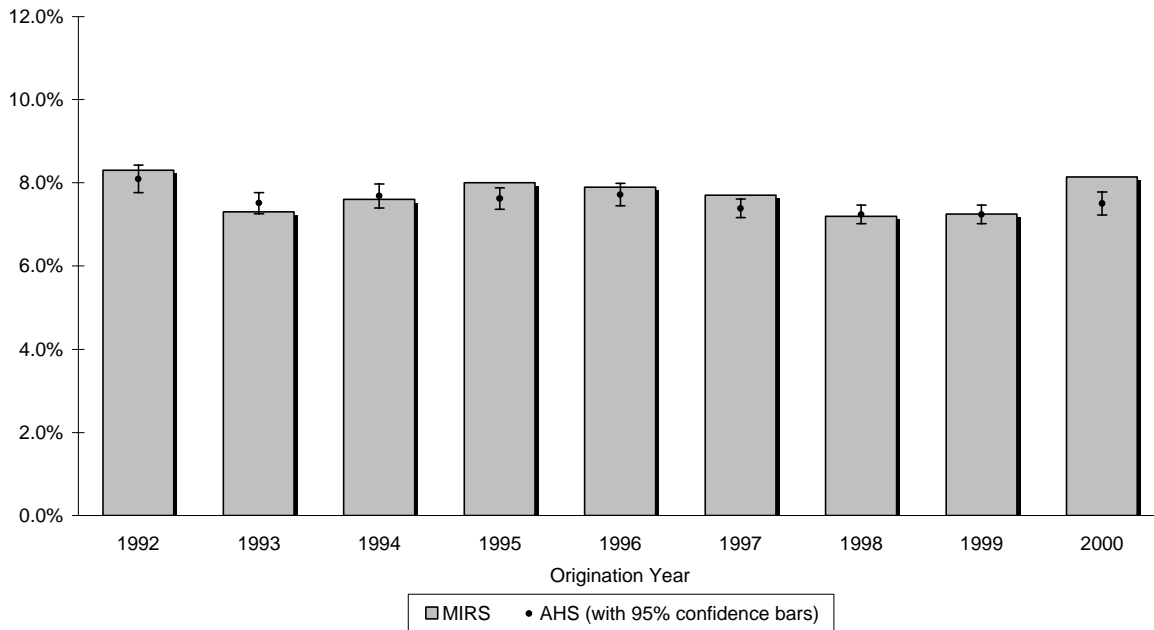
For some cohorts (1993, 1996, and 1998) of the loans in the conforming segment of the market, the AHS mean estimates are slightly higher than the benchmarks. Interest rates for loans originated in 2000 are understated in the AHS. These findings are further supported by an examination of the additional statistics (the 25th, 50th and 75th percentiles, standard deviation, and distribution of loans by interest rate categories), provided in Exhibit D-19-20 of Appendix D. Moreover, the distribution of the AHS variable for loans in this segment of the market is more dispersed, with a larger standard deviation. Relative to MIRS, the AHS data show a larger proportion of such loans in the high interest rate categories. This discrepancy, however, may not be an indication that the AHS interest rate information collected for this market segment is unreliable. Rather, this probably reflects the fact that the MIRS surveys did not include most of the financial institutions that specialize in serving the subprime mortgage market, which often is associated with loans with relatively high interest rates.

By contrast, the AHS interest rate distribution for loans in the jumbo segment of the market is closer to the MIRS benchmarks. Standard deviations derived from the two data sources are more similar and the AHS does not contain a larger share of such originations in the high interest rate categories relative to the MIRS. This is probably because there are very few subprime originations in this segment of the mortgage market.

**Exhibit 4-19:
Average Interest Rate for
Fixed-Rate Conventional Purchase Mortgages,
Conforming
AHS vs. MIRS**



**Exhibit 4-20:
Average Interest Rate for
Fixed-Rate Conventional Purchase Mortgages,
Jumbo
AHS vs. MIRS**



Loan-to-value (LTV) Ratio at Origination

While the AHS does not include a variable for the LTV, it does report the loan amount and both the purchase price and the owner's estimate of the current value of the property that can be used to estimate the LTV. For home-purchase loans, the purchase price information reported by the borrowers is the most appropriate measure of the house value. For refinance mortgages, the owner's estimate of the house value should be used.

The principal data sets that contain estimates of LTV that can be used to evaluate the AHS estimates are the FHA administrative data and MIRS.⁴⁶ HMDA does not report such a variable. While loan amounts are reported, the data do not include information on sales price or appraised house value. Thus, it is impossible to estimate LTVs from HMDA records.

Exhibit 4-21 compares the average LTV for the cohorts of home-purchase FHA loans originated between 1989 and 2000, using data from the AHS and FHA administrative records. The estimates track reasonably well between the two data sets. According to the FHA data, the average LTV has been rising steadily over the last decade, from 91 percent in 1989 to 96 percent in 2000. The average is 94 percent for the period examined. Estimates from the AHS reflect this upward trend. However, there is a tendency for the AHS to understate the average LTV value in this segment of the market. The average LTV across the twelve cohorts is 91 percent in the AHS.

Further analysis is displayed in Exhibit 4-22, which compares the LTV distributions between the two data sources using four LTV categories (less than 80 percent, 80 to 90 percent, 90.1 to 95 percent, and greater than 95 percent). In general, there is not very good correspondence between the distribution of loans by LTV category in the AHS and that found in FHA's data. Relative to the FHA benchmarks, the AHS has consistently identified a higher proportion of home-purchase loans in the lowest LTV category, as well as in the lower two LTV categories combined (loans with a LTV less than 90 percent) for all origination cohorts examined. The category of loans with LTV above 95 percent may be of particular importance for policy makers, as this category corresponds to borrowers who manage to purchase a home with relatively little equity. Prior to 1996, the AHS estimates appeared to be unreliable, with share estimates that were much too high from 1989 to 1991 and much too low for this category from 1992 to 1995. From 1996 to 1998 the AHS estimates were quite accurate for this category. Large discrepancies appeared again in the 1999 and 2000 cohorts of loans. Given the general lack of accuracy in the LTV distribution overall, these results may simply be the result of sampling variability. Sampling errors from the loan amount and house value variables in the AHS may have exaggerated the differences in share estimates between the two data sets. For this analysis, we have focused on point estimates for the average rather than confidence intervals. Since the LTV value is computed from two variables, its confidence intervals are very difficult to derive. We also compare the share of loans by LTV category, but since sample sizes are fairly small, a small change in the sample data can have a large effect on the share estimates for the LTV categories.

⁴⁶ The FHA administrative database contains several variables pertaining to LTV. This study makes use of the LTV estimates that exclude any financed upfront MIP. LTV information collected in the MIRS data are rarely used by analysts and policy makers. Therefore, the quality of this data element is unknown.

Exhibit 4-21:
Average Loan-to-Value Ratio of FHA-Insured Mortgages at
Origination, AHS vs. FHA,
Home-Purchase Loans

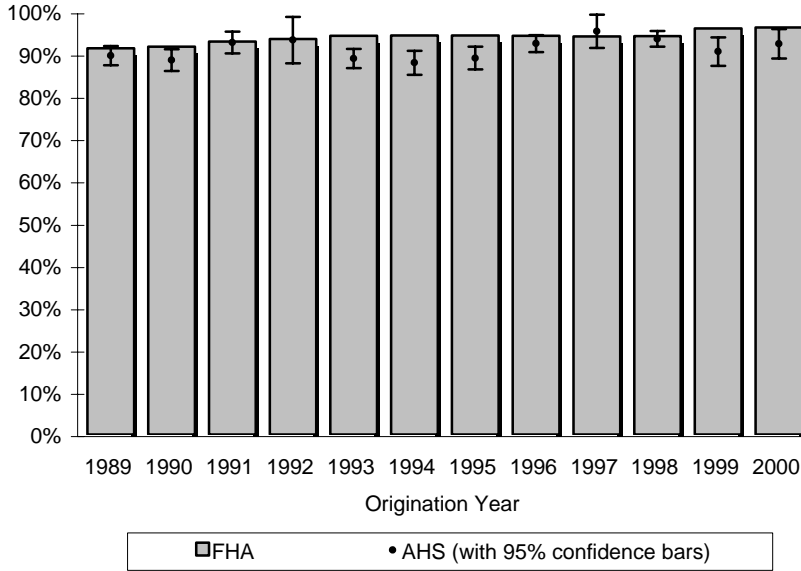
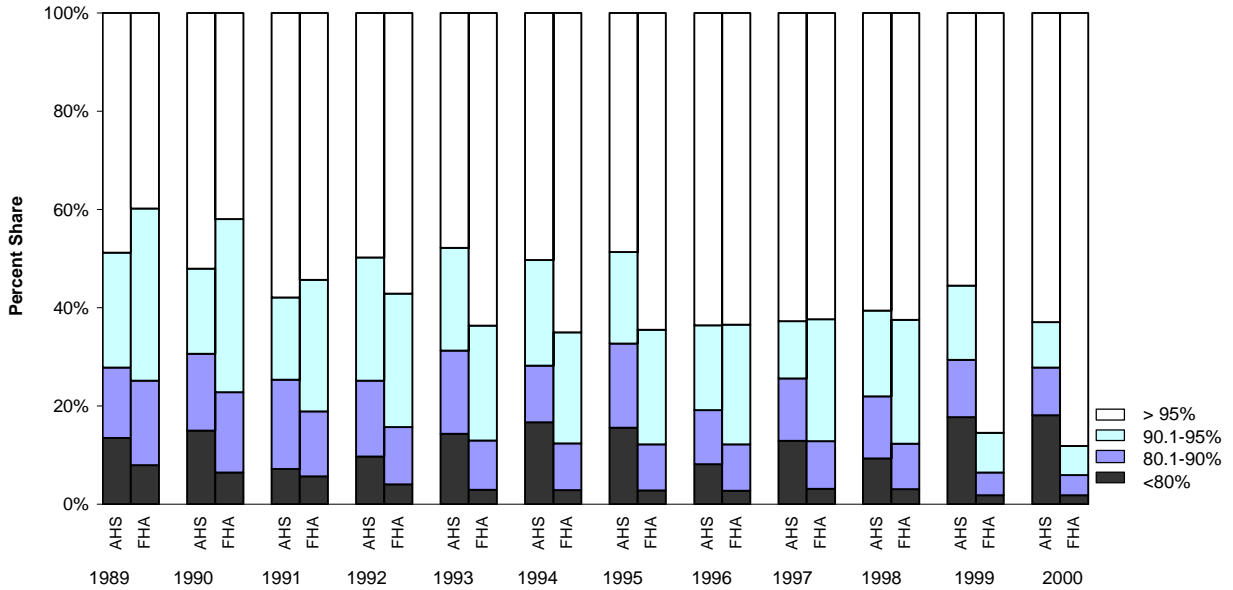


Exhibit 4-22:
Distribution of Loan-to-Value Ratio of
FHA-Insured Mortgages at Origination, AHS vs. FHA,
Home Purchase Loans



The discrepancies found in the LTV distribution are somewhat surprising given our finding (described above) that the AHS has captured loan amount information quite accurately for FHA mortgages. This leads us to the conclusion that these discrepancies must result mainly from the fact that home values are measured differently in the two data sets. While the purchase price (reported by the homeowner) is used in the AHS LTV calculation, FHA administrative data depend on the appraised house value (determined by FHA-approved appraisers in the loan application process). Purchase price information is not available from the FHA data. Comparisons of the distribution (the mean, 25th, 50th, and 75th percentiles) of home values for this segment of loans between the two data sets confirm that the AHS house values are consistently higher than FHA counterparts. (See Exhibit D-21-22a of Appendix D.)

Comparison of LTV distributions for refinance FHA loans between the AHS and FHA data is highly problematic. The house value information collected in the FHA administrative records for such loans only reflects the appraised house value at purchase, rather than the updated market value of the house when the refinancing happened. For most of the country, house values have appreciated over the time period examined in this study. Therefore, this comparison is not shown.

We now turn to the replication analysis of LTVs for home-purchase loans in the conventional market segment using information provided in the MIRS data. Since both the loan amount and purchase price variables are top-coded in the public-use version of AHS files, our investigation focuses exclusively on conforming loans. One again, for this comparison, we have excluded from the AHS data any mortgages for mobile homes and for housing units with more than one unit in structure because the sampling universe for the MIRS data does not include those loans. Exhibit 4-23 presents a comparison of the average LTV between the two data sets. According to MIRS, the average LTV for originations in this segment of the market has risen gradually, from about 74 percent to 80 percent, over the 1989-2000 period. The AHS estimates for 1999 and 2000 stand out as having a much greater magnitude of error than other years. Excluding these two years, the AHS estimates are remarkably close to the MIRS benchmarks. Over the 1989-1998 period, the average LTV for conforming purchases is 79 percent in the AHS, as compared to 77 percent in MIRS. The discrepancy for individual year is never greater than 5 percentage points.

Further comparisons of the distribution of loans by LTV category between the two data sets, as shown in Exhibit 4-24, reveal that there are significant differences in the distributions of loans by LTV category between the AHS and MIRS. In particular, the AHS has a smaller proportion of loans in the lowest LTV category (LTV less than 80 percent) and a higher proportion of loans in the highest LTV category (LTV larger than 95 percent) for all the origination cohorts examined. The difference in the share estimates between the two data sets for these LTV categories ranges from 15 to 20 percent for most of the cohorts. The shares of loans in the middle two LTV categories are, in general, fairly close between the two data sets. Once again, sampling variability may have exaggerated the differences in share estimates between the two databases in this comparison. The distortions occur at thin points in the distribution, where a small change in the sample data can have a large effect on the share estimates. Interestingly, given that the average LTVs between the two data sets are not that different, the larger share of high LTV loans in the AHS must be essentially offset by the smaller share of low LTV loans.

Exhibit 4-23:
Average Loan-to-Value Ratio of Conventional Home-Purchase
Mortgages at Origination, AHS vs. MIRS, Conforming Loans

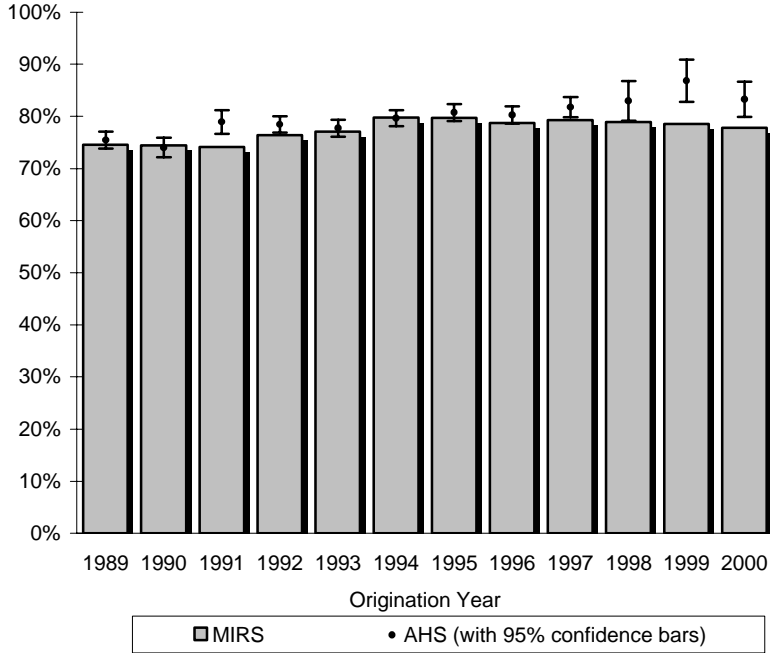
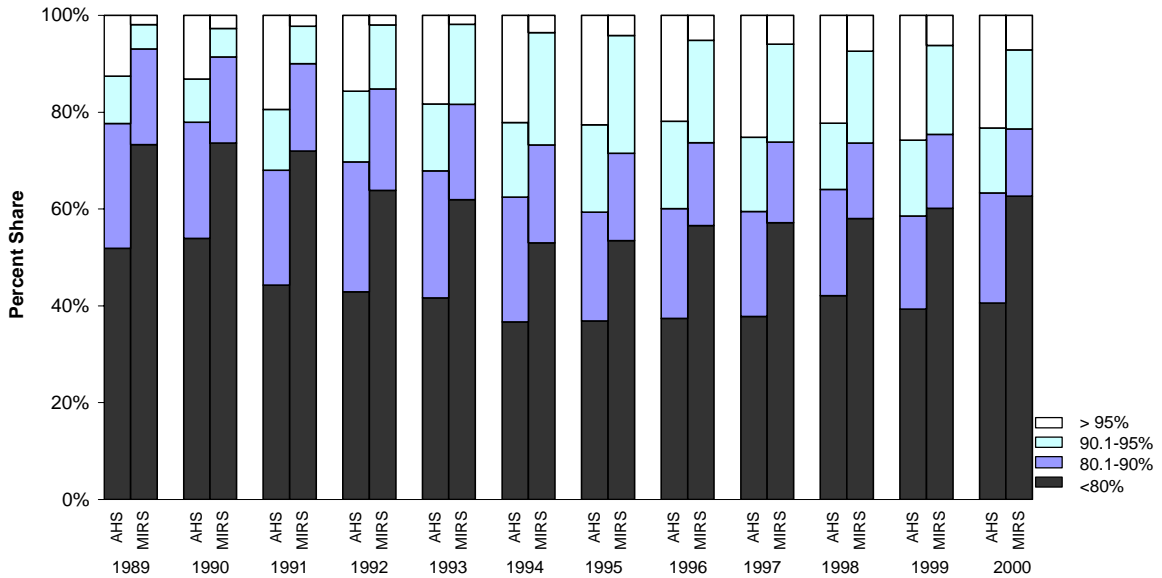


Exhibit 4-24:
Distribution of Loan-to-Value Ratio of
Conventional Home-Purchase Mortgages at Origination, AHS vs. MIRS,
Conforming Loans



To shed additional light on these discrepancies, we examine the home values and loan amounts between the two data sets. Home values are based on purchase house prices provided in both data sets. A comparison of this variable between the two databases suggests that the distribution (mean, 25th, 50th, and 75th percentiles) of AHS house price is consistently lower relative to the MIRS values. (See Exhibit D-23-24a in the Appendix) On the other hand, as discussed in a previous section, we found that the reported mortgage amounts for this segment of loans in AHS tend to be smaller than the benchmark estimates. The discrepancy of the LTV estimate in the AHS would depend on the relative size of these two differences, since both are in the same direction relative to the MIRS benchmarks.

One explanation for the lower value estimates in the AHS is the misreporting of mortgage insurance status for this segment of loans identified in AHS. A portion of borrowers might have incorrectly classified their FHA-insured loans, which tend to have lower loan amounts and higher LTVs, as conventional loans. However, this explanation is in conflict with our earlier finding that the AHS reports the origination volumes and market shares of FHA and conventional loans fairly accurately.

Mortgage Payment Plan

The AHS does not ask homeowners explicitly about the mortgage payment plan or product type – that is, whether the loan is fixed rate, adjustable rate, balloon payment, etc. Rather, adjustable rates and other payment plan features are identified in the survey by asking the respondents a series of questions about whether the monthly mortgage payments are fixed over time or whether they vary with interest rates and other factors. However, data elements related to these payment plan features in the public-use file do not allow the AHS users to easily and unambiguously classify each mortgage into a mutually exclusive product type. To do so, users can run a SAS computer code written by the Census Bureau staff, which reconciles all the relevant information collected. The computer program classifies mortgages into the following seven mutually exclusive types:

- Fixed payment, self-amortizing mortgage;
- Adjustable rate mortgage (ARM);
- Adjustable term mortgage;
- Graduated payment mortgage;
- Balloon mortgage;
- Combination of the above; and
- Other.

The replication analysis was used to test whether this indirect method of identifying payment plan features has reliably identified mortgage product types in the AHS. This analysis focuses primarily on comparing the AHS estimates of the fixed-rate and ARM market shares to external benchmark estimates derived from other mortgage market data sources. Given the relatively small numbers of loans in product types besides fixed-rate and ARMs in the market, such loans are grouped together into a single category called “other” in this analysis.

The current AHS survey instruments are incapable of identifying borrowers with more sophisticated payment product types such as hybrid ARMs and two-step mortgages.⁴⁷ These mortgages, among other forms of hybrid product types, would generally fall under the “combination of the above” or “other” categories, if the respondents have indeed indicated that both the fixed payment and adjustable rate features apply to their mortgages. Otherwise, they are mistakenly grouped under the ARM or fixed-rate categories in the AHS data. As a result, it is not possible to calculate the share of borrowers who took out hybrid ARMs in AHS. While FHA did not offer mortgage insurance for such loans over the study period, the share of hybrid ARMs and two-step mortgages in the conventional market in recent years is not trivial. This is a clear limitation of the AHS.

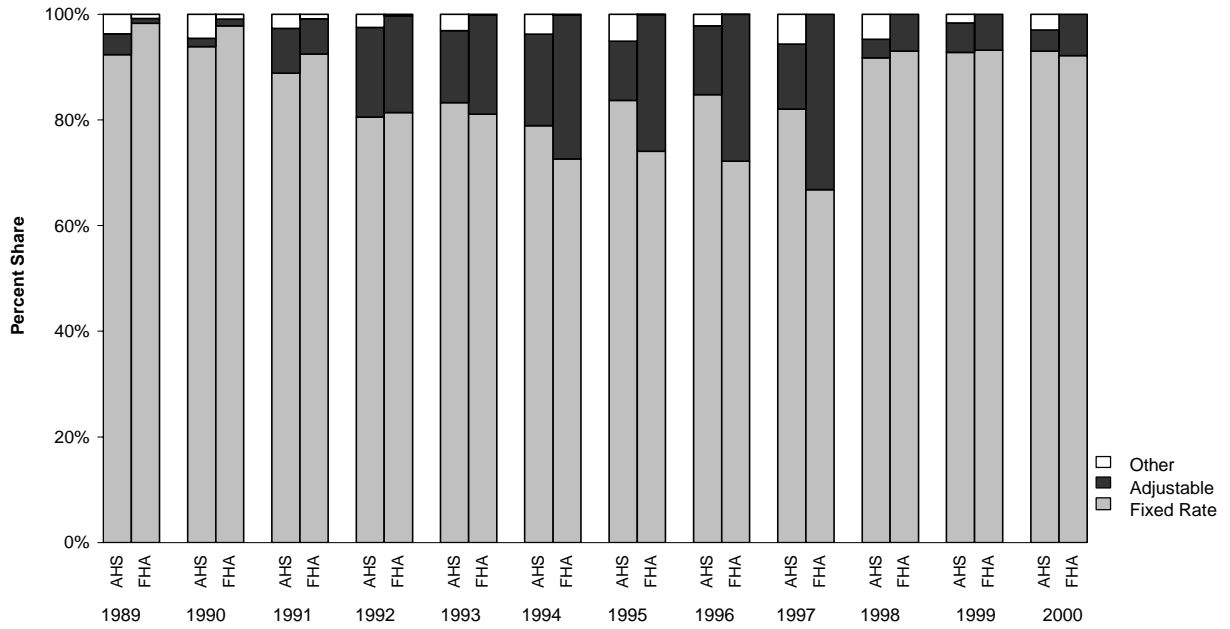
There are two principal sources of data for testing mortgage payment plan information from the AHS. For loans with FHA insurance, the natural choice is the FHA administrative database. MIRS data are used to access the accuracy of fixed-rate and ARM share information in AHS for loans in the conventional market segment. Payment plan information is not available from the HMDA records.

The analyses for home-purchase and refinance FHA loans are displayed in Exhibits 4-25 and 4-26. Stratification by loan purpose is crucial here because of the high correlation between product type and loan purpose (refinance loans tend to be fixed rate products), in addition to the problem of low coverage of refinances in the AHS for most of the years examined. An important observation about the type of loans insured by FHA is that there are very few FHA loans other than fixed-rate or adjustable rate loans. FHA does insure graduated payment loans, but these are always less than one percent of purchase mortgage originations, and have become exceedingly rare since the early 1990s. Despite the rarity of this class of loans, the AHS reports a share of “other” loans that ranges from 2 to 6 percent over the period. This suggests that the AHS has consistently over reported such product types.⁴⁸

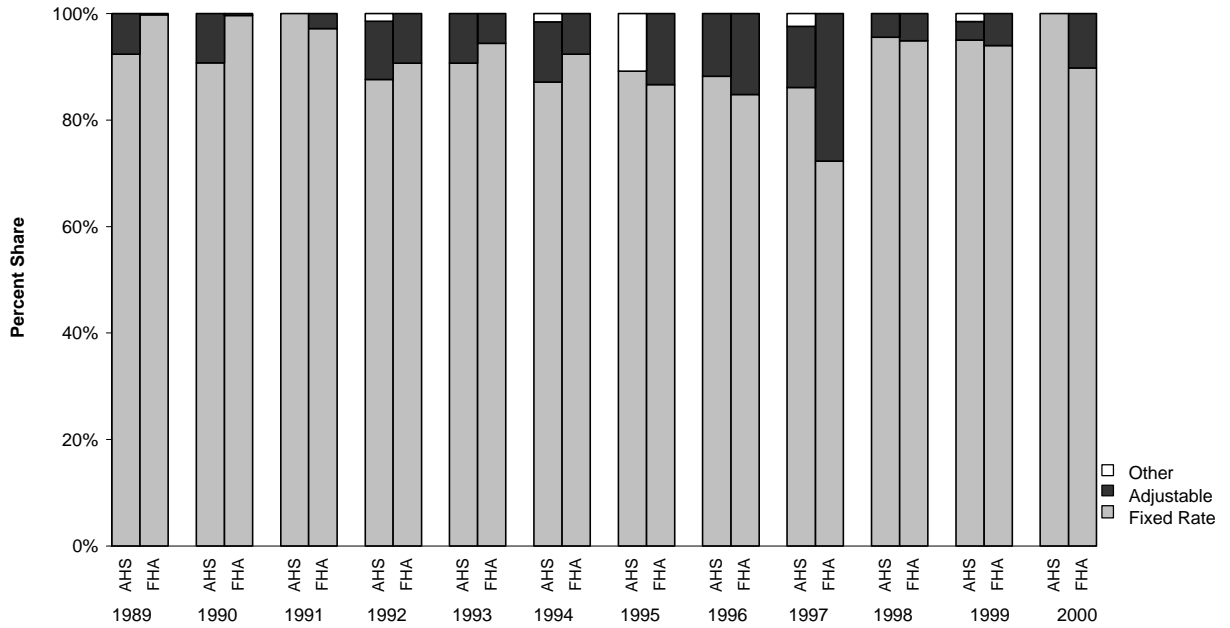
⁴⁷ Hybrid ARMs are mortgages that have a fixed interest rate for the first three, five, seven or ten years and then the interest rate adjusts annually thereafter. These are commonly referred to as 3/1, 5/1, 7/1 and 10/1 Hybrid ARMs. Because a hybrid ARM converts to an adjustable rate mortgage after the initial fixed interest rate period, lenders can offer these loans with an initial interest rate that is lower than the interest rate for a 30-year fixed rate mortgage. Two-step mortgages, such as 5/25 and 7/23, offers an initial five (or seven) years fixed rate; after that, the rate is adjusted once for the remaining twenty-five (or twenty-three) years of the loan.

⁴⁸ It is also possible that the respondents have misidentified these loans as FHA-insured.

**Exhibit 4-25:
Distribution of Home-Purchase FHA Mortgages
by Payment Plan Type
AHS vs. FHA**



**Exhibit 4-26:
Distribution of Refinance FHA Mortgages
by Payment Plan Type
AHS vs. FHA**

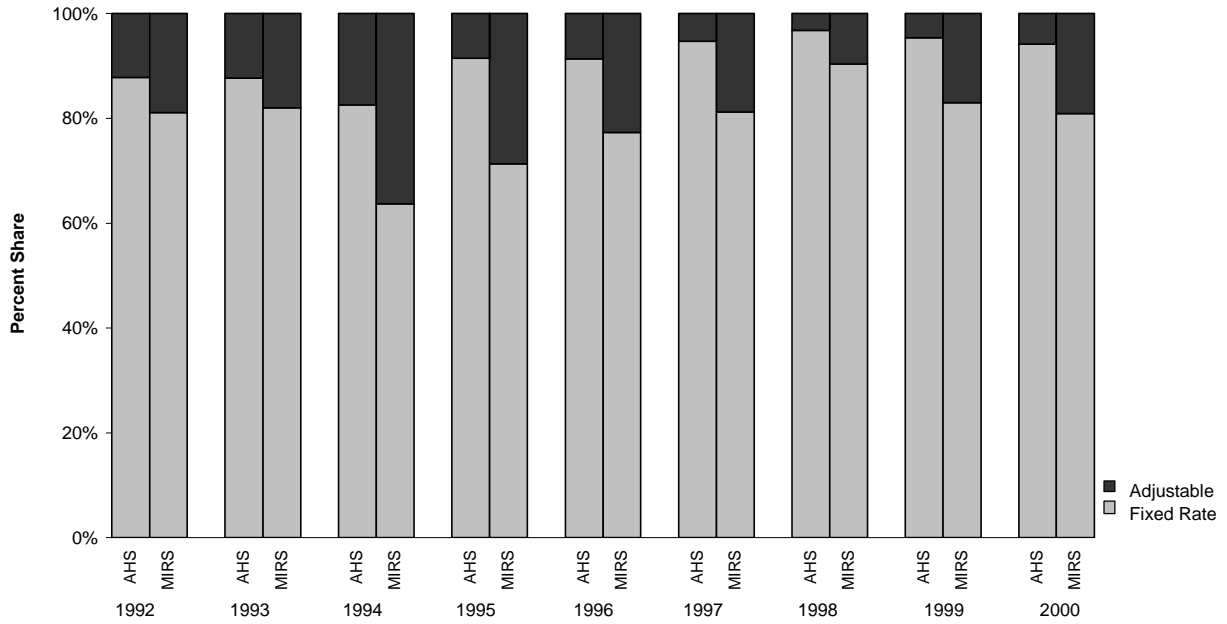


With regard to ARMs, the AHS reports shares that are very close to the FHA benchmarks for the four early cohorts (1989-1992). The differences are only one percentage point in three of these years, and none of the differences is statistically significant. From 1989 to 1991, the AHS proportion of fixed-rate loans is lower than the one in the FHA data by an amount that is nearly equal to the “other” loan type share. During this period, the loans reported as “other” in the AHS would appear to be most likely fixed-rate loans. From 1993 to 1997, ARMs became a more common type of FHA loan. However, the AHS does not accurately reflect the level of increase in ARM shares. The AHS has consistently understated the ARM share of home-purchase mortgages and generally overstated the fixed-rate share during this period. Both data sets show that ARM shares have declined substantially since 1998. The AHS share estimates for fixed-rate loans and ARMs identified in the 1999 and 2000 cohorts track reasonably well with the FHA data. In any case, it seems likely that most of the loans reported as “other” in the AHS are ARMs.

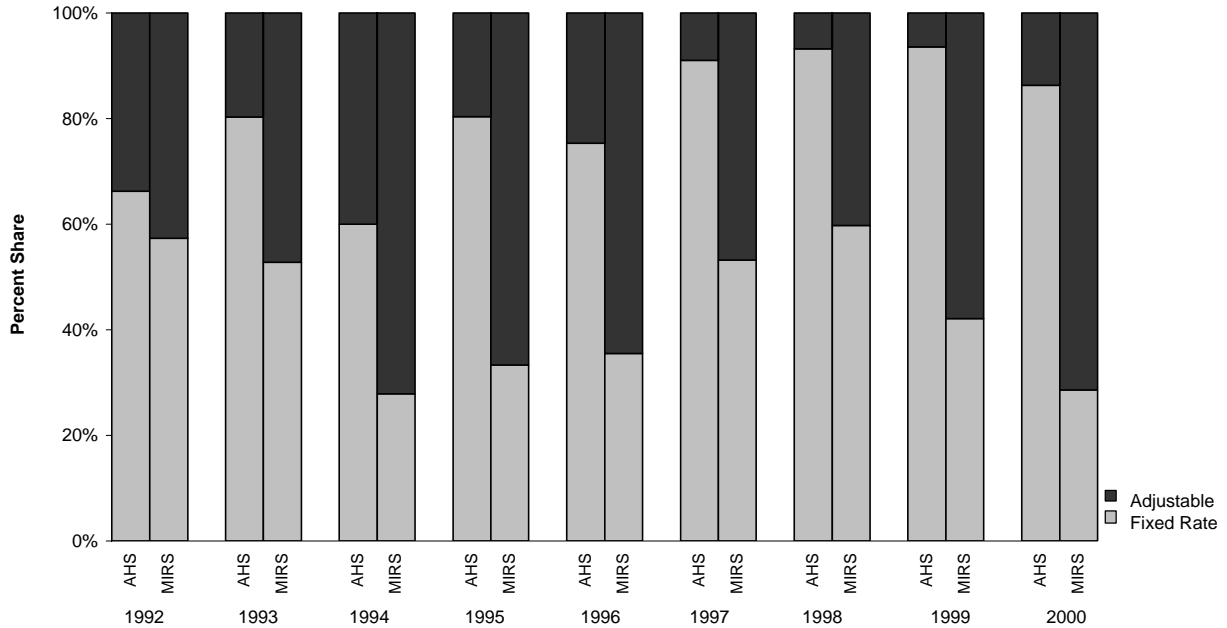
It is challenging to conduct similar comparisons for refinance FHA loans because of the small number of such mortgages captured in the AHS. For most of the cohorts, the AHS share estimates are associated with very wide confidence intervals that make it difficult to draw comparisons with the benchmarks that are statistically reliable. Focusing on the point estimates alone, it appears that the AHS has overestimated the ARM shares for the earlier cohorts, while the proportion of ARMs is understated for originations in the late 1990s. For the 1995 cohort, the share of loans in the “other” product type category has also been seriously overstated relative to the FHA benchmark.

Analyses for loans in the conventional market segment are shown in Exhibit 4-27 (conforming) and Exhibit 4-28 (jumbo), using data from MIRS as the comparison. Comparisons are restricted to home-purchase mortgages for non-mobile home one-unit structures. In addition, for this analysis, we have excluded loans identified in the “other” category in the AHS, since the MIRS survey does not include loans other than ARMs and fixed-rate level-payment products. Overall, we found that the AHS has consistently understated the share of ARM products in this segment of the mortgage market, an indication that the indirect method employed by the survey has not been reliable in identifying such mortgages. ARM shares derived from the AHS for the nine cohorts (1992-2000) are all smaller than the MIRS benchmarks, with differences significant at the 95-percent level. The undercounting problem appears to be most pronounced for the cohorts of loans originated since 1994.

**Exhibit 4-27:
Distribution of Conforming Conventional Home-Purchase Mortgages
by Payment Plan Type
AHS vs. MIRS**



**Exhibit 4-28:
Distribution of Jumbo Conventional Home-Purchase Mortgages
by Payment Plan Type
AHS vs. MIRS**



The use of ARMs is more prevalent among borrowers in the jumbo segment of the market, as Exhibit 4-28 indicates. While the AHS data have been able to reflect this fact, the AHS-based ARM share estimate for each cohort of jumbo origination is much smaller than the MIRS benchmark. Except for the 1992 cohort, estimates from the two data sets are statistically different at the 95-percent level. Last but not least, the AHS share estimates are associated with very wide confidence intervals because of the small sample size of such mortgages reported.

In interpreting the findings regarding the product type shares for conventional conforming and jumbo loans, it is worth noting that there have long been data quality concerns regarding the ability of MIRS to distinguish ARMs from fixed-rate level payment products. A study by Cotterman and Pearce (1996) evaluating MIRS data from 1989 to 1993 suggested that lenders frequently misreported ARMs as fixed-rate loans for that period.⁴⁹ Because the Federal Housing Finance Board now screens the data more effectively for errors in reporting, the data quality of MIRS is said to have improved.⁵⁰ However, to the extent that the bias in MIRS was to over report fixed-rate loans, this problem would not explain why the AHS consistently shows much higher levels of fixed-rate loans.

Borrower Characteristics

This section examines the reliability of borrower characteristics reported in the AHS. The approach is similar to the method we used for testing loan characteristic variables: we compare the distribution of the numeric values and the shares of borrowers associated with different characteristics with those based on other mortgage market data sources. Since the focus is on distribution (rather than on the number of loans originated), this approach allows us to evaluate whether the AHS accurately reports borrower characteristics even if it may not accurately capture volumes for certain type of loans.

The variables we have examined are:

- First-time homebuyer status;
- Race/ethnicity; and
- Income.

The comparison databases are Chicago Title, HMDA, FHA and RFS. Results of this analysis are mixed. We found that the shares of first-time homebuyers among purchases reported in AHS track very well with the Chicago Title data. The AHS first-time buyer shares also appear to match well those from RFS in the conventional and VA/FmHA/RHS market segments, but not for FHA. Further

⁴⁹ Robert F. Cotterman and James E. Pearce, “The Effect of the Federal National Mortgage Association and the Federal Home Loan Mortgage Corporation on Conventional Fixed-Rate Mortgage Yields” in *Studies On Privatizing Fannie Mae and Freddie Mac*. U.S. Department of Housing and Urban Development, May 1996.

⁵⁰ David Torregrosa, “Interest Rate Differentials Between Jumbo and Conforming Mortgages, 1995-2000.” Congressional Budget Office (CBO) paper, 2001.

evidence is found in the comparison with the FHA data that the AHS has reported a smaller share of first-time buyers among FHA home-purchase mortgages. This difference is most likely due to how first-time buyer status is defined in the two data – the AHS estimates are more consistent with a strict notion of first-time buyers. Regarding race/ethnicity, AHS tracks reasonably well with both the FHA data and HMDA in describing the racial distribution of homebuyers overall and by market segment. Nonetheless, there are some discrepancies. They most likely result from the fact that race/ethnic categories are defined and reported differently across the data sets. Benchmarking the borrower’s income variable is difficult. FHA and HMDA data define income quite differently from the AHS. Perhaps because of this limitation, we found that borrower incomes in the AHS are consistently higher than those in HMDA for the vast majority of the origination cohorts examined.

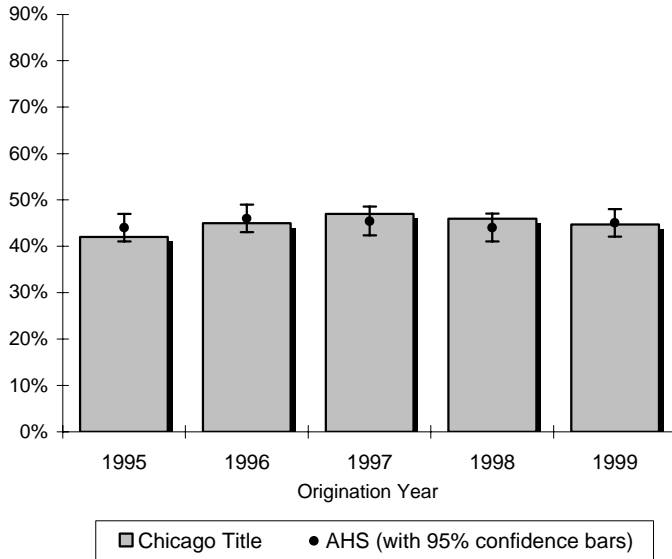
First-time Homebuyer Status

The principal focus of the replication analysis for the first-time buyer variable is to evaluate the share of home purchase mortgages in different segments of the market accounted for by borrowers who identified themselves as first-time buyers. First-time buyers are of interest to policy makers who seek to increase homeownership rates. Among the comparison databases, only three have information on first-time buyer status: Chicago Title Company Annual Survey of Recent Homebuyers, FHA administrative data, and the RFS.

A comparison of the 1995-1999 first-time buyer share estimates between the AHS and Chicago Title data for all segments of the market is presented in Exhibit 4-29. Chicago Title data for other years are not available. The Chicago Title estimates are based on an annual survey of homebuyers in twenty major metropolitan areas. Therefore, to make the estimates comparable, we have limited the AHS tabulations to home-purchase loans made in those metropolitan areas. As the Exhibit shows, share estimates from the two data sources are very close for each origination cohort. For all five years, the Chicago Title estimates are located within the 95-percent confidence interval of the AHS estimates, indicating that any differences are not statistically significant. There is no consistent pattern to the differences in the estimates, as the estimate from the 1997 AHS was slightly higher while the estimate from the 1999 AHS was slightly lower. Overall, the AHS reporting is consistent with the Chicago Title estimates.⁵¹

⁵¹ The Chicago Title data are based on a survey of homebuyers. Its estimates are therefore subject to sampling errors. However, no standard errors or confidence interval estimates are provided in the *Who’s Buying Homes in America* report, so we were not able to compare AHS confidence intervals with Chicago Title confidence intervals.

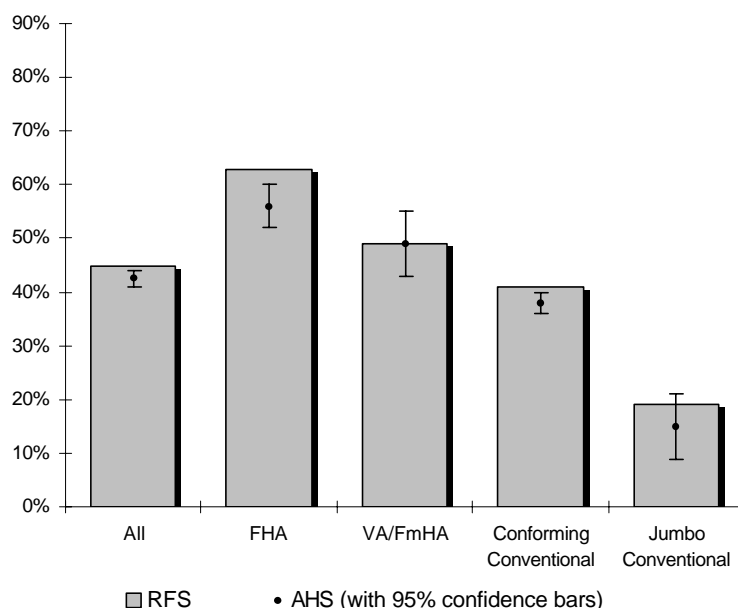
**Exhibit 4-29:
Share of First-Time Buyers
Among Home-Purchase Loans
AHS vs. Chicago Title Data, Metropolitan Areas Only**



To evaluate the accuracy of AHS first-time buyer shares of earlier origination cohorts, we use the RFS data. The RFS permits analysis by mortgage market segment. Exhibit 4-30 depicts the comparisons of first-time buyer shares among all home-purchases and by market segments for loans originated in the 1989-1991 period. The RFS estimates are subject to sampling errors, and the associated confidence intervals of the point estimates are reported in Exhibit D-30 of Appendix D.

Focusing on the point estimates, the AHS first-time buyer share in aggregate matches up fairly well with the RFS, although it is slightly lower (43 percent compared to 45 percent in the RFS). Since first-time buyer status is self-reported in both surveys, we have no reason to conclude that one estimate is more reliable than the other. Findings by market segment are consistent with this overall finding. The AHS first-time buyer share is lower than the RFS share in each market segment, with the exception of the VA/FmHA segment, where the estimates are the same. Of note, however, the AHS and RFS do show very consistent differences in first-time buyer shares across market segments, in keeping with expectations. *A priori*, we know that the FHA segment should have the largest share of first-time buyers, while it is expected that the share of such borrowers is smallest among those who take out jumbo loans; first-time buyer shares for VA/RHS and conventional conforming purchases should be between these two extremes. Both the RFS and AHS data correctly reflect this pattern, as shown in Exhibit 4-30.

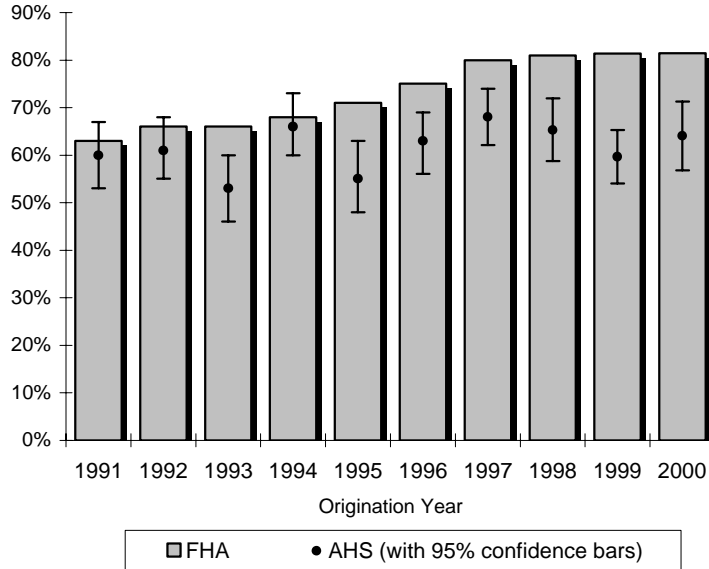
**Exhibit 4-30:
Share of First-Time Buyers
Among Home-Purchase Loans, 1989-1991
by Mortgage Market Segment
AHS vs. RFS Data**



Comparisons for the 1991-2000 origination cohorts between FHA loans estimated from the AHS and reported in the FHA administrative data are presented in Exhibit 4-31.⁵² They confirm a pattern that the AHS consistently reports a lower share of first-time buyers for all ten cohorts examined, although in 1991, 1992 and 1994 the actual share falls within the 95-percent confidence interval. While it is possible that these lower shares relative to the FHA data are indicative of underreporting in the AHS, this discrepancy is most likely the result of a difference in definition – that is, how this information is defined and solicited in the two data sets. The AHS survey asks whether any of the owners in a household ever owned a home before, while the FHA data define a first-time buyer as a household in which any of the owners has not owned home in the last three years. Therefore, in the FHA database, a borrower who had previously owned a house but had rented for the past three years would be identified as a first-time buyer. Moreover, it is not known how careful FHA lenders are in inquiring about first-time buyer status in the loan application process. All in all, this discrepancy in estimates between the AHS and FHA data is not an indication that the AHS reporting is inaccurate. The AHS definition and estimates are more consistent with a strict notion of first-time homebuyers.

⁵² Loans in the 1989 and 1990 cohorts are excluded from this analysis because the first-time buyer status variable is not sufficiently populated in the FHA administrative data.

**Exhibit 4-31:
Share of First-Time Buyers Among
FHA-Insured Home-Purchase Loans
AHS vs. FHA Data**



Race/Ethnicity

In evaluating the reliability of borrower race/ethnicity information in the AHS, FHA administrative records are used as the principal comparison data source for FHA-insured originations. For loans in the other segments of the market, HMDA data provide an appropriate comparison, given HMDA’s broad coverage of market segments and its widespread use by analysts and regulators for examining the lending patterns of minority borrowers and neighborhoods. The borrower information in the FHA administrative files also allows us to cross check the accuracy of the racial/ethnic composition of the FHA market segment in HMDA.

A major challenge in this investigation is to define racial/ethnic groupings of borrowers that are consistent and comparable across the AHS and the two benchmark data sets. Race and Hispanic origin information is collected separately from two questions in the AHS. A respondent can identify himself/herself as both white and Hispanic. This is handled differently in HMDA and FHA data, in which Hispanic origin and other racial identities are reported as mutually exclusive categories.⁵³ We are also concerned about sample sizes, given the small number of records identified in the AHS for minority borrowers in some of the market segments.

⁵³ In AHS, the race categories are white, black, American Indian, Aleut, or Eskimo, Asian or Pacific Islander, or other. Hispanic origin of the householder is identified in a separate question. The categories reported in FHA data are non-Hispanic white, non-Hispanic black, American Indian or Alaskan, Asian or Pacific Islander, Hispanic, and other. HMDA records classify borrowers into the following eight categories: American Indian or Alaska native, Asian or Pacific Islander, black, Hispanic, white, other, information not provided, or not applicable.

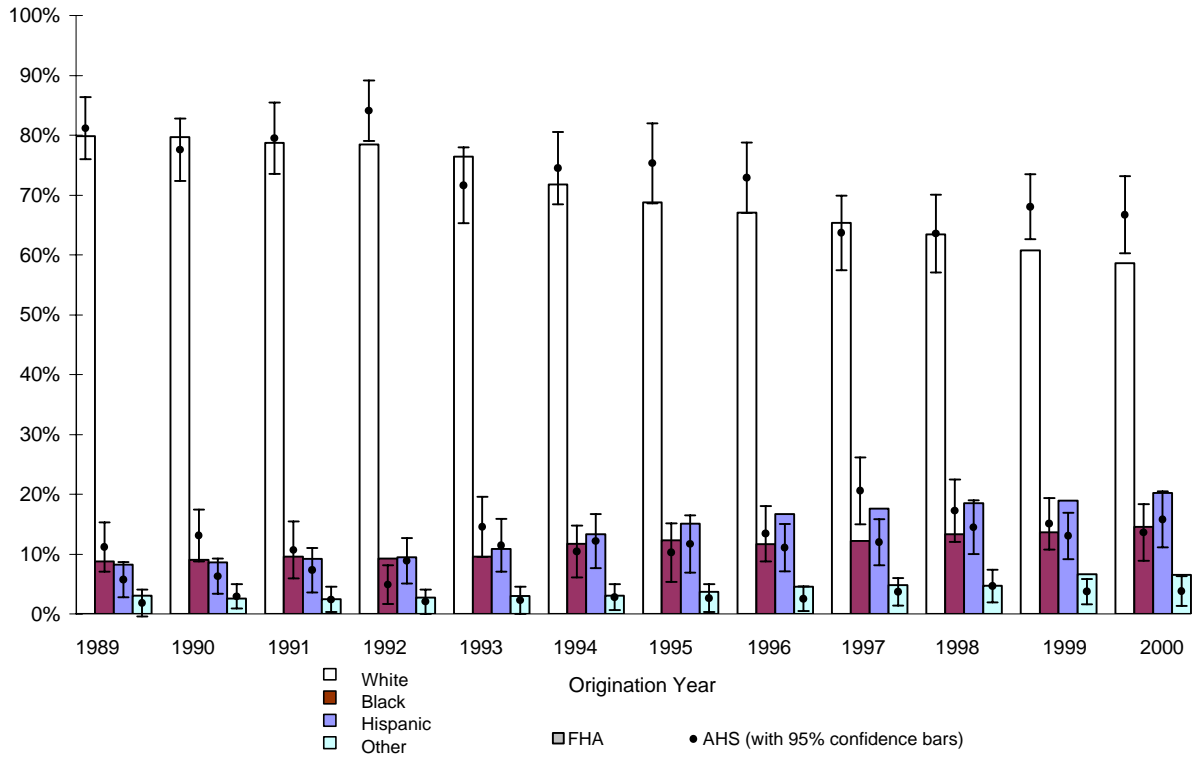
For the purpose of this analysis, borrowers are grouped into four mutually exclusive categories in all three data sets:

- Non-Hispanic white;⁵⁴
- Non-Hispanic black;
- Hispanic; and
- Other.

Borrowers who did not provide race/ethnicity information in either the AHS or comparison databases are excluded from this analysis.

Comparisons for racial composition (in terms of shares) of borrowers who took out home-purchase mortgages are presented in Exhibits 4-32 to 4-35 for the four market segments.

Exhibit 4-32:
Share of Home-Purchase FHA-Insured Mortgages by Race/Ethnicity
AHS vs. FHA Data



⁵⁴ Although Hispanic origin and other racial identities are reported as mutually exclusive categories in HMDA, we still cannot tell whether a borrower who identifies himself/herself as white or black is in fact non-Hispanic.

Exhibit 4-33:
Share of Home-Purchase VA/FmHA Mortgages by Race/Ethnicity
AHS vs. HMDA Data, Metropolitan Areas Only

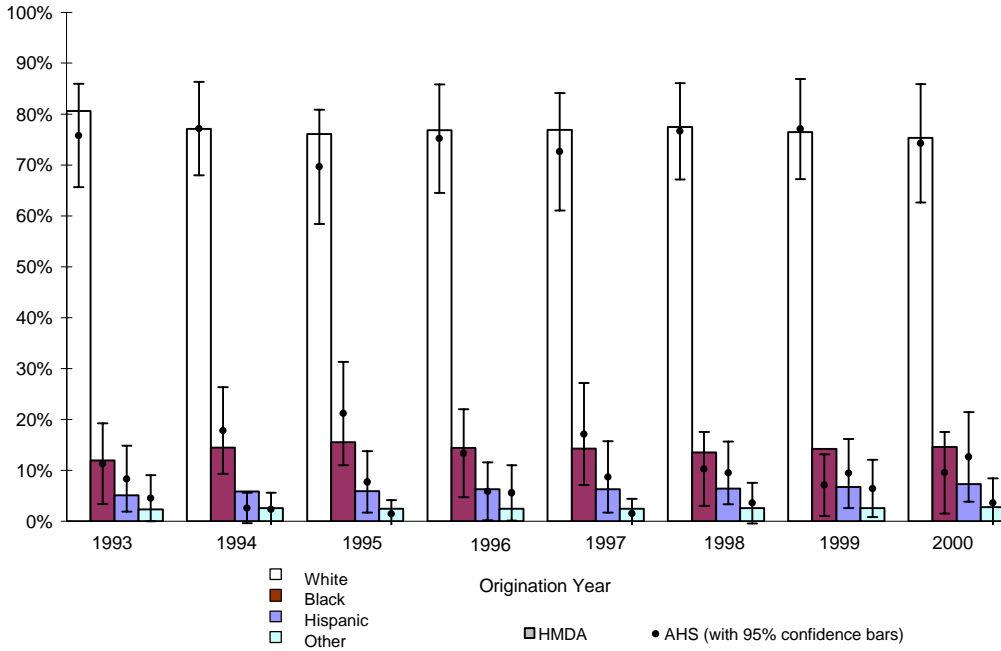
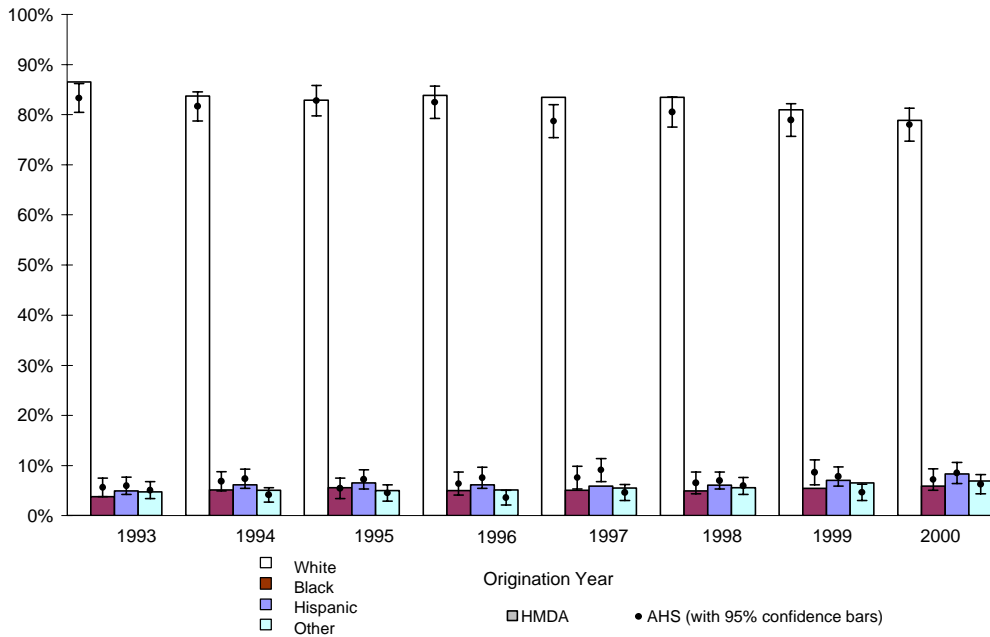
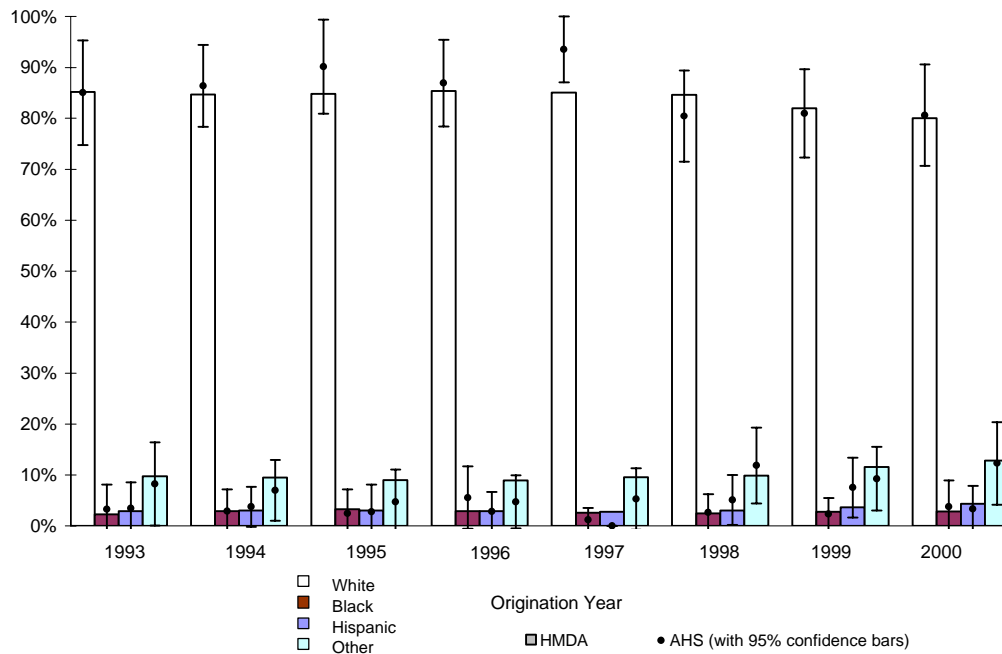


Exhibit 4-34:
Share of Home-Purchase Conforming Conventional Mortgages by Race/Ethnicity
AHS vs. HMDA Data, Metropolitan Areas Only



**Exhibit 4-35:
Share of Home-Purchase Jumbo Conventional Mortgages by Race/Ethnicity
AHS vs. HMDA Data, Metropolitan Areas Only**



When compared to FHA benchmarks, the AHS appears to do reasonably well in describing the racial composition of borrowers in that segment of the market. This is especially the case if the sampling variability of the AHS estimates is taken into account. FHA benchmarks are located within the boundaries of the AHS confidence intervals for most of the origination cohorts examined, as Exhibit 4-32 indicates. According to the FHA data, the share of white borrowers who took out FHA mortgages has experienced a steady decline over the decade. At the same time, shares of loans taken out by racial minorities (especially by Hispanics in recently years) are rising. Estimates from the AHS reflect this overall trend.

Some discrepancies are worth pointing out. Focusing on the point estimates of borrower shares, we found that the AHS has overstated the proportion of loans taken out by blacks in eight out of the twelve cohorts, while the shares by Hispanic and “other” borrowers are slightly understated in all cohorts except one. To quantify the differences in borrower share estimates between the two data sources, we calculate the average ratio of the AHS estimate to the FHA benchmark, which indicates the percentage difference in the estimates. For shares associated with black borrowers, the average ratio is 1.15 across the twelve cohorts, while the corresponding estimate for Hispanic and other shares is 0.8. While the difference in terms of ratios appears to be large, however, the absolute magnitude of the discrepancies is not large. For example, over the 1989-2000 period, the average market share for black borrowers was 13 percent in the AHS compared to 11 percent in the FHA data. The average

ratio of share estimates for white borrowers between the two databases is 1.04 over the decade, an indication that the AHS has captured the share of such borrowers quite accurately.⁵⁵

For home-purchase loans in the non-FHA segments of the mortgage market, racial distributions reflected in the AHS are remarkably close to the HMDA benchmarks. To identify the general tendency of discrepancies in borrower share estimates between the two databases, we again compute the ratio of AHS point estimate to HMDA benchmark for each of the three market segments. Results are displayed in Exhibit 4-36 and they seem to indicate that the AHS has a tendency to overstate the borrower shares for Hispanics and blacks in all three markets. However, these discrepancies most likely result from the fact that race/ethnicity categories are defined differently across the two data sets, notwithstanding our reconciliation efforts. In other words, they are not an indication that the AHS reporting is inaccurate.

**Exhibit 4-36:
Average Ratio of AHS Borrower Share Estimate to Benchmark:
Home-Purchases Originated between 1993-2000**

	MARKET SEGMENT			
	FHA ^a	VA/FmHA ^b	Conventional Conforming ^b	Conventional Jumbo ^b
Non-Hispanic White	1.04	0.97	0.97	1.02
Non-Hispanic Black	1.15	0.95	1.34	1.10
Hispanic	0.8	1.29	1.20	1.11
Other	0.8	1.43	0.88	0.77

Sources: Authors' tabulation of the 1991-2001 AHS, 1993-2000 HMDA, 1989-2000 FHA administrative data.

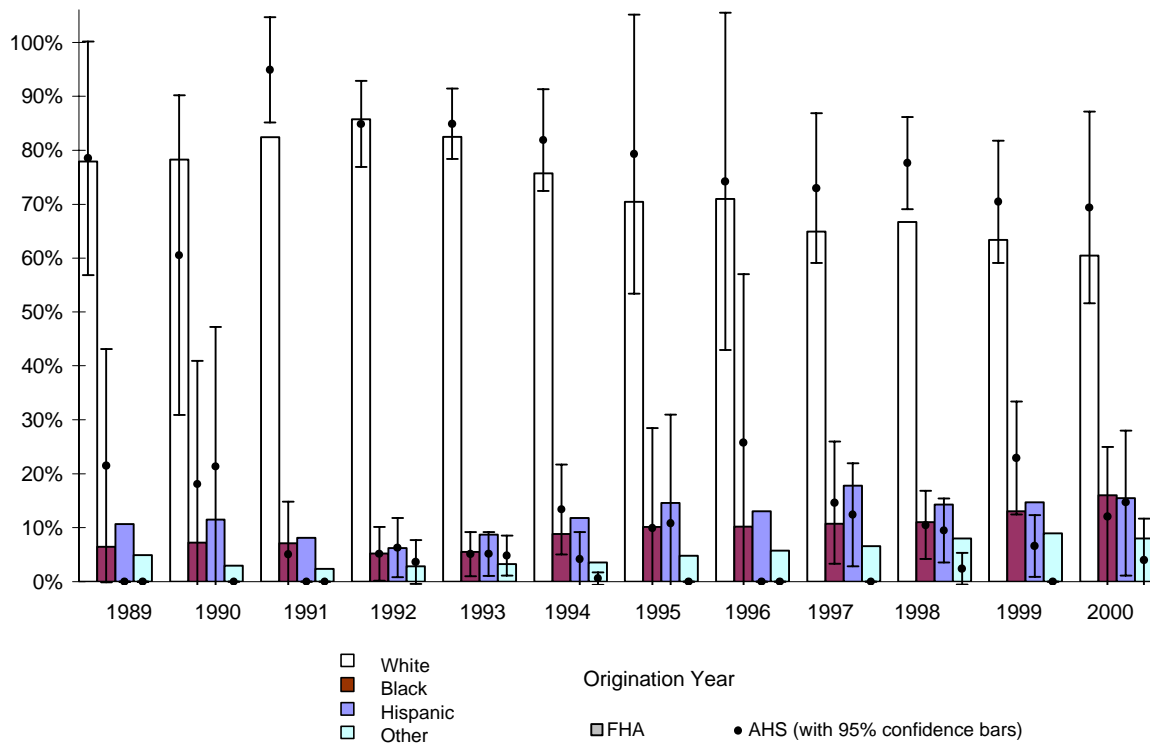
Notes: ^aBased on data from the AHS and FHA administrative records for the whole nation.

^bBased on data from the AHS and HMDA files for loans made in metro areas.

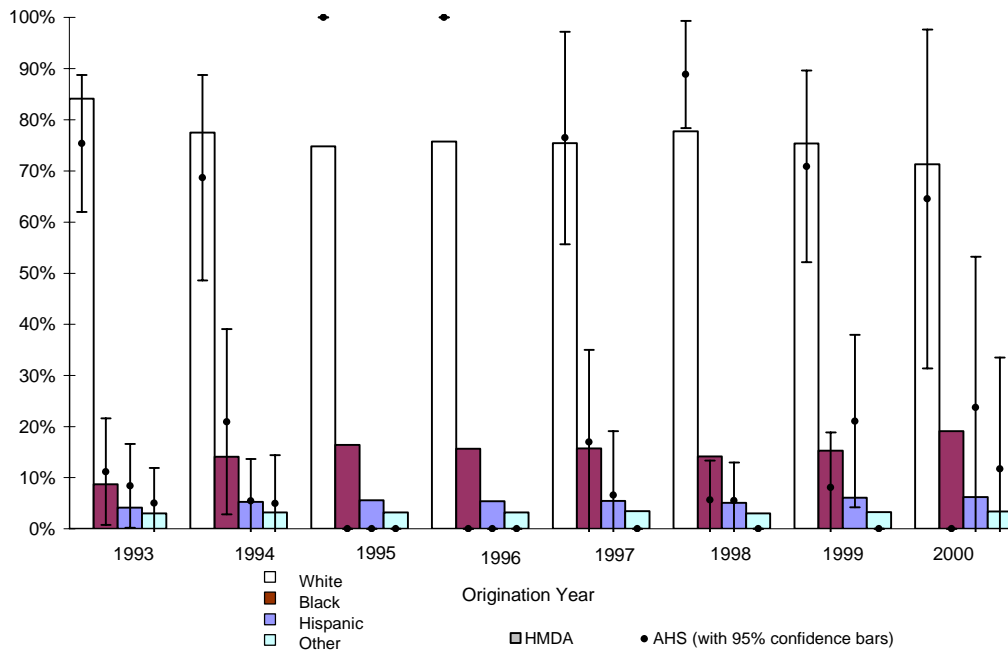
⁵⁵ The FHA administrative data also allow us to cross check the accuracy of race information collected in HMDA. Tabulations are restricted to loans made in the metro areas. Results for the 1993 to 2000 origination cohorts are presented as parts of Exhibit D-31-38, with separate analyses for home-purchase and refinance mortgages. The borrower distribution by race in HMDA was found to track remarkably well with the FHA benchmark for each cohort of home purchases and refinances.

Comparisons for borrower shares of refinance mortgages by race/ethnicity are presented in Exhibits 4-37 to 4-40, separately for each of the four market segments. It is virtually impossible to draw any statistically valid findings for borrower shares of refinance loans in the FHA, VA/FmHA and conventional jumbo segments of the market, given the wide confidence interval associated with the AHS estimates. Therefore, we focus our attention on the conventional conforming originations of refinance loans. As Exhibit 4-39 reveals, it appears that the AHS has exaggerated slightly the share of borrowers who took out such loans who are white. In seven out of the eight cohorts examined, the share of white borrowers estimated from the AHS is higher than the HMDA benchmark.

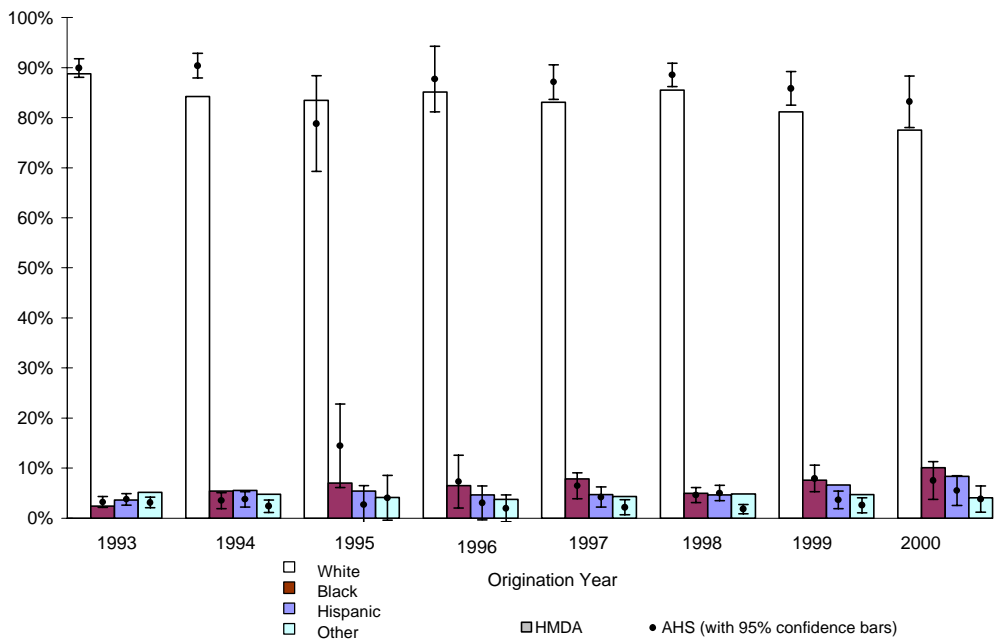
Exhibit 4-37:
Share of Refinance FHA-Insured Mortgages by Race/Ethnicity
AHS vs. FHA Data



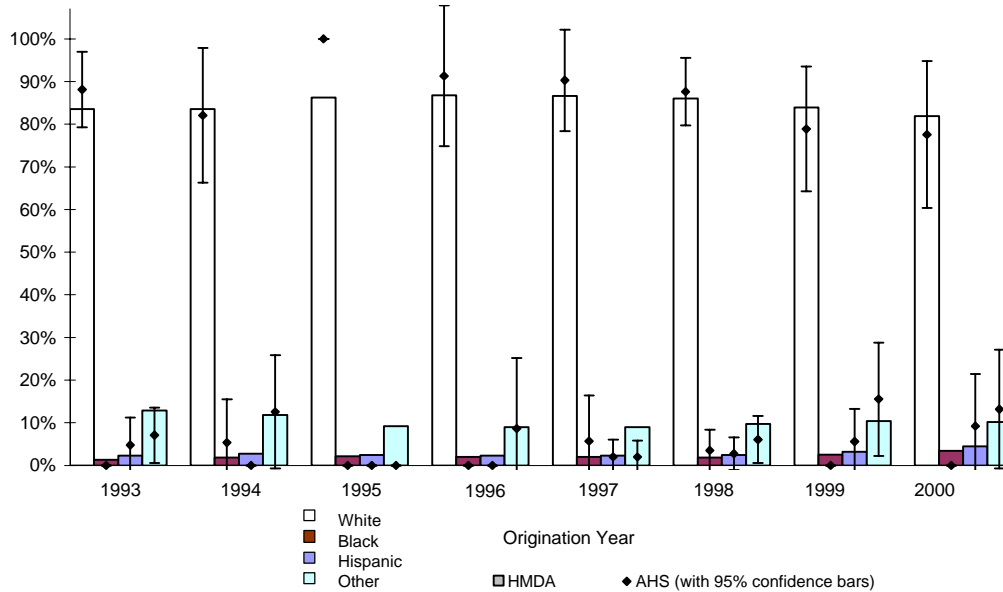
**Exhibit 4-38:
Share of Refinance VA/FmHA Mortgages by Race/Ethnicity
AHS vs. HMDA Data, Metropolitan Areas Only**



**Exhibit 4-39:
Share of Refinance Conforming Conventional Mortgages by Race/Ethnicity
AHS vs. HMDA Data, Metropolitan Areas Only**



**Exhibit 4-40:
Share of Refinance Jumbo Conventional Mortgages by Race/Ethnicity
AHS vs. HMDA Data, Metropolitan Areas Only**



Despite our attempt to create consistent categories for comparison, we suspect that reporting differences result in different self-characterized categories across the data sets. For instance, when asked a collapsed race/ethnicity question in the HMDA and FHA data, we suspect that black Hispanics call themselves Hispanics and Hispanic whites identify themselves as whites. This would explain the differences in share estimates. It is also important to note how the race/ethnicity data are collected in these different mortgage databases. In AHS the information is self-reported by the respondents. For FHA and HMDA, lenders report the data. Lenders may either ask the borrower his or her race/ethnicity or may complete this field based on the lender’s observation of the borrower. Because Hispanic origin is not necessarily apparent based on skin color, lenders who report race/ethnicity based on their own observations may be prone to misclassify Hispanics. Because of these concerns, discrepancies in race/ethnicity distribution between the AHS and HMDA/FHA should not be interpreted as an indication that the AHS data are not accurate.

Income

In general, respondents in household surveys have a tendency to underreport incomes. The Statistical Research Division of the Census Bureau has carried out a couple of studies in the early 1990s comparing the aggregate income of all households reported in the AHS with benchmarks derived from four independent sources: GNP accounts, the Social Security Administration, the Veterans Administration, and the Current Population Survey (CPS). The AHS estimates were found to be

consistently lower.⁵⁶ Recent research by Scott Susin of the Census Bureau confirms this finding. He found that average household income in 1999 is 9 percent lower in the AHS than in the CPS, while family earnings are almost the same between the two data sets. The problem is most likely a failure of many respondents to report any non-wage income, rather than the wrong amounts being reported. Average non-wage income is 32 percent lower in the AHS than in the CPS. The largest potential sources of underreported no-wage income are interest, dividends, social security and pensions, “other income,” and business income. Moreover, the discrepancy has become worsened over time, especially since the 1995 survey. Susin’s analysis speculates that this might be the result of the move from paper to computerized questionnaires (CATI).⁵⁷

Testing the accuracy of the income distribution of homeowners in the AHS data against benchmarks from other mortgage market databases presents a challenge. While borrower’s income is available from many of the mortgage market data considered in this study (including HMDA and FHA administrative data), it is measured quite differently across databases. In the AHS survey, respondents are asked to report the gross income of the household, which includes any wage and salary income, welfare receipts, rental income, and dividends from stocks and interest from savings accounts for all household members, before any deductions such as taxes, Social Security, and union dues.⁵⁸ The income information in the FHA database, on the other hand, consists of income of the borrower *net of taxes*, according to the data dictionary.⁵⁹ Clearly, the two are not comparable, since in general after-tax income is smaller than gross income.⁶⁰ The income definition in HMDA appears to be somewhat closer to the AHS definition.⁶¹ Nonetheless, in many cases, it appears that the information captured in the HMDA income variable may not be gross income for the entire household. According to the reporting guidelines, the income in HMDA should reflect the gross

⁵⁶ See *Current Housing Report, No. H121/95-1, American Housing Survey: A Quality Profile*, by Rameswar P. Chakrabarty, Census Bureau (1996); and “Comparison of 1989 AHS and CPS Income Reporting.” Internal Census Bureau memo, by Barbara T. Williams (1992).

⁵⁷ Scott Susin, *Discrepancy Between Measured Income in the American Housing Survey (AHS) and the Current Population Survey (CPS): Final Report*. U.S. Census Bureau, March 27, 2003. It is worth noting that the CPS questionnaire is more detailed than the AHS instrument in terms of income reporting, especially regarding components of non-wage income: the CPS asks about the receipt and amounts of 19 types of non-wage income for each person, while the AHS asks about the receipt of 9 categories of non-wage income, and a total amount for each family.

⁵⁸ In the AHS, there is a distinction between household income (ZINC2) and family income (ZINC) in the public-use file. The former measure may include incomes from non-family members or non-relatives who reside in the same household. This analysis makes use of ZINC2 because of the smaller sample size associated with the ZINC variable. Our exploratory work using ZINC has generated tabulations that are very similar to those presented in this chapter.

⁵⁹ In addition, the field is not populated for the cohorts of loans originated in the early 1990s. The FHA started collecting borrower income information in electronic format in the F-42 MIS system in 1993.

⁶⁰ An important exception in this context is for very-low-income households taking the Earned Income Tax Credit (EITC). In general, EITC boosts their after-tax income.

⁶¹ *A Guide to HMDA Reporting: Getting It Right!* Federal Financial Institutions Examination Council (FFIEC), 1998. For definition of income in HMDA, see page A-15 of Appendix A and page D-12 of Appendix D of that guidebook.

income that the individual financial institution relied upon in making the credit decision regarding the mortgage application. The guidelines do not specify clearly what income components should be included. When a mortgage is taken out by a dual-earner household, for example, it is conceivable that income of the secondary earner may not be reported if it is not needed to support the creditworthiness of the loan application. For the same reason, the HMDA income may also exclude other miscellaneous and non-salary components that are part of the AHS income measure.

Nevertheless, given the prominence of HMDA data in mortgage market research, it is valuable to check how the income distribution of borrowers reported in HMDA compares with the distribution from the AHS.

Exhibits 4-41 to 4-44 display a comparison of the average borrower income for home-purchase originations between the two data sets, stratified by mortgage market segment. Once again, all comparisons are restricted to loans made in metro areas. We found that average income in the AHS does not track well with average income in HMDA for the vast majority of cohorts of loans examined, regardless of market segment. Except for loans in the conventional jumbo segment of the market, homeowners in AHS have on average reported a higher level of income relative to the HMDA benchmark. The gap appears to be widening in the more recent loan cohorts. Overall, this is consistent with our hypothesis that household income reported in HMDA sometimes contains fewer components. Findings for conventional jumbo purchases are mixed. When compared to the HMDA estimates for such loans in the 1993 to 1996 cohorts, average borrower income was consistently lower in the AHS. This finding is contrary to the pattern observed in the other segments of the market (FHA, VA/FmHA, and conventional conforming purchases) and probably results from the fact that the AHS income variable is top-coded for confidentiality reasons. Estimates for jumbo purchases for the remaining four cohorts (1997 and 2000) are quite close between the two data sources.

Additional statistics (25th, 50th and 75th percentiles) that show income distributions from the two databases are provided in Exhibit D-41-44 and confirm these findings.

Analysis for borrowers who took out refinance mortgages is not presented here. Because of the relatively small number of such originations reported in the AHS, it is difficult to draw any statistically reliable comparisons with the estimates from HMDA. Interested readers can find the tabulated results of such analysis in Exhibits D-41-44 of Appendix D.

Exhibit 4-41:
Average Borrower Income of
Home-Purchase FHA Mortgages
AHS vs. HMDA Data, Metropolitan Areas

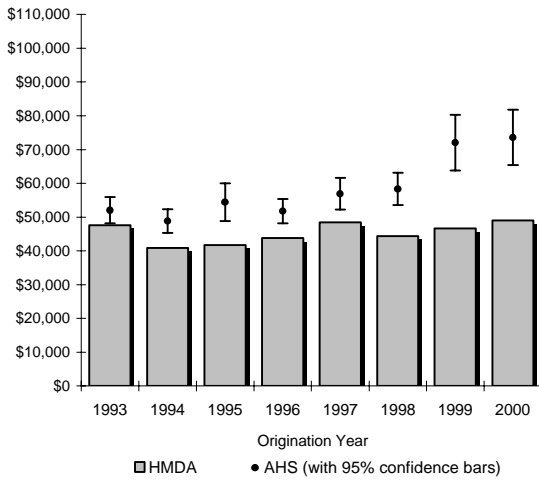


Exhibit 4-42:
Average Borrower Income of
Home-Purchase VA/FmHA Mortgages
AHS vs. HMDA Data, Metropolitan Areas

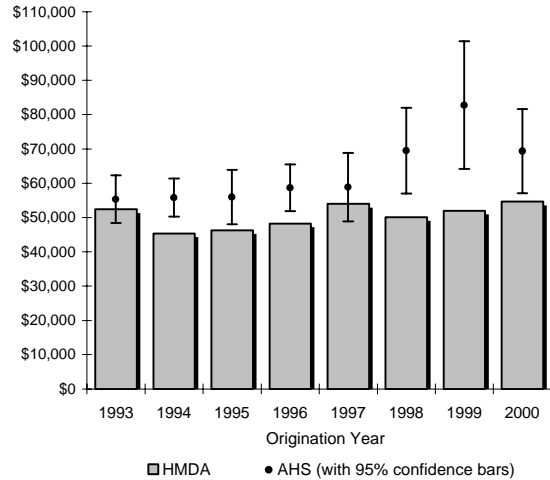


Exhibit 4-43:
Average Borrower Income of
Home-Purchase Conforming
Conventional Mortgages
AHS vs. HMDA Data, Metropolitan Areas

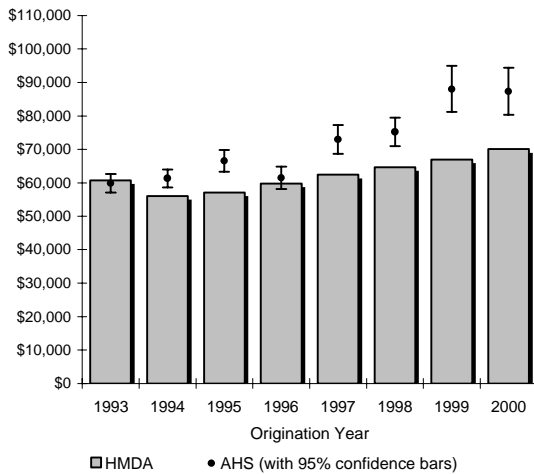
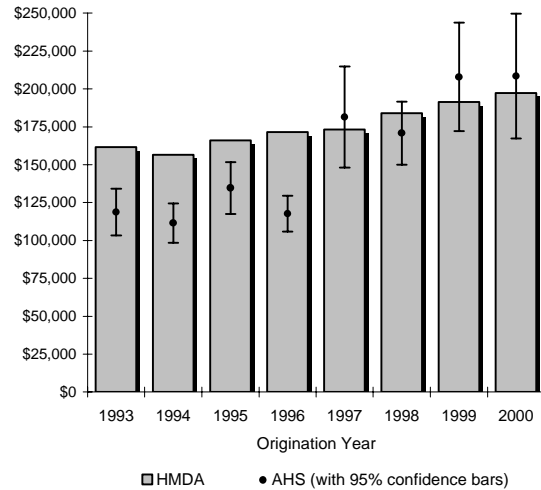


Exhibit 4-44:
Average Borrower Income of
Home-Purchase Jumbo Conventional Mortgages
AHS vs. HMDA Data, Metropolitan Areas



Summary of Major Findings on AHS Mortgage Variables Tested in the Replication Analysis

Exhibit 4-45 summarizes the analyses presented in this chapter. It identifies the AHS mortgage variables we have benchmarked, the comparison databases used, the cohorts of mortgage activity

covered, and any stratification needed when the replication analysis was performed. Major findings for each variable are displayed in the last column of the Exhibit.

**Exhibit 4-45:
Major Findings for AHS Mortgage Market Measures Tested in the Replication Analysis**

MORTGAGE MARKET MEASURE	BENCHMARK DATA & ORIGINATION COHORTS EXAMINED	STRATIFICATION	MAJOR FINDINGS
Mortgage Origination Volume	<ul style="list-style-type: none"> • FHA: 1989-2000 for FHA • HMDA: 1993-2000 for VA/RHS, Conventional Conforming and Conventional Jumbo 	<ul style="list-style-type: none"> • Mortgage Market Segment^a • Loan Purpose^b 	<p>Home-purchase loans:</p> <ul style="list-style-type: none"> • Generally accurate for volume estimates of origination cohorts. Over the 1993 to 2000 period, the average coverage ratio was 99 percent for all loans combined (Exhibits 4-3, 4-4, and 4-5). • Coverage ratio declined for the 1998 to 2000 cohorts. Average coverage ratio was in the 83-89 percent range for all loans combined during that period (Exhibit 4-4). • Relatively low volume estimates for conventional jumbo originations (Exhibit 4-5). <p>Refinance loans:</p> <ul style="list-style-type: none"> • Low volume estimates across all mortgage market segments. Average coverage ratio was 48 percent for 1993 to 2000 (Exhibits 4-7, 4-8, and 4-9). • Especially low estimates for cohorts since 1993 (1995 survey) probably because of question wording (a lack of clear distinction between cash-out refinance mortgages and lump-sum home equity loans) or coding errors in the survey (Exhibit 4-8). • Loan purpose questions added in the 2001 survey did not significantly improve the survey's ability to capture refinance loans (Exhibit 4-8).

MORTGAGE MARKET MEASURE	BENCHMARK DATA & ORIGINATION COHORTS EXAMINED	STRATIFICATION	MAJOR FINDINGS
<i>Loan Characteristics</i>			
Mortgage Market Segment	<ul style="list-style-type: none"> • RFS: 1989-1991 • HMDA: 1993-2000 • SCF: 1993-1994 	<ul style="list-style-type: none"> • Loan Purpose 	<p>Home-purchase loans:</p> <ul style="list-style-type: none"> • For loans originated in 1989-1991, slight underestimate of FHA share and overestimate of conventional conforming share (Exhibit 4-10). • Accurate market shares for loans in the FHA, VA/RHS and conventional segments, 1993-2000. Discrepancies are mostly in the 1 to 2 percent range (Exhibit 4-11). • Slight underestimate of market share of conventional jumbo loans and overestimate of conventional conforming loans for most of the cohorts (Exhibit 4-11). <p>Refinance loans:</p> <ul style="list-style-type: none"> • Small sample sizes result in wide confidence intervals that make it difficult to draw statistically reliable comparisons. • Accurate shares for loans in the 1989-1991 cohorts (Exhibit 4-10). • Underestimate of shares for conventional jumbo loans for most cohorts (Exhibit 4-12).
Origination Loan Amount	<ul style="list-style-type: none"> • FHA: 1989-2000 for FHA • HMDA: 1993-2000 for VA/RHS, Conventional Conforming 	<ul style="list-style-type: none"> • Mortgage Market Segment • Loan Purpose 	<p>Home-purchase loans:</p> <ul style="list-style-type: none"> • Accurate for loans in the FHA market segment (Exhibit 4-13). • Slight underestimate, by about 10 percent, of loan amounts in the VA/RHS segment (Exhibit 4-15). • Accurate for 1995, 1997, 1998, 1999 and 2000 conventional conforming loans. Underestimate, by about 10 percent, of those in 1993, 1994 and 1996 cohorts (Exhibit 4-15). <p>Refinance loans:</p> <ul style="list-style-type: none"> • Underestimate of loan amounts of FHA loans originated in 1995 and 1996 (Exhibit 4-14). • Small sample sizes for VA/RHS loans (Exhibit 4-16). • Accurate for 1995, 1997 and 1998 conventional conforming loans. Underestimate for 1993, 1994 and 1996 cohorts. Overestimate for 1999 and 2000 cohorts (Exhibit 4-16).

MORTGAGE MARKET MEASURE	BENCHMARK DATA & ORIGINATION COHORTS EXAMINED	STRATIFICATION	MAJOR FINDINGS
Mortgage Interest Rate (Fixed-Rate Loans)	<ul style="list-style-type: none"> FHA: 1989-2000 for FHA MIRS: 1992-2000 for Conventional Conforming and Conventional Jumbo Home Purchases 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose Payment Product Type^c 	<ul style="list-style-type: none"> Accurate averages for both home-purchase and refinance loans in segments of the market examined. Discrepancy in the averages is generally one to two tenths of a percent. Overall distribution appears to be more dispersed (Exhibits 4-17 and 4-18). Compared to MIRS, the AHS reports a higher proportion of conventional conforming purchases with relatively high interest rates, probably because subprime loans are included in the AHS but not in MIRS (Exhibits 4-19 and 4-20).
LTV at Origination	<ul style="list-style-type: none"> FHA: 1989-2000 for FHA Home Purchases MIRS: 1992-2000 for Conventional Conforming Home Purchases 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> Reasonably accurate mean LTVs by cohort for FHA home purchases. Distribution of loans by LTV categories does not track well with benchmarks (Exhibits 4-21 and 4-22). Reasonably accurate mean LTVs by cohort for conventional conforming purchases. Distribution of loans by LTV categories does not track well with benchmarks (Exhibits 4-23 and 4-24). LTVs in the other segments of the mortgage market cannot be tested because of the lack of comparable benchmarks.
Mortgage Payment Product Type	<ul style="list-style-type: none"> FHA: 1989-2000 for FHA MIRS: 1992-2000 for Conventional Conforming and Conventional Jumbo Home Purchases 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> Understates ARM shares for FHA loans (Exhibits 4-25 and 4-26). Overstates shares of loans other than fixed-rate and ARMs in the FHA segment (Exhibits 4-25 and 4-26). Understates ARM shares for conforming and conventional jumbo home purchases (Exhibits 4-27 and 4-28).
<i>Borrower Characteristics</i>			
First-time Homebuyer Status	<ul style="list-style-type: none"> Chicago Title: 1995-1999 RFS: 1989-1991 FHA: 1991-2000 for FHA 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> Estimate of first-time buyer shares for all home purchases consistent with RFS and Chicago Title data (Exhibits 4-29 and 4-30). While differences with RFS are not statistically significant, the share of first-time buyers in AHS is lower for FHA and conventional jumbo borrowers but about equal for VA/RHS and conventional conforming (Exhibit 4-30). Discrepancies between the AHS and FHA data are due to how first-time buyer status is defined. The AHS estimates are more consistent with a strict notion of first-time buyers.

MORTGAGE MARKET MEASURE	BENCHMARK DATA & ORIGINATION COHORTS EXAMINED	STRATIFICATION	MAJOR FINDINGS
Borrower Race/Ethnicity	<ul style="list-style-type: none"> FHA: 1989-2000 for FHA HMDA: 1993-2000 for VA/RHS, Conventional Conforming and Conventional Jumbo 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> Different race/ethnicity questions and method of gathering information between AHS and administrative benchmark data sets make interpretation of differences difficult. Overall, distributions are reasonably close to the benchmarks (Exhibits 4-32 to 4-40).
Borrower Income	<ul style="list-style-type: none"> HMDA: 1993-2000 for FHA, VA/RHS, Conventional Conforming and Conventional Jumbo 	<ul style="list-style-type: none"> Mortgage Market Segment Loan Purpose 	<ul style="list-style-type: none"> Incomes are defined differently between AHS and benchmark data, making it difficult to interpret discrepancies. Household incomes in AHS are consistently higher than in HMDA, probably because in HMDA some borrowers do not need to report all income components to lenders to qualify for loans (Exhibits 4-41 and 4-44).

Notes: ^aMortgage market segments are FHA, VA/FmHA/RHS, conventional conforming, and conventional jumbo.

^bLoan purposes are home purchase and refinance mortgages.

^cPayment product types are fixed-rate and adjustable rate mortgages.

Chapter Five

Internal Consistency of AHS Mortgage Variables Across Surveys

This chapter presents the study's findings on the internal consistency of the key AHS mortgage data elements across survey years. We examine whether respondents are consistent in their answers to questions regarding mortgages that are still in place in successive surveys. The first section discusses the goals and analytic approach. The findings on consistency of the mortgage variables are presented in the following section.

Goals and Approach

A key feature of the AHS design is that it is a panel survey of a sample of housing units. In each AHS survey year, updated information on the housing units, households, and mortgages is collected from the same sample of housing units. Unique identifiers (unit control numbers) in the public-use files allow analysts to link these observations across surveys.⁶² An important consequence of this feature for mortgage-related research is that researchers can monitor the use of mortgage debt over the course of an owner's occupancy of a given unit. This longitudinal feature is a strength of the AHS compared with other mortgage market databases because most of them only capture loan characteristics at origination. For these mortgage databases, it is not possible to identify the characteristics of loans that were replaced by a refinance, for example, or to examine other dynamic housing finance decisions made by homeowners.⁶³ In contrast, the AHS is well suited for this type of analysis. Our preliminary literature review indicates that the mortgage variables in AHS have mostly been used for such analyses.⁶⁴

In order for longitudinal analysis to be valid, homeowners residing in the sample housing units must report consistent information about their mortgages across survey years. Analysis of internal

⁶² Beginning in 1985, the Census Bureau used a new and redesigned sample selected from the 1980 Census. Therefore, it is not possible to link AHS observations since then to those collected in the pre-1985 era.

⁶³ Notable exceptions are the Census Bureau's Survey of Income and Program Participation (SIPP), the Panel Study of Income Dynamics (PSID) conducted by the Survey Research Center at the University of Michigan, and the FHA administrative data maintained in the HUD Single-Family Data Warehouse (SFDW). Both the SIPP and PSID data have relatively small sample sizes compared to the AHS. While the mortgage payment and termination status of the FHA mortgages is updated regularly in the SFDW, all the borrower (such as household income) and loan (such as interest rate) characteristics are only captured at the time of mortgage origination. In addition, the FHA data do not report whether a prepayment results from a house sale or refinancing, and, in the case of a refinance, what the new loan is.

⁶⁴ For example, Wayne R. Archer, David C. Ling and Gary A. McGill, "The effect of income and collateral constraints on residential mortgage terminations." *Regional Science and Urban Economics*, 26: 235-261 (1996); Wayne R. Archer, David C. Ling and Gary A. McGill, "Demographic versus option-driven mortgage terminations." *Journal of Housing Economics*, 6: 137-163 (1997).

consistency can also be viewed as an assessment of the quality of the AHS mortgage variables for analysis of the stock of mortgages that are in place at a particular time. It sheds light on whether these cross-sectional estimates, especially on seasoned loans, are likely reliable when used at the household or housing unit level. The previous chapter evaluated the reliability of mortgage variables at origination. The findings from this chapter will clarify whether there is any notable degradation in the quality of the AHS variables as the time from origination increases. In sum, our analysis is a validation of the use of the AHS as a micro data set in general.

Our investigation focuses on the AHS national surveys conducted in the 1990s.⁶⁵ To take maximum advantage of the AHS data on loans originated during the 1990s, we have constructed four study panels. Each panel tracks a cohort of loans originated in the period shortly before the interview across subsequent survey years. Specifically, they are:

- A 1991-1993-1995-1997-1999 panel for mortgages originated in the 1989-1990 period
- A 1993-1995-1997-1999 panel for mortgages originated in the 1991-1992 period
- A 1995-1997-1999 panel for mortgages originated in the 1993-1994 period
- A 1997-1999 panel for mortgages originated in the 1995-1996 period

Only loans originated in the two-year period prior to the interview were included, because these homeowners are expected to remember more accurately facts about their mortgages than those with mortgages originated in earlier years. Our analyses also assume that mortgage information provided by the homeowners in the base survey year of each panel is likely to be the most accurate report.⁶⁶

Each panel is used to examine whether the mortgage-related information from the first survey year is reported consistently for the same mortgage across subsequent survey years.

Construction of the AHS Panels

In this section, we provide a brief description on how we linked observations from multiple years of the AHS national surveys to form a longitudinal file, using the construction of the 1991-1993-1995-1997-1999 study panel as an example. Details of the data elements and merging steps are presented in Appendix E.

To construct this panel, we selected all the owner-occupied housing units with primary mortgages identified in the 1991 survey as having been originated in the 1989-1990 period. This resulted in a total of 3,528 records, representing an estimated 7,567,469 loans.

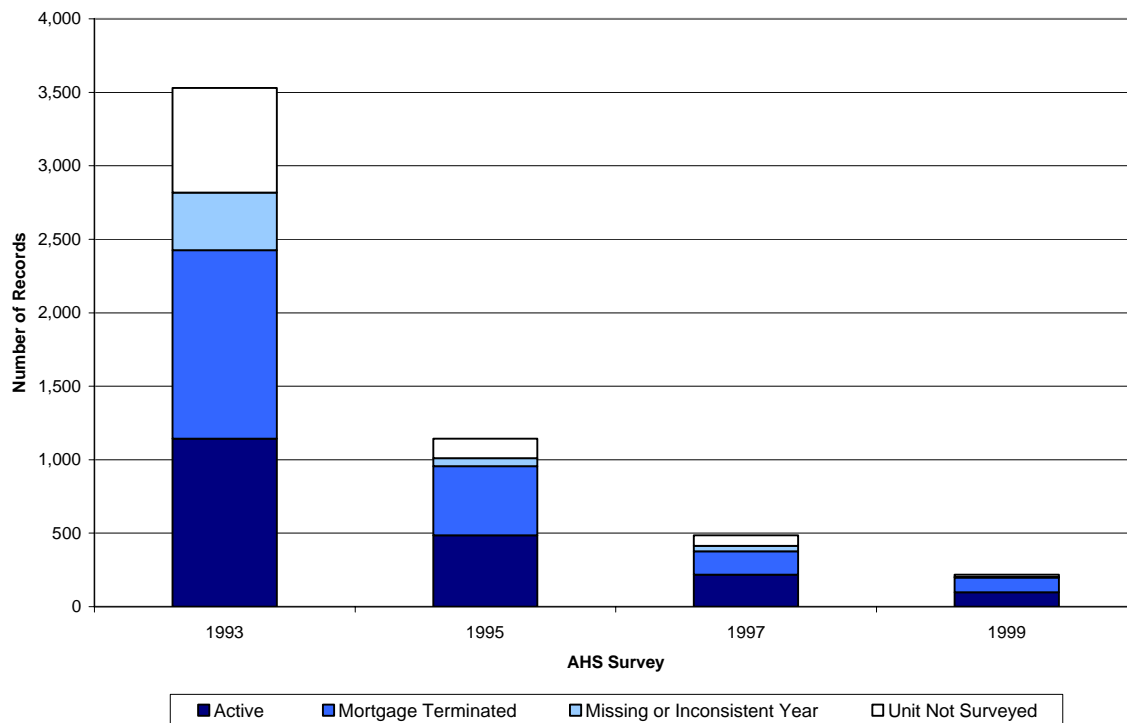
⁶⁵ For the 1999 survey, we used the public-use file released on November 14, 2002. The Census Bureau indicated that they have corrected the errors in the WHNGET variable associated with the previous release of the data.

⁶⁶ It is of course possible that some of first reports were incorrect and some of later reports provided by the same household in the subsequent survey years are in fact correct.

We then tracked the status of these loans in the 1993 survey. Since the AHS follows housing units rather than households or mortgages, a sequence of data steps is needed to ensure that we are following the same set of households and loans across survey years. Details are provided in Appendix E. In brief, records were merged to the later survey year file by unit control number, additional information on the mortgage, housing unit and household. The initial match by unit control number yielded a total of 2,824 records. Additional data steps were performed to ensure that these represent the same mortgages as those in the 1991 file. In the end, only 1,143 of the 3,528 loans still appeared active in the 1993 survey. There are three reasons why some of the loans reported in the 1991 survey would not be observed in the 1993 survey. First, a sizable portion of the housing units was not even included in the 1993 survey, primarily because they were part of a special over-sampled group in 1991. Second, a number of the mortgages were terminated because of household relocation, mortgage prepayments, or defaults. For the purpose of this analysis, if the date of mortgage origination reported in the successive survey was later than the date reported in the 1991 survey, we considered this a mortgage termination. Finally, for some observations the 1993 survey reported earlier (inconsistent) origination dates or did not provide sufficient information to identify whether the 1989-1990 loan was still active.

These 1,143 records were then matched to the 1995, 1997 and 1999 survey data. By 1999, only 99 records (representing an estimated 223,752 mortgages originated in the 1989-1990 period) are left in this study panel. Exhibit 5-1 shows, in a bar-chart format, the attrition of this cohort of loans as they are tracked across subsequent surveys.

Exhibit 5-1
Status of 1989-90 Mortgage Originations from 1991 AHS Over Subsequent Surveys



The other three study panels were created in a similar fashion. Appendix F documents the sample sizes and loan attrition situation for each panel.

This method of using the origination date information for mortgage tracking has its limitations. While we consider a change in the origination date to a later year as an indication of mortgage termination, it could be the result of an error in the homeowner's reporting. In this study, only cases where the origination date reported in the successive survey is earlier than the previously reported year are considered as clear errors. Furthermore, the consistency of loan purpose information (purchase vs. refinance) cannot be evaluated, since loan purpose is a user-derived variable based on the origination date. We cannot tell whether a homeowner refinanced the mortgage since the first survey or he/she reported the mortgage origination date incorrectly in one of the interviews.

Rather than relying on a comparison of the mortgage origination date across surveys, we might have used a set of loan characteristics (such as original mortgage amount, payment plan type and loan term), in addition to the unit control number, as matching criteria. However, we then would not have been able to assess whether the homeowners reported these mortgage characteristics consistently across survey years.

Consistency of Mortgage Variables Across Time

Mortgage Origination Year

If a household reports an earlier mortgage origination date in the subsequent survey than was reported in the base year survey of a study panel, we consider this an inconsistency. For example, in the 1993-1995-1997-1999 study panel (which tracks the cohort of loans originated in 1991-1992), 232 homeowners (16 percent of 1,409 active mortgage records) in the 1995 survey reported that they obtained their mortgages before the 1991-1992 period. This is inconsistent because in the 1993 survey these 232 borrowers reported that they obtained their mortgages in the 1991-1992 period. Exhibit 5-2 shows the proportion of such observations (among active loans) for each survey year of the corresponding study panel.

**Exhibit 5-2:
Inconsistent Origination Date
Proportion of Records (Among Active Loans) Reporting Origination Date
Earlier than the One Indicated in Base-Year Survey**

AHS Survey	Loan Cohort			
	89-90	91-92	93-94	95-96
93	9%	-	-	-
95	5%	16%	-	-
97	8%	16%	25%	-
99	1%	1%	2%	4%

Source: Authors' calculations of the AHS National Surveys

The results show a relatively high inconsistency rate for the 1991-1992 and 1993-1994 loan cohorts for information reported in the 1995 and 1997 AHS surveys. However, the 1995 and 1997 surveys provide much more reliable information on the 1989-1990 cohort, so it does not appear that the problem reflects general issues with those surveys. The inconsistency rates for all survey years for the 1989-1990 cohort are fairly modest, at less than 10 percent. Data from the 1999 AHS appear to yield the lowest inconsistency rates for all loan cohorts examined, all less than 4 percent. This improvement is most likely the result of the extensive use of “dependent interviewing” in that survey year (for the house acquisition year variable). This data collection technique was used to skip some of the survey questions, if logically the answers should not have changed since the previous survey year.

Other Mortgage Characteristics

This section discusses whether other mortgage-related information (besides origination year) was reported consistently across survey years. The variables we have examined are:

- Mortgage market segment (FHA, VA, FmHA, or conventional);⁶⁷
- Original mortgage amount;
- Mortgage type (payment plan);
- Mortgage term;
- Mortgage Interest rate (for fixed-rate mortgages only);⁶⁸
- Monthly mortgage payment amount (for fixed-rate mortgages only); and
- First-time homebuyer status.

Mortgage information reported in later survey years is compared to information in the first survey year of the study panel. In this context, consistency means that, among loans that did not terminate in subsequent surveys, the homeowner has provided the same answer about his/her mortgage as before. For example, from the 1991-1999 panel, observations from the 1993, 1995, 1997 and 1999 surveys were each compared to information reported in the 1991 survey.⁶⁹ For each *categorical* mortgage variable in each survey year, we computed the proportion of homeowners who reported exactly the same information as in the base year of that panel. For *continuous* variables (original mortgage amount, loan term, interest rate and monthly payment amount), a 15-percent difference was allowed. Interest rate information provided in the AHS public-use files has been rounded down to the nearest

⁶⁷ In this analysis, no distinction is made between conventional conforming and jumbo loans.

⁶⁸ AHS homeowners are asked only about the current interest rate on the mortgage. Given the lack of information about the timing of adjustments to ARMs or the index and margin used to make these adjustments, it is not possible to test the internal consistency of interest rate information for such loans.

⁶⁹ It is possible that those who provide inconsistent mortgage information vary from survey to survey. In other words, the respondents who misreport in one survey may not be the same as those who misreport in another survey, even though both surveys have the same percentages of consistent responses.

1/4 percentage point in the 1997 survey and previous surveys. In the 1999 survey, the variable has been rounded down to the nearest 1/8 percentage point.

A challenge for the evaluation of the interest rate (or monthly mortgage payment) information in this context is that variation in the reported amount over time may reflect either inconsistency in the owners' recall of the interest rate (or monthly mortgage payment) or errors in the reported type of mortgage product (FRM vs. ARM). While we are not able to distinguish between these types of errors, we can report on the prevalence of large changes in the interest rates or monthly mortgage payments of fixed-rate loans.

Exhibit 5-3 presents the estimates of the percentage of responses that are consistent with the base year for each of the four panels. The results are similar for all four panels. Therefore, we focus our discussion only on findings from the 1991-1999 panel.

Exhibit 5-4 presents the results for the 1991-1999 panel in a bar-chart format. It shows that, regardless of AHS years and mortgage variables considered, the vast majority of the consistency scores are in the 80-90 percent range. Confirming our expectation, homeowners remember information about their mortgages more accurately in the consecutive survey (1993) than in the later survey years.

Among the variables, first-time homebuyer status (87-90 percent) and mortgage term (88-92 percent) have the highest consistency scores, indicating the survey respondents tend to remember these two aspects of their mortgages most accurately over time. The mortgage market segment and payment plan variables have a decent level of internal consistency, almost always above 80 percent. In contrast, homeowners were more likely to provide inconsistent information about the interest rate and monthly payment amount of their fixed-rate mortgages in subsequent surveys. The consistency scores for the interest rate variable are in the 43-74 percent range, while the scores for the monthly payment amount variable are in the 68-85 percent range. For both variables, the proportion of respondents reporting information that was consistent with the base year drops dramatically in the third and fourth survey years. A similar decline of consistency scores was observed for the original mortgage amount variable, but the drop was relatively modest and a high proportion (86 percent) of respondents provided consistent information in the second survey year (1993).

It is important to note that, since monthly mortgage payments may include property taxes and homeowner's insurance, a portion of this inconsistency for the payment variable may simply reflect the fact that the payment for taxes and insurance has risen over time for some of these homeowners. The relatively low consistency rates for the payment and interest rate variables may also indicate that some borrowers had incorrectly identified their loans as a fixed-rate product in the base survey year. It is equally likely that some of those who reported inconsistent payments or interest rates may have correctly reported their loans as adjustable-rate in the later surveys, having incorrectly reported them as fixed-rate initially.

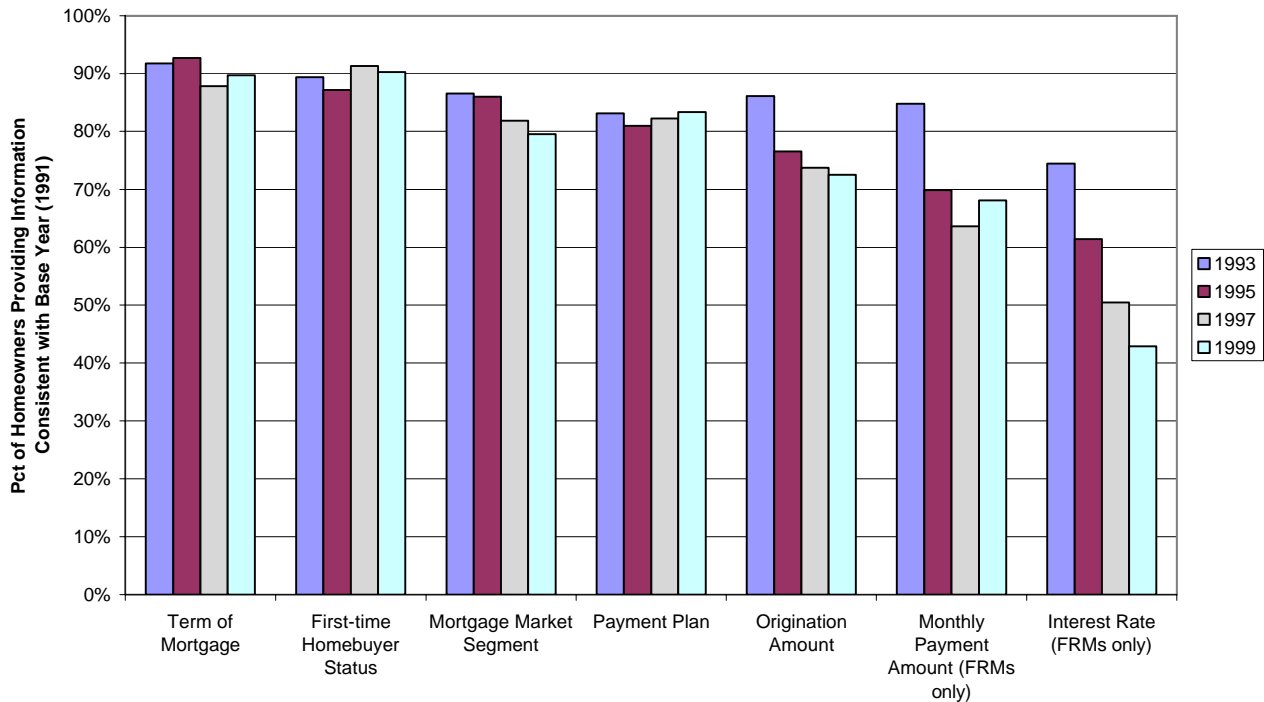
**Exhibit 5-3:
Consistency of Mortgage Variables Over Time
(Percentage of Responses Consistent with Base-Year Survey)**

	1991-1993-1995-1997-1999 Panel					1993-1995-1997-1999 Panel		
	1993	AHS Year		1999		1995	AHS Year	
		1995	1997			1997		
Mortgage Market Segment	87%	86%	82%	80%	Mortgage Market Segment	88%	82%	86%
Origination Amount	86%	77%	74%	73%	Origination Amount	75%	78%	72%
Term of Mortgage	92%	93%	88%	90%	Term of Mortgage	89%	90%	86%
Interest Rate (FRMs only)	74%	61%	50%	43%	Interest Rate (FRMs only)	81%	79%	74%
Payment Plan	83%	81%	82%	83%	Mortgage Type (Payment Plan)	87%	87%	85%
Monthly Payment Amount (FRMs only)	85%	70%	64%	68%	Monthly Payments Amount (FRMs only)	78%	77%	75%
First-time Homebuyer Status	89%	87%	91%	90%	First-time Homebuyer Status	88%	88%	88%
Number of Observations	1,143	486	219	99	Number of Observations	1,409	545	263

	1995-1997-1999 Panel			1997-1999 Panel
	AHS Year			AHS Year
	1997	1999		1999
Mortgage Market Segment	86%	85%	Mortgage Market Segment	82%
Origination Amount	75%	74%	Origination Amount	70%
Term of Mortgage	92%	87%	Term of Mortgage	85%
Interest Rate (FRMs only)	84%	82%	Interest Rate (FRMs only)	79%
Payment Plan	85%	85%	Payment Plan	84%
Monthly Payment Amount (FRMs only)	76%	74%	Monthly Payment Amount (FRMs only)	74%
First-time Homebuyer Status	88%	89%	First-time Homebuyer Status	100%
Number of Observations	1,423	718	Number of Observations	1,326

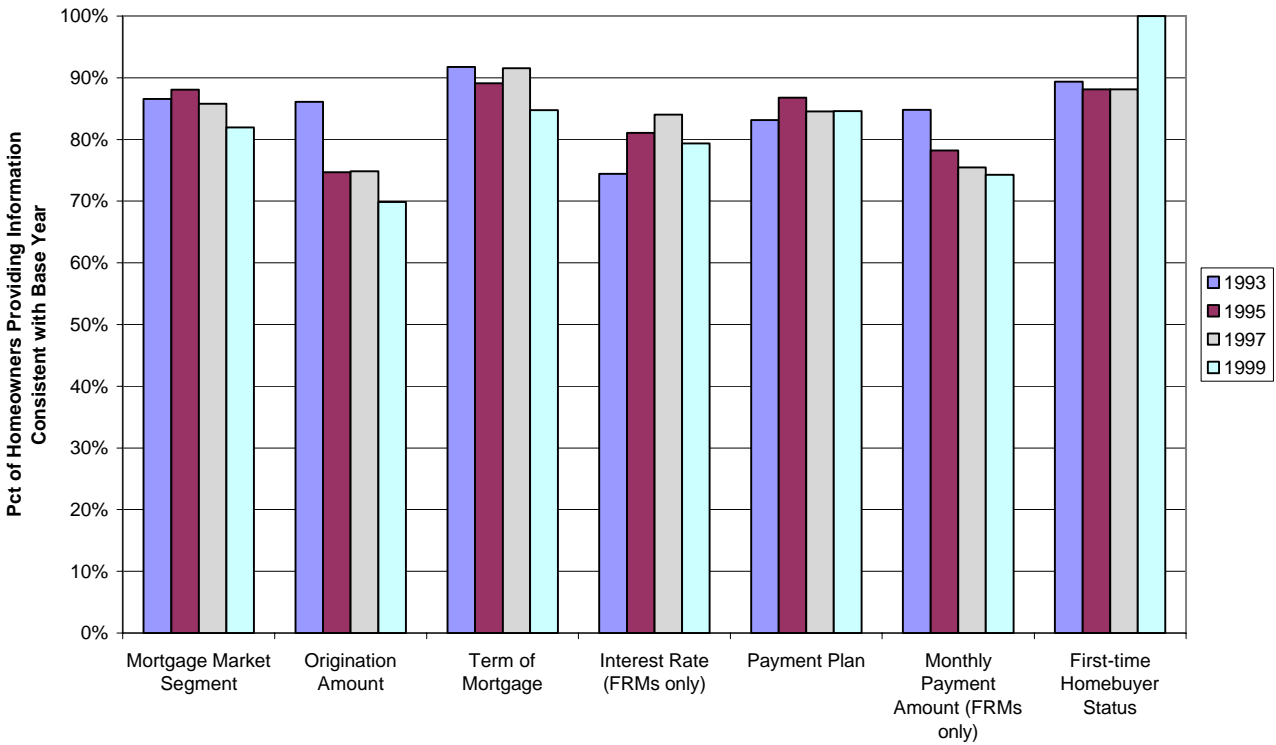
Source: Authors' calculations of the AHS National Surveys

**Exhibit 5-4:
Consistency of Mortgage Variables Across AHS Surveys
1991-1993-1995-1997-1999 Panel**



To examine whether there is a *survey year effect* for the consistency rates, we selected all the consistency score estimates for the second survey year from each of the four study panels. This allows us to determine whether a particular survey year's data has outperformed the others. Results are shown in the Exhibit 5-5. They offer no firm evidence of a clear survey year effect overall. However, the 1999 survey has the lowest consistency scores among the four survey years for four of six mortgage variables examined. The 100-percent consistency rate for the first-time homebuyer status variable for the 1999 survey data most likely results from the “dependent interviewing” technique used in the data collection process for that variable in that year. This technique is used to avoid repeating questions to the household if the answer should not have changed since the previous survey. The first-time buyer status question was not asked in the 1999 survey if the same household occupied the housing unit and a valid answer had been obtained in a previous survey.

**Exhibit 5-5:
Second Survey Year Consistency of Mortgage Variables of Each AHS Survey**



Finally, to assess whether homeowners with government insured or guaranteed loans (FHA/VA/FmHA loans) tend to remember information about their mortgages differently from conventional loan borrowers, separate consistency scores were calculated for each variable for these two types of borrowers. Estimates for all panels are presented in Exhibit 5-6. Exhibit 5-7 highlights the consistency of the mortgage market segment variable from the 1991-1999 panel in bar-chart format, while a comparison for estimates from the second survey year of the four panels are presented in Exhibit 5-8.

**Exhibit 5-6:
Consistency of Mortgage Variables Over Time, by Mortgage Market Segment
(Percent Reporting Responses Consistent with Base-Year Survey)**

1991-1993-1995-1997-1999 Panel

	AHS Year							
	1993		1995		1997		1999	
	FHA/VA/FmHA	Conventional	FHA/VA/FmHA	Conventional	FHA/VA/FmHA	Conventional	FHA/VA/FmHA	Conventional
Mortgage Market Segment	77%	91%	71%	93%	66%	90%	65%	89%
Origination Amount	94%	83%	83%	74%	84%	70%	100%	59%
Term of Mortgage	97%	90%	95%	92%	91%	87%	95%	87%
Interest Rate (FRMs only)	85%	69%	75%	55%	53%	49%	67%	29%
Payment Plan	94%	79%	92%	77%	93%	78%	95%	78%
Monthly Payment Amount (FRMs only)	91%	82%	79%	65%	74%	58%	79%	61%
First-time Homebuyer Status	91%	89%	93%	84%	95%	90%	100%	83%

1993-1995-1997-1999 Panel

	AHS Year					
	1995		1997		1999	
	FHA/VA/FmHA	Conventional	FHA/VA/FmHA	Conventional	FHA/VA/FmHA	Conventional
Mortgage Market Segment	76%	92%	66%	88%	71%	90%
Origination Amount	85%	72%	86%	76%	81%	70%
Term of Mortgage	97%	87%	96%	88%	93%	84%
Interest Rate (FRMs only)	90%	79%	79%	79%	89%	71%
Payment Plan	92%	86%	94%	85%	97%	82%
Monthly Payment Amount (FRMs only)	90%	75%	88%	73%	80%	74%
First-time Homebuyer Status	87%	89%	90%	87%	86%	88%

1995-1997-1999 Panel

	AHS Year			
	1997		1999	
	FHA/VA/FmHA	Conventional	FHA/VA/FmHA	Conventional
Mortgage Market Segment	70%	90%	65%	91%
Origination Amount	79%	74%	74%	75%
Term of Mortgage	96%	91%	92%	86%
Interest Rate (FRMs only)	89%	83%	91%	80%
Payment Plan	86%	84%	90%	85%
Monthly Payment Amount (FRMs only)	84%	73%	86%	71%
First-time Homebuyer Status	90%	88%	91%	88%

1997-1999 Panel

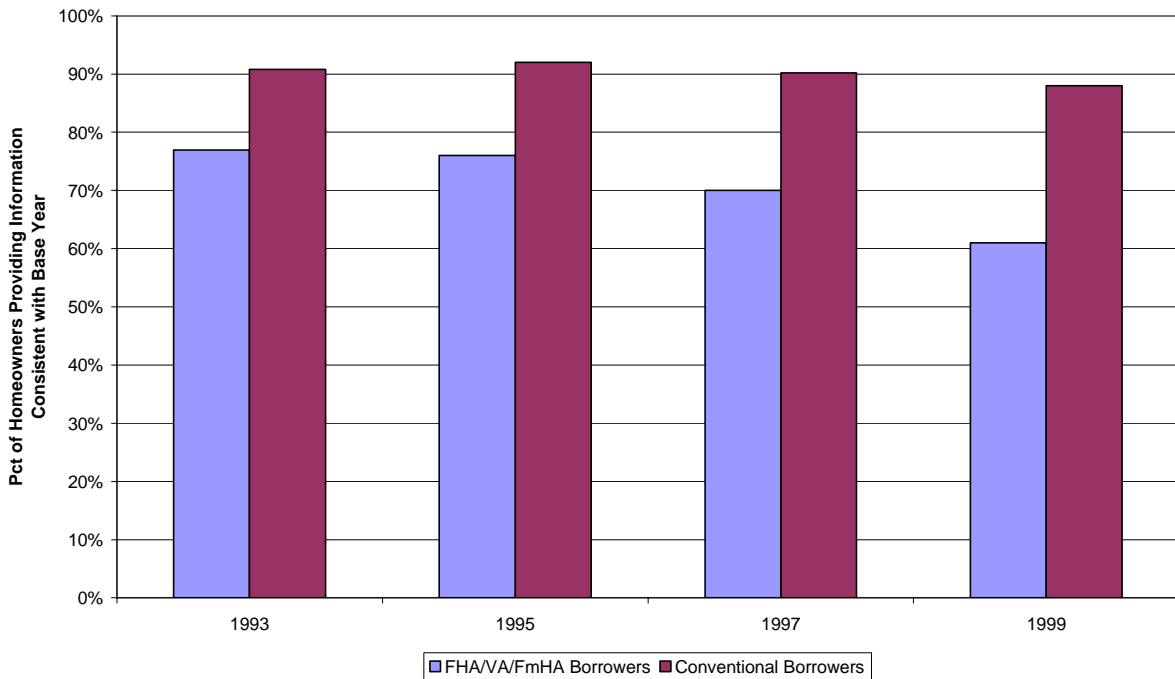
	AHS Year	
	1999	
	FHA/VA/FmHA	Conventional
Mortgage Market Segment	62%	88%
Origination Amount	84%	67%
Term of Mortgage	92%	83%
Interest Rate (FRMs only)	86%	77%
Payment Plan	88%	84%
Monthly Payment Amount (FRMs only)	85%	72%
First-time Homebuyer Status	100%	100%

Source: Authors' calculations of the AHS National Surveys

**Exhibit 5-7:
Consistency of Mortgage Market Segment Information Across AHS Surveys
1991-1993-1995-1997-1999 Panel**



**Exhibit 5-8:
Consistency of Mortgage Market Segment Information Across AHS Surveys
Second Survey Year Consistency**



Conventional wisdom has suggested that many borrowers with FHA/VA/FmHA loans are not aware of the fact that their mortgages are government insured or guaranteed. It is suspected that many homeowners have difficulties distinguishing between the types of mortgage insurance and guarantee provided by FHA, VA and FmHA/RHS. Our analysis provides empirical evidence. As shown in Exhibits 5-7 and 5-8, a larger proportion of respondents who said they had a FHA, VA, or FmHA loans in the base year provided a different answer about their mortgage insurance type in the subsequent survey years than borrowers with conventional mortgages, who provided more consistent information over time. For example, among borrowers who reported that they had a conventional mortgage in the 1991 survey, 91 percent continued to report that they had a conventional mortgage in the 1993 survey. In contrast, among those with a FHA, VA or FmHA mortgage in 1991, only 71 percent of them consistently identified the same insurance type for their mortgages in 1993.

Interestingly, as shown in Exhibit 5-6, borrowers with a government-insured or -guaranteed loan tend to report most of the other characteristics about their mortgages more consistently across surveys when compared to those with a conventional loan.

Summary of Major Findings

To summarize, we found that a large majority of homeowners have reported mortgage information consistently across time in the AHS surveys. In general, the number of responses that are consistent with the base year survey tends to decrease in the later survey years, regardless of the variable examined. Among the variables and survey years we examined, first-time homebuyer status and mortgage term were reported most consistently, while information on interest rate and monthly mortgage payment was least consistent across surveys. The inconsistency of these variables may be due to errors in reporting loans as fixed rate that were actually adjustable rate mortgages. Most homeowners provided consistent answers on questions about their loan's mortgage insurance type and payment plan type.

A potential use of the AHS is to estimate unpaid principal balances (UPBs) and use them to derive contemporaneous loan-to-value (LTV) ratios for the stock of mortgages that are in place. A key question is whether this calculation would be problematic in light of our finding that interest rates reported by homeowners tend to fluctuate across waves of the survey. Recall that UPBs are not collected in the survey and interest rates are one of the variables required in the calculation of UPBs. To see the impact of interest rate discrepancies on the accuracy of the UPB value, let us consider a 30-year fixed-rate loan with an original mortgage amount of \$100,000 and 7 percent interest rate. If the interest rate is off by 1 percentage point after five years, the UPB will be off by 1 percent. The discrepancy in UPB will be around 2 percent if the interest rate is off by 2 percentage points. The differences in UPB are about twice as large after ten years. Therefore, the impact of inaccurate interest rates on the UPB calculation appears to be fairly small.

While our findings have indicated that the proportion of inconsistent responses is not in alarming magnitude for many of the mortgage variables, HUD and the Census Bureau staff should consider the use of "dependent interviewing" technique for the crucial data elements such as mortgage insurance type and payment plan. This technique would be especially appropriate for survey questions with relatively high inconsistency rate, such as interest rate and principal payment amount for fixed-rate

mortgages.

Chapter Six

Conclusions and Recommendations

This chapter summarizes findings and draws lessons from the analyses of the AHS and comparison databases presented in the two previous chapters. It is intended to serve as a guide to the reliability of various mortgage variables and market segments covered in the AHS. The discussion provides guidance to AHS researchers on which types of mortgage analysis they should and should not attempt and places where they should adjust the data or use it with caution. Areas for further possible investigation of the reliability of AHS mortgage variables are discussed, as are modifications that HUD and the Census Bureau might consider making to the AHS survey instrument and interview methods.

Reliable AHS Mortgage Variables

We conclude that the following AHS variables and AHS-based estimates are reliable and can be used for mortgage market analysis, subject to the caveats described in each section:

- **Mortgage Originations for Home-Purchase Loans.** Using the adjustments and timeframes described in this report, the AHS can be used to support analyses of the total volume of recent originations of loans for home purchases.

Since only surviving loans are observed in the AHS, it is necessary to adjust volume estimates upward to account for the number of loans that terminated prior to the interview. We recommend that all AHS volume estimates should be adjusted upward using the termination rates of the market segment to which the mortgage belongs, when that is feasible. Because of the limited data availability for this study of termination rates for other market segments, the analysis in the report used FHA historical termination rates to adjust the volume of loans in all segments of the market.⁷⁰

The AHS is best used to estimate volumes of recent originations, but not of originations that occur in the same year as the survey. Because of the timing of AHS interviews, spanning the last quarter of the year, the AHS will miss a large number of loans originated during the survey year. Therefore, information collected in the AHS should be used to examine the volume and loan characteristics of mortgages originated in the two years prior to the

⁷⁰ Ideally, the loan volume estimates from the AHS should be adjusted using historical termination rates from the corresponding market segments and mortgage product types, since default and prepayment speeds are likely to be different across market segments. Borrowers with conventional mortgages are more likely to refinance their mortgages than their FHA counterparts, given the same interest rate environment. On the other hand, FHA loans have higher claim rates. We expect that the termination rates for FHA loans can serve as a reasonable proxy for the other market segments because, on average, the low FHA prepayment rates should be offset by the high claim rates of FHA loans.

interview year. For example, the 1997 survey should be used to derive volume estimates for loans originated in 1995 and 1996.

- **Volume and Share by Market Segment.** Starting with 1993 originations, researchers can feel comfortable using the AHS to examine mortgage activity by market segment using the following division: FHA, VA/FmHA/RHS, and conventional. Volumes and share estimates for these separate market segments were found to be reliable. The AHS slightly undercounts the share of jumbo loans. It is possible that some borrowers in the conventional segment of the market underestimate their original mortgage amounts. An alternative explanation is that homeowners with jumbo mortgages are underrepresented in the survey as a result of non-participation. Given the fact that participation in the AHS survey commits one to repeated long interviews, more wealthy homeowners may be less willing to participate. It is also likely that, in a declining interest rate environment, those with jumbo loans are particularly quick to refinance their mortgages multiple times, which would not be captured in the AHS.
- **Original Mortgage Amount.** Original loan amounts for mortgages in the FHA segment generally are accurately reported. There is a slight underestimate of the mortgage amounts for borrowers in the VA/FmHA and conventional conforming segments. Users should consider applying an adjustment factor to inflate the loan amounts for such mortgages. In addition, the loan amount variable provided in the public-use file is top-coded for confidentiality reasons. Since the top-coded value has been above the conventional conforming loan limit, the truncation of the loan amount primarily affects analysis of jumbo loans. Therefore, the AHS cannot be used to evaluate jumbo loan amounts.
- **Interest Rate.** Interest rates reported by borrowers in the AHS track the benchmarks for home-purchase fixed-rate mortgages in the FHA, conventional conforming, and conventional jumbo segments of the market. Because of the lack of benchmark information, we have not been able to verify the reliability of interest rates for non-fixed-rate loan types reported in the AHS.
- **First-Time Homebuyer.** The AHS is generally consistent with the RFS and Chicago Title Surveys regarding the share of borrowers who are first-time homeowners. The AHS does consistently report a lower share of first-time buyers than the FHA administrative data. But the cause of this discrepancy is most likely a difference in definition – that is, how this information is defined and solicited in the two data sets. While AHS asks whether any household members have previously owned a home, the FHA mortgage application form inquires whether the buyer has owned a home in the past three years. The AHS definition and estimates are thus more consistent with a strict notion of first-time homebuyers.

The AHS can be used to support research on mortgage activity by market segment as long as all conventional loans are grouped together.

- **Borrower Race/Ethnicity.** Racial and ethnic characteristics reported in the AHS track the benchmark database estimates reasonably well, although there is a tendency to report a slightly higher market share for African Americans and a lower share for Hispanics. These discrepancies most likely result from differences between the AHS and the administrative

data of FHA and HMDA both in definitions of race/ethnicity categories and in the manner in which the race/ethnicity information is gathered.

The principal advantage of using the AHS for mortgage market and housing finance analysis is its detailed household, housing, loan, and geographic characteristics available to the users.⁷¹ In addition to supporting detailed cross tabulations, these variables can also be used as micro data to conduct multiple regressions on the cross sectional data and other loan-level statistical analyses on the longitudinal panels.

One of the findings of this study is the importance of recognizing the sampling variability of estimates obtained from the AHS. We found many cases in which the small sample sizes of particular segments of the mortgage market examined resulted in very wide confidence intervals surrounding individual point estimates. To take sampling variability into account, researchers should report confidence intervals in addition to the AHS point estimates. When the data are used to support loan-level econometric analysis, we recommend that statistical software packages such as STATA[®] and SUDAAN[®] be employed to address fully the sampling errors in such analysis.

Unreliable or Uncertain AHS Mortgage Variables

Users should avoid using the AHS for mortgage market research that would require the following, or use caution in some cases:

- **Refinances.** Refinance originations are not captured adequately in the AHS. Across most of the survey years and all mortgage market segments, the AHS consistently undercounts the number of refinance loans originated. Surprisingly, the loan purpose questions added in the 2001 survey did not provide any noticeable improvement in this underreporting problem. Researchers thus should not use the surveys to produce volume estimates of such mortgages and should not attempt other analysis for which accurate identification of the market share of refinance loans is critical.

However, the distributions of certain loan characteristics such as mortgage market segment, original mortgage amount, and interest rate among those refinances that are captured in the AHS appear to be consistent with the benchmarks for most of the years. Therefore, users may be able to use the data for these loans for analysis pertaining to such loan characteristics. In particular, these records may be used as micro data to conduct regression modeling and other loan-level analyses.

The reason for this underreporting is unknown. One possible explanation is that the questionnaire does not provide a clear guidance to the homeowners regarding the distinction between lump-sum home equity loans and mortgages that are used to refinance an existing

⁷¹ Detailed geographic characteristics of the records are identified in the internal-use version of the AHS file. Abt Associates, through another HUD contract, has geocoded the address information of the 1995 AHS national sample to 1990 census tracts. The work was done under Contract No. C-OPC-05978, Task Order 13: Analysis of the HUD Property Owners & Managers Survey (POMS) and the American Housing Survey (AHS).

loan in order to take out housing equity (so-called cash-out refinances). As a result, some survey respondents might have incorrectly classified their refinance mortgages as home equity loans.

The mediocre performance of the loan purpose question (REFI variable) in the 2001 survey could be because 2001 was such a big refinancing year that the proxy adjustment rates have failed to account for many of the loans that had terminated prior to the interview. In particular, some borrowers might be able to refinance their mortgages multiple times. The REFI variable may perform better in future survey years.

- **Jumbo Loans.** The AHS generally understates the share of mortgages that exceed the conforming loan limit, although it is fairly reliable at identifying conventional loans overall. Possible explanations for this underreporting are offered in the previous section titled “Volume and Share by Market Segment.”
- **Payment Product Types.** Payment product types are not identified accurately in the AHS. In general, ARMs are underreported, and the survey instrument is incapable of identifying borrowers with more sophisticated payment product types such as hybrid ARMs and two-step mortgages. The share of fixed-rate mortgages is generally over estimated, although when measured as a percent of the overall share, the magnitude of the discrepancy is smaller than for ARMs. Thus, the AHS should not be used for analyses of mortgage product choice, and analysts should use caution when including mortgages other than fixed rate mortgages in other types of analyses.
- **Loan-to-Value (LTV) Ratios.** Although average LTV ratios track well the benchmark estimates for both FHA and conventional loans, the distributions of the ratios do not track as well. Of particular note for policy makers, it appears that the AHS estimates of the share of mortgages with very high LTVs may be overstated. This may result from differences in the way house values are recorded in different databases. Discrepancies may also be caused by the extreme values of the loan amount and house value variables that were inaccurately reported in the AHS. Further investigation is needed. Thus, researchers should be cautious in their use of LTV estimates based on the AHS.
- **Borrower Income.** Borrowers’ income in AHS does not track the values reported in HMDA and the FHA data well, mainly because the AHS variable includes different components of income. The AHS measure is more complete and thus yields estimates that are consistently higher. On the other hand, when incomes captured in the Current Population Survey (CPS) and Social Security Administration data are used as benchmarks, studies by the Census Bureau have shown that the AHS respondents have a tendency to underreport incomes. Researchers need to make a judgment independent of the benchmarking to HMDA and FHA data conducted for this study on whether income as identified by the AHS is an appropriate measure of income for their mortgage market research questions.

Potential Use of the AHS for Research on Mortgages

Given the nature of the AHS, there are three broad areas of mortgage-related research that it might be used to address:

- New mortgage originations;
- The use of mortgage finance by homeowners over their tenure in the home; and
- The stock of existing mortgages.

We list below some of the many research topics that have been considered as potential candidates for housing finance research using the AHS and draw lessons for researchers who might be considering using the AHS to investigate those topics.

New Mortgage Originations

Loan originations are of special interest to policy makers and mortgage industry participants because they reflect the current product choices made by borrowers and because they provide insights into the changing underwriting standards of credit suppliers and changes in the supply and demand for mortgage credit over time. Specific topics of interest to researchers and policy audiences may include:

Mortgage market segments for new mortgage originations

- What are the characteristics of borrowers served by the different market segments?
- What are the characteristics of properties served by the different market segments?
- What are the differences in the geographic areas served by the different market segments?
- What are the characteristics of borrowers who use private mortgage insurance (PMI)?
- How do LTV, housing cost-to-income ratios, and interest rates vary across market segments?

The AHS would be a good data source for answering the first three questions for new originations of home purchase loans but not for refinance loans. For home purchase originations, the AHS reliably estimates the volume of loans for each market segment and could be used to study borrower, property, and geographic characteristics of such loans. Thus, for example, the AHS could be used to estimate shares of home purchase loans made to different racial/ethnic and demographic groups. The failure of the AHS to capture refinance loans means that it may not have sufficient sample size to analyze the characteristics of borrowers and properties using refinance loans or the geographical distribution of such loans.

The AHS could also be used to study how interest rates and housing cost-to-income ratios vary across market segments for mortgages originated for home purchase, but not for refinance loans. Because

the distributions of loan-to-value ratios do not track benchmark data well, we recommend caution in using the AHS to study LTVs for home purchase originations.⁷²

Interest rates for new mortgage originations

- How do interest rates vary by borrower characteristics?
- How do interest rates vary by mortgage market segment and other loan characteristics?
- How do interest rates vary with other risk factors such as housing cost-to-income ratios or LTV?
- How do interest rates vary by geographic location?

We conclude that the AHS can be used to study how interest rates vary by borrower characteristics, housing cost-to-income ratios, and geographic location, but only for home purchase loans that have fixed-rates of interest. The accuracy of interest rates for non fixed-rate loans is uncertain because of a lack of reliable benchmark data. When combined with the housing cost-to-income ratios, the AHS interest rate information could potentially be used to identify subprime mortgages (if analysis is limited to fixed-rate originations).

The reliability of LTVs estimated from the AHS is still uncertain.

Cost to income ratios for new mortgage originations

- How has the distribution of housing cost-to-income ratios been changing on newly originated mortgages?
- How does this ratio vary by borrower type?
- How many earners' incomes are supporting mortgage payments (as a measure of the potential resilience of households to the loss of one income)?

The AHS can be used to study these questions, for the mortgage market as a whole and across market segments, but only for home purchase loans and not for refinance loans. Since the AHS does not permit accurate identification of the terms of loans other than fixed rate loans, cost-to-income estimates based on current payment levels for such loans should be interpreted with caution.

Mortgage payment type for new mortgage originations

- What are the characteristics of borrowers using ARMs? How have the characteristics of borrowers using ARMs changed over time?
- How has the ARM share changed over time?

⁷² See below for suggested further analysis that might determine that the AHS can be used reliably to study LTVs.

The AHS should not be used to study mortgage product choice or to study the characteristics of borrowers using ARMs, since it is not possible to identify ARMs and non-fixed rate loans accurately in the AHS.

First-time Homebuyers

- What are the characteristics of first-time homebuyers?
- How have these characteristics changed over time?
- What are the property characteristics of first-time buyers?
- What is the share of first-time buyers who do not finance their house with a mortgage? What are the demographic and property characteristics of these homeowners?
- What are the mortgage characteristics of first-time buyers?

The AHS is well suited to studying the characteristics of first-time homebuyers. However, study of the mortgage characteristics of first-time homebuyers will be limited by the inability of the AHS to identify the terms of mortgages other than fixed rate mortgages.

Mortgage Finance Use over the Life Cycle of an Owner's Tenure in the House

A key feature of the AHS is that it is a panel survey of a sample of housing units. An important consequence of this feature for mortgage-related research is that researchers can monitor the use of mortgage debt over the course of an owner's occupancy of a given unit. This feature also makes AHS a potential source of data for examining dynamic housing finance decisions. In order for such longitudinal analyses to be valid, homeowners residing in the sample housing units must report consistent information about their mortgages across survey years. We found evidence that a large majority of them did so. For most variables tested, between 80 and 90 percent of responses are consistent across survey years. In general, the number of responses that are consistent with the base year survey tends to decrease in the later survey years regardless of the variable examined, reflecting respondents' tendency to misremember mortgage characteristics as time since origination increases. Among the variables and survey years we examined, first-time homebuyer status and mortgage term were reported most consistently, while information on interest rate and monthly mortgage payment was least consistent across surveys. Most homeowners also provided consistent answers on questions about their loan's mortgage insurance type and payment plan type. Information on payment amounts and interest rates differs substantially over time, making these variables less reliable. Undoubtedly, part of this discrepancy is because adjustable rate loans are mistakenly identified as fixed-rate loans.

Potentially, the longitudinal AHS files may be used to study a variety of housing finance questions, including homeowners' decisions to refinance, their use of second mortgages and home equity loans, and the way in which loan-to-value (LTV) ratios affect those decisions. Given our findings regarding the LTV distribution and the survey's inability in identifying refinances, users should exercise caution when using the AHS for such analysis.

Refinancing

- Under what circumstances do borrowers refinance their mortgages?
- What are the homeowners' reasons for refinancing their mortgages? How many are cash-out refinances?
- What are the demographic, property, and loan characteristics of the borrowers who choose not to refinance their homes during a refinance boom (such as 1998)?
- When they refinance, under what circumstances do they change the type of mortgage insurance?
- What is the share of prepayments that are associated with refinancing rather than moves to a different housing unit?

The underidentification of refinance loans may make it problematic to analyze the reasons for refinancing. However, despite the underreporting of refinance mortgages in the survey, many of the loan characteristics appear to remain representative. Therefore, it appears that the AHS may be useful for conducting pooled regression analysis and other longitudinal loan-level modeling pertaining to those variables.

Use of Second Mortgages and Home Equity Loans

- How often are second mortgages or home equity loans used in conjunction with different types of primary mortgages?
- What is the total housing cost-to-income ratio of homes with more than one mortgage?
- What are the characteristics of borrowers and properties with more than one mortgage?

The inability of the AHS survey instrument in distinguishing between home equity loans and refinance mortgages may limit its use for such analysis.

LTV Post Origination

- How much home equity is tapped with the refinancing?
- How much of a home's equity is tapped through second mortgages or home equity loans?

Again, the weakness of the AHS in correctly identifying refinance mortgages limits its use for such analysis. There is evidence that interest rates and loan amounts are reported inconsistently across waves of the survey. Nonetheless, the magnitude of this instability appears to be small. Given the slow amortization of mortgages, the movement of contemporaneous LTV values across time is likely to be dominated by the change in house prices. Other research has demonstrated that the homeowner's estimate of house price is by and large reliable. Therefore, the AHS could be used to derive contemporaneous LTVs for non-refinance mortgages.

The Stock of Existing Mortgages

The AHS is one of the few sources of comprehensive information on the stock of mortgages, including what share of the housing stock has no mortgage. Other sources include the RFS, which is only conducted every ten years, and the SCF, which has smaller sample sizes and less geographic detail. Compared to these other data sources, the AHS is available more frequently and has a much larger sample size. Once again, the detailed borrower and property characteristics and representative national coverage are extremely valuable for supporting research on the nation's stock of mortgages.

The stock of mortgages at a given point in time is of interest because it reflects the potential market for refinancing and home equity mortgages. Furthermore, analysis of the stock of existing mortgages can shed light on the vulnerability of outstanding mortgages and the households that have these financial obligations.

Since the key variables of interest related to the mortgage stock have already been touched upon in the discussion of issues related to new originations and to the use of mortgage financing by an owner over the course of their occupancy of a home, they will not be repeated here.

Below we highlight some of the research questions related to the characteristics of the stock of mortgages at a given point in time:

Cost burdens for homeowners

- What is the distribution of housing cost-to-income ratios for the stock of outstanding mortgages?
- How does this ratio vary by borrower type?
- How many earners' incomes are supporting mortgage payments (as a measure of the potential resilience of households to the loss of one income)?

By design, the AHS appears to be particularly well suited for studying the stock of outstanding mortgages. For analyses of housing cost burdens, the fact that the AHS underestimates refinance originations should lead analysts to use caution (and perhaps to consider making upward adjustments to the estimates of the stock of such loans), but probably not to exclude those loans from the analysis. The weaknesses of the AHS in identifying the terms of mortgages other than fixed-rate mortgages mean that researchers should interpret with caution the implications of current mortgage payments reported for those mortgages.

Vulnerability of households with mortgage debt

- What are the characteristics of borrowers using ARMs? How have the characteristics of borrowers using ARMs changed over time?
- What are the characteristics of borrowers and properties with little or no debt who may be candidates for reverse mortgages or prone to efforts of predatory lenders?

- What are the characteristics of households with ARMs who are exposed to greater interest rate risk?
- How many households are holding mortgages with above market interest rates?
- How many households have negative equity in their properties and what are their characteristics?

Analysis that depends on the accurate identification of ARMs or other mortgage products that are not fixed rate mortgages probably should not be undertaken. On the other hand, the AHS could be used to study the characteristics of borrowers with little or no debt, or with above market interest rates (if the analysis is limited to borrowers who took out fixed-rate loans).

Areas Requiring Further Investigation

- The substantial underreporting of refinance mortgages in the AHS deserves further examination. Given the longitudinal nature of the AHS survey design, linked AHS files might be used to explore whether refinances can be identified more accurately from AHS panel data. Looking at refinances from a panel setting might also shed new light on the structural cause of underreporting of such mortgages and suggest modifications that could be made to the survey. In particular, responses linked across surveys could potentially be used to identify cases where a previous mortgage was retired and a new home equity loan was originated, which could be counted as a refinance.
- For longitudinal analysis, additional research is needed to examine whether there are systematic ways to delete problematic cases from the AHS files to produce a core of internally consistent records with mortgage variables that are reliable. Would the remaining sample be too small to conduct any meaningful analysis? Would deletion of such cases distort the composition of the remaining sample relative to the benchmarks?
- LTVs are another area where further study is needed before determining that the AHS is incapable of supporting LTV research. We suspect that some of the discrepancies in the LTV estimates between the AHS and other mortgage market data sets are caused by outliers in loan amounts and house values reported to the AHS. The distribution of loans by LTV category might be usable if problematic outlier cases were removed by the analyst.
- The micro data for the new round (2001) of RFS will be available to researchers in the summer of 2003. Extensive benchmarking of the 2001 AHS to the new RFS data should be undertaken to gain insight into the impact of post-1995 AHS survey changes on the accuracy of the AHS data. The RFS can be used to test mortgage market measures of recent cohorts as well as the entire stock of mortgages that were in place in 2001. It will be especially valuable if the internal version of the RFS file can be used. The RFS is particularly valuable as a point of comparison since information on mortgages is collected directly from lenders. These data elements should be quite accurate, since the lenders' information is likely supported by administrative records (rather than the homeowners' memory as in data based only on household surveys).

Recommendations for Modifications to the Survey

Based on the research conducted for this study, we have identified several potential modifications to the AHS that would make it substantially more useful for research on mortgage market issues:

- Questions should be added to identify the number of refinances that have taken place during the interval since the last survey. The home equity questions also should be changed to ask about lines of credit rather than “home equity loans,” since there is no real distinction between lump-sum home equity loans and standard mortgages.
- In the current AHS questionnaire, for homeowners who obtained the mortgage the same year the home was purchased, the survey does not obtain the year of loan origination from the respondents; the origination year information can only be retrieved from the WHNGET variable in the “purchase” module of the AHS survey, which reports the year when the house was obtained. For other homeowners, the survey asks explicitly what year the owner obtained the mortgage and the information is stored in the YRMOR variable. HUD and the Census Bureau should consider asking *all* homeowners the month and year of mortgage origination explicitly and storing the information in one variable.
- HUD and the Census Bureau should evaluate changes to specific questions that might elicit more accurate reporting of mortgage payment product types. One of the issues to be considered is the cost-effectiveness of making such changes, given the rapid evolution of mortgage products.
- New questions should also be considered to collect information on mortgage payment status (delinquency and default) and the termination of mortgages that were previously in place.
- HUD and the Census Bureau should consider asking an explicit question on the unpaid principal balance (UPB). UPB is useful for estimating the current home equity and LTV of each homeowner. Currently, UPB is not collected in the AHS. Users can calculate an estimate using information pertaining to interest rate, original amount and mortgage term. Given the lack of details on adjustment terms for non-fixed rate loans and the instability of the interest rates and mortgage amounts reported across waves of the survey, this method is not likely to yield reliable UPB estimates. Although some borrowers may not be able to report the UPB accurately if asked in the survey, the self-reported estimates can nonetheless provide an internal validity check against the user-calculated amounts.
- Given our findings on the instability of several key variables over time, we recommend that the “dependent interviewing” technique be extended to include mortgage insurance type, payment plan, interest rate (of fixed-rate loans), and principal payment amount. This technique is used to avoid repeating questions to the household if the answer should not have changed since the previous survey. For example, the first-time buyer status question was not asked in the 1999 survey if the same household occupied the housing unit and a valid answer had been obtained in a previous survey.

Appendix A

GSE Conforming Loan Limits, 1980-2000

Origination Year	Number of Units in Structure			
	One-Unit	Two-Unit	Three-Unit	Four-Unit
1980	\$93,751	\$120,000	\$145,000	\$180,000
1981	\$98,500	\$126,000	\$152,000	\$189,000
1982	\$107,000	\$136,800	\$165,100	\$205,300
1983	\$108,300	\$138,500	\$167,200	\$207,900
1984	\$114,000	\$145,800	\$176,100	\$218,900
1985	\$115,300	\$147,500	\$178,200	\$221,500
1986	\$133,250	\$170,450	\$205,950	\$256,000
1987	\$153,100	\$195,850	\$236,650	\$294,150
1988	\$168,700	\$215,800	\$260,800	\$324,150
1989	\$187,600	\$239,950	\$290,000	\$360,450
1990	\$187,450	\$239,750	\$289,750	\$360,150
1991	\$191,250	\$244,650	\$295,650	\$367,500
1992	\$202,300	\$258,800	\$312,800	\$388,800
1993	\$203,150	\$259,850	\$314,100	\$390,400
1994	\$203,150	\$259,850	\$314,100	\$390,400
1995	\$203,150	\$259,850	\$314,100	\$390,400
1996	\$207,000	\$264,750	\$320,050	\$397,800
1997	\$214,600	\$274,550	\$331,850	\$412,450
1998	\$227,150	\$290,650	\$351,300	\$436,600
1999	\$240,000	\$307,100	\$371,200	\$461,350
2000	\$252,700	\$323,400	\$390,900	\$485,800

Notes: Loan limits for mortgages made in Alaska, Hawaii and Guam (and the U.S. Virgin Islands as of 1/1/93) are 50% higher than the ones for the rest of the country. Loan limits for second mortgages are half of the 1-unit limits.

Appendix B

Summary Information of Comparison Databases

Exhibit B-1 HUD FHA Administrative Data

Strengths	<ul style="list-style-type: none"> • Universe of FHA-insured mortgages • Loan status is tracked monthly from origination to maturity or termination • Loan-level records cover a rich set of mortgage characteristics
Weaknesses	<ul style="list-style-type: none"> • Only FHA-insured mortgages are covered • Borrower characteristics are only available for loans originated in 1992/1993 or later (F42 records)
Source	Administrative databases HUD FHA uses to process mortgage insurance applications and underwriting, and to keep track of loan performance over the life of the loans. The Single-Family Data Warehouse System (SFDWS) has incorporated monthly loan records from systems such as A43, F42, and CHUMS.
Frequency	Since mid-1970. Loan status is updated monthly.
Geographic Coverage	National
Relevant Finance Variables	<ul style="list-style-type: none"> • Comprehensive coverage on basic mortgage characteristics including: <ul style="list-style-type: none"> • Year of origination • Loan purpose • Original mortgage amount • Loan term • Contract interest rate • Payment plan • Mortgage insurance premium • Upfront premium and fees • Appraised house value • Sales price • Loan-to-value (LTV) ratio at origination • Property type
Market Segments	<ul style="list-style-type: none"> • FHA-insured mortgages
Demographic Variables	<ul style="list-style-type: none"> • Age • Gender • Race/ethnicity • Marital status • Household income • First-time home owner
Neighborhood Variables	Most loan records have been geo-coded. The availability of Census tract identification numbers allows the possibility of linking Census tract-level neighborhood variables to each loan records.

Exhibit B-2
Home Mortgage Disclosure Act (HMDA) Data

Strengths	<ul style="list-style-type: none"> • Availability of Census tract identifier • Micro-level data on mortgage applications and originations including information on the race, income, age and gender of borrower • Data available annually since 1990
Weaknesses	<ul style="list-style-type: none"> • Very limited information on mortgage terms and costs • Poor coverage in rural areas • Poor coverage of second mortgages
Source	<p>Mortgage lenders meeting the following criteria:</p> <ul style="list-style-type: none"> • Depository institutions with at least \$29 million in assets, a branch or home office in an MSA, and is either: federally insured/regulated, issues loans that are either insured, guaranteed or supplemented by a federal agency, or sells loans to Fannie Mae or Freddie Mac. • Non-depository institutions that are for-profit, have home purchase or refinance originations totaling 10% or more of its total loan originations in dollars, have a branch office in an MSA or review 5 loan applications from any single MSA, have at least \$10 million in assets or originate 100 or more home purchase or refinance loans.
Frequency	Annual
Geographic Coverage	National but there are known limitations in coverage. It is estimated that HMDA covers about 75-80% of all home purchase loans (including refinancings). In general, there appears to be a poor coverage in non-metro areas and better coverage in metro areas.
Relevant Finance Variables	<p>Number and dollar amount of mortgages originated. Loan stratifications include:</p> <ul style="list-style-type: none"> • Loan Purpose: Purchase, Refinance, Home Improvement, and Multifamily. Home Improvement and Multifamily are believed to have very poor coverage of the markets because of exclusions related to reporting of home equity loans and the organizations that are required to report, respectively. • Loan Type: Conventional, FHA, VA, and FmHA • Owner Occupancy: Owner-occupied, Not Owner Occupied, Not Applicable (for multifamily loans or for loans outside metro areas).
Market Segments	Residential first mortgages, with limited coverage of home equity loans
Demographic Variables	<ul style="list-style-type: none"> • Income • Race of borrower and co-borrower • Gender of borrower and co-borrower
Neighborhood Variables	None provided directly, but Census tract number is reported for loans made in MSAs and county code is reported for virtually all loans. This allows researchers to link any data that are available at the Census tract or county level.

Exhibit B-3

**Federal Housing Finance Board's (FHFB) Monthly Interest Rate Survey (MIRS)
(Monthly Survey of Rates and Terms on Conventional 1-Family Nonfarm Mortgage Loans)**

Strengths	<ul style="list-style-type: none"> • Long time-series • Both monthly and annual information is available • Both aggregate estimates and micro-level loan data are available
Weaknesses	<ul style="list-style-type: none"> • FHA, VA, multifamily, mobile home loans, balloon and other product types, second mortgage, and refinance loans are not covered • No demographic information on borrowers • Studies have questioned the internal consistency and reliability of the data, particularly with respect to classification of loans as adjustable- or fixed-rate.
Source	Monthly survey of loan-level lending activities of a sample of approximately 300 mortgage lenders, including savings associations, mortgage companies, commercial banks, and savings banks. It covers fully amortized conventional loans for the purchase of 1-unit single-family, nonfarm homes.
Frequency	Monthly survey
Geographic Coverage	<ul style="list-style-type: none"> • National and state estimates are reported annually • Major metropolitan areas, and Federal Home Loan Bank (FHLB) district estimates are reported quarterly or annually
Relevant Finance Variables	<p>Aggregate data include information such as mortgage interest rates, loan terms, and house prices by property type (all, new, and previously occupied), by loan type (fixed- or adjustable-rate), and by lender type (savings associations, mortgage companies, commercial banks, and savings banks).</p> <p>Loan-level data with sample weights can be purchased at a cost of approximately \$100 per year of data. Relevant variables include:</p> <ul style="list-style-type: none"> • Fixed or adjustable rate loan • Effective interest rate (contract rate plus fees and charges amortized over a 10-year period) • Initial fees and charges • Lender type (OTS-regulated thrift, mortgage company, commercial bank, or FDIC regulated thrift) • Loan-to-value (LTV) ratio • Purchase house price • Principal mortgage amount • Loan purpose (existing home, new home, combined construction) • Mortgage term • Month and year of origination
Market Segments	Fully amortized conventional mortgages for the purchase of single-family, non-farm houses
Demographic Variables	None
Neighborhood Variables	None

Exhibit B-4
Residential Finance Survey (RFS)

Strengths	<ul style="list-style-type: none"> • Micro-level data with very large sample size • Comprehensive coverage of housing finance and mortgage-related variables
Weaknesses	<ul style="list-style-type: none"> • Available every ten years. Most recent 1991. • Relatively old (survey conducted about 10 years ago) • Limited geographic identifiers
Source	Survey on a nationally representative sample of approximately 70,000 properties providing detailed information of the financing of homeowners and rental properties, including characteristics of the mortgage, properties, and property owners. Separate questionnaires were sent to the property owners and lenders; both were asked to provide information about the mortgages and financing. Different sets of survey questions were designed for single-family homeowners and rental and vacant property owners.
Frequency	Available every ten years. The latest available is 1991.
Geographic Coverage	National, four Census regions, inside/outside MSAs and central cities, and four highly-populated states (California, Florida, New York, and Texas). The metropolitan area boundaries were based on the 1984 OMB definition.
Relevant Finance Variables	Very comprehensive. The public-use micro data cover almost the same set of mortgage and housing finance data items as the AHS.
Market Segments	Mortgages secured by non-farm and privately-owned residential properties
Demographic Variables	<ul style="list-style-type: none"> • Age • Gender • Race/ethnicity • Veteran status • Household income from various sources
Neighborhood Variables	None

Exhibit B-5
Federal Reserve Board's Survey of Consumer Finance (SCF)

Strengths	Micro-level data contain comprehensive information on mortgage, finance and household demographic characteristics.
Weaknesses	<ul style="list-style-type: none"> • Small sample size • Crude geographic identifiers • Data only available every three years; survey years may not correspond to the AHS
Source	Survey of a nationally representative sample of about 4,500 households. Detailed information is collected on the household assets, liabilities, mortgages and other relevant financial information. Information related to the financial balance sheets, pension plans, income, use of financial institutions, and selected demographic characteristics are also available.
Frequency	Triennial. The most recent data available are for 1998.
Geographic Coverage	National, four Census regions, and nine Census divisions
Relevant Finance Variables	Very comprehensive. The public-use micro data cover almost the same set of mortgage and housing finance data items as the AHS.
Market Segments	Mortgages secured by privately-owned residential properties
Demographic Variables	<ul style="list-style-type: none"> • Race/ethnicity • Gender • Age • Marital status • Household income • Education level
Neighborhood Variables	<ul style="list-style-type: none"> • Type of neighborhood (based on interviewer's impressions): mostly residential, mostly non-residential • Characterization of neighborhood residents (based on interviewer's impressions): all black, all non-black

Exhibit B-6**Chicago Title and Trust Co. Annual Survey of Recent Home Buyers**

Strengths	<ul style="list-style-type: none">• Tabulation tables cover a comprehensive set of mortgage as well as household characteristics• Available annually
Weaknesses	<ul style="list-style-type: none">• Only home sale activities in 20 major metropolitan housing markets were covered• Metropolitan area boundary definitions may not be compatible with the ones used in the AHS• Only tabulations in published reports are available for public use• Micro-data are proprietary and expensive to acquire
Source	Survey of a nationally representative sample of approximately 1,800 home buyers in 20 major metropolitan housing markets across the nation. Summary level data (in cross-tabulation format) are published annually in the <i>Who's Buying Homes in America</i> report and can be downloaded from the company's Internet site.
Frequency	Annual. The 24 th survey covers home sales in 1999.
Geographic Coverage	20 major metropolitan housing markets
Relevant Finance Variables	Tabulations contain a wealth of information on the mortgages, properties and home buyers' demographic and economic characteristics.
Market Segments	Home-purchase loans in major metropolitan markets
Demographic Variables	<ul style="list-style-type: none">• Age• Gender• Race/ethnicity• Marital status• Family size• Household Income• First-time home buyer
Neighborhood Variables	None

Exhibit B-7**National Association of Realtors' (NAR) Existing Home Sales and Metropolitan Median Prices Data**

Strengths	<ul style="list-style-type: none">• Long time-series• Data available monthly, quarterly and annually
Weaknesses	<ul style="list-style-type: none">• Aggregate data for large geography• New homes, mobile homes, and properties with more than one unit are not covered
Source	Compiled by the Economic Research Group of the National Association of Realtors, source data come from a national sample of Multiple Listing Service sales.
Frequency	Monthly, quarterly, and annually
Geographic Coverage	National, state, and MSA
Relevant Finance Variables	<ul style="list-style-type: none">• Volume of existing single-family home sales• Average and median sales price
Market Segments	Single-family owner-occupied existing home sales
Demographic Variables	None
Neighborhood Variables	None

Appendix C

Alternative Approaches to Adjusting the Mortgage Origination Volume Estimates in AHS

A challenge for conducting the replication analysis was to adjust the AHS estimates to account for mortgage activity between survey waves that is not captured by the AHS. Records in administrative data sets such as the FHA and HMDA files are organized by year of mortgage origination. Information collected in the AHS, in contrast, represents a snapshot of the mortgage stock as of the time of the interview. This means that mortgages observed in the survey are surviving loans – loans still outstanding at the time of the interview. As a result, it is not possible for the AHS to capture completely mortgage activity that happened between waves of the survey. This appendix identifies alternative methods to account for mortgage activity between survey waves and discusses the pros and cons of these methods.

Since the AHS is generally conducted between July and December of the survey year, mortgage activity in the second half of the survey year is not captured completely. As a result, using the AHS to estimate origination volume for the calendar year of the survey will understate the actual volume. Our tabulations of the 1991, 1993, 1995, 1997 and 1999 data have indicated that the vast majority of the interview activity was carried out between August and November, on average a quarter of the year. Hence one way to address this problem is to inflate the volume estimate for that calendar year by a factor of $\frac{1}{4}$. We call this the *partial year adjustment* method. It depends on the assumption that mortgage activity in the second half of the calendar year does not differ from that in the first half of the year, and this is not likely the case because of the seasonality of home purchase activity. In addition, a small portion of the loans originated in the first half of that year would have already been terminated by refinancing or household relocation by the time the interview happened. The $\frac{1}{4}$ inflation factor does not adjust for terminated loans. Given the challenge of adjusting current year's activity, we do not recommend using the AHS to look at mortgage originations in the same year the interview is conducted.

Instead, this study recommends that information collected in the AHS be used to estimate the volume and characteristics of mortgages originated in the two years prior to the interview year. For example, we use the 1997 survey data to study the characteristics and volume of loans originated in 1995 and in 1996. Mortgage activity that occurred in 1997 is studied using data collected in the 1999 survey. Exhibit C-1 illustrates the correspondence between the AHS survey years and the calendar years of mortgage activity examined in the study.

To account for the loans that have terminated prior to the interview, all volume estimates are adjusted upward using historical termination (prepayment plus claim) rates derived from the FHA administrative data. The adjustment factors are based on the cumulative prepayment and claim rates for different loan cohorts published in the *Annual Actuarial Review of the FHA's Mutual Mortgage Insurance Fund (Fiscal Year 2000)*. This report contains historical termination rates tabulated by origination cohort and policy year. Ideally, one would want to adjust the loan volume estimate from the AHS using historical termination rates of loans for each market segment and mortgage product

type, since different loan products are likely to have different prepayment speeds. For example, borrowers with conventional mortgages are more likely to refinance their mortgages than their FHA counterparts, given the same interest rate environment. On the other hand, FHA loans have higher claim rates. Historical prepayment and claim information for loans in the non-FHA segments by origination year and policy year is not readily available. The rates for FHA loans can serve as reasonable proxies as we expect that on average the low FHA prepayment rates should be offset by the high claim rates of FHA loans. It is possible that these adjustments could result in underestimated origination totals for loans in the non-FHA segments of the market, especially for surveys conducted during refinancing boom (such as the 2001 survey). Further research is needed in this regard. An alternative approach is to use prepayment (termination) rates of securities issued by the GSEs (Ginnie Mae, Fannie Mae, and Freddie Mac). These data could be purchased from a number of Wall Street firms. But this method may be problematic because GSE securities represent pools of mortgages and their prepayment and claim speeds could potentially be different from their loan-level counterparts.

Thus, all AHS volume estimates reported in Chapter Four, regardless of market segments, have been adjusted upward using the FHA historical termination rates of the appropriate origination year and policy year.⁷³ Exhibit C-2 presents the adjustment rate for each origination.

**Exhibit C-1:
Correspondence between AHS Survey Years and Calendar Years of Mortgage Activity Examined**

AHS SURVEY YEAR	CALENDAR YEARS OF MORTGAGE ACTIVITY EXAMINED
1991	1989 1990
1993	1991 1992
1995	1993 1994
1997	1995 1996
1999	1997 1998
2001	1999 2000

⁷³ These are derived from the sum of cumulative prepayment and claim rates of appropriate origination year and policy year, published in a series of tables in Appendix H of *Annual Actuarial Review of the FHA's Mutual Mortgage Insurance Fund (Fiscal Year 2000 and 2001)* by Deloitte & Touche. In these tables, "policy year" represents the number of years since loan origination. As an example, when we estimate the 1993 loan volume using the 1995 AHS, the raw (unadjusted) estimate represents the volume of surviving loans of that origination cohort in their 3rd policy year. The table in the *Actuarial Review* indicates that, according to the experience of FHA loans, 13.6 percent of the loans have already terminated – 12.9 percent because of prepayment/refinancing, and 0.7 percent because of claim/foreclosure. Therefore, we adjust the 1995 AHS raw volume estimate for 1993 mortgages by 13.6 percent.

**Exhibit C-2:
Adjustment Rate for AHS Volume Estimates by Origination Year**

AHS SURVEY YEAR	ORIGINATION YEAR	ADJUSTMENT RATE
1991	1989	8.2%
1991	1990	3.0%
1993	1991	28.8%
1993	1992	8.8%
1995	1993	13.6%
1995	1994	4.5%
1997	1995	21.2%
1997	1996	5.4%
1999	1997	37.3%
1999	1998	13.3%
2001	1999	17.8%
2001	2000	18.6%

Source: Authors' calculations based on historical prepayment and claim rates of FHA loans.

An alternative approach to adjusting the AHS volume estimates was suggested by HUD staff. It seeks to recover the number of home-purchase originations that have terminated prior to the interview by using information on house purchase years and recent refinances. Specifically, a number of homeowners would indicate to the AHS that they bought their homes within the two-year period prior to the interview. They might also report that their current mortgage was taken out after the house was purchased. These loans would logically be treated as refinances rather than home purchase originations. But we could also infer that these homeowners must have taken out a first mortgage when they purchased their house and that the loan was subsequently terminated or refinanced. Hence, these “imputed” first mortgages could be added to the count of home-purchase originations for the house purchase year.

This adjustment method, however, depends critically on the assumption that recent refinancing activity prior to the interview has been adequately captured in the AHS. As the analysis in Chapter Four shows, refinance loans are not captured well in the AHS, especially since the 1995 survey. Furthermore, this method would not be able to capture any refinancing activity that happened within the same year of the house purchase. Therefore, we find this adjustment method inferior to the approach that makes use of the historical FHA termination rate information.

Appendix D
Tabulations Supporting the Analysis in Chapter
Four

Exhibit D-3-7: Number of FHA-insured Mortgages, by Origination Year and by Loan Purpose (AHS vs. FHA)

Home Purchases

AHS Survey	Origination Year	AHS				FHA	AHS Coverage
		Unadjusted	Adjusted	95% Conf. Interval			
1991	1989	546,723	594,090	518,316	669,864	675,695	88%
1991	1990	597,725	617,843	542,204	693,481	702,307	88%
1993	1991	479,857	621,254	533,074	709,434	603,718	103%
1993	1992	521,862	570,403	493,692	647,114	563,941	101%
1995	1993	502,817	577,167	497,833	656,502	616,111	94%
1995	1994	499,967	528,069	454,833	601,305	583,789	90%
1997	1995	425,169	519,300	437,247	601,352	562,571	92%
1997	1996	612,549	651,213	564,819	737,606	677,831	96%
1999	1997	625,355	865,593	754,728	976,458	745,041	116%
1999	1998	567,461	647,576	559,100	736,052	775,038	84%
2001	1999	751,160	893,314	791,177	995,450	900,609	99%
2001	2000	623,345	743,694	638,632	848,756	806,110	92%

Refinances

AHS Survey	Origination Year	AHS				FHA	AHS Coverage
		Unadjusted	Adjusted	95% Conf. Interval			
1991	1989	30,265	32,757	15,483	50,031	47,709	69%
1991	1990	24,616	25,360	10,154	40,566	44,158	57%
1993	1991	45,895	59,173	33,224	85,122	66,070	90%
1993	1992	190,768	207,844	162,386	253,302	215,567	96%
1995	1993	261,137	297,103	241,543	352,663	627,062	47%
1995	1994	162,343	169,920	128,774	211,065	314,033	54%
1997	1995	20,918	25,363	8,704	42,022	59,641	43%
1997	1996	21,090	22,225	6,669	37,782	121,797	18%
1999	1997	91,954	126,300	84,628	167,972	108,728	116%
1999	1998	199,777	226,379	178,194	274,563	347,369	65%
2001	1999	159,795	188,247	140,093	236,402	243,716	77%
2001	2000	67,020	79,468	49,185	109,750	62,916	126%

**Exhibit D-5-9: Number of Mortgages, by Origination Year and by Loan Purpose (AHS vs. HMDA)
Loans Originated in Metro Areas Only**

Home Purchases

Origination Year	Market Segment	AHS			HMDA	AHS Coverage	
		Unadjusted	Adjusted	95% Conf. Interval			
1993	FHA	456,804	524,351	450,509	598,193	496,403	106%
1994		456,586	482,250	413,333	551,166	456,154	106%
1995		370,561	452,602	377,800	527,403	472,019	96%
1996		555,814	590,897	514,553	667,240	583,970	101%
1997		555,610	769,055	665,116	872,993	629,653	122%
1998		500,444	571,097	487,738	654,456	663,540	86%
1999		683,042	812,305	714,565	910,044	754,482	108%
2000		544,689	649,852	561,002	738,701	714,660	91%
1993	VA/FmHA/RHS	190,247	217,143	167,968	266,319	195,081	111%
1994		228,169	239,770	187,677	291,863	185,196	129%
1995		163,742	199,157	151,494	246,820	180,005	111%
1996		169,978	179,685	135,952	223,417	189,935	95%
1997		142,564	196,233	146,236	246,230	176,071	111%
1998		196,411	223,260	174,136	272,384	199,434	112%
1999		175,561	207,403	158,741	256,065	178,974	116%
2000		151,929	180,478	133,647	227,310	149,957	120%
1993	Conforming	1,611,620	1,914,291	1,769,608	2,058,974	1,601,621	120%
1994		1,808,067	1,989,002	1,851,778	2,126,226	1,838,419	108%
1995		1,613,177	2,035,078	1,896,863	2,173,293	1,771,977	115%
1996		1,778,559	1,938,676	1,803,211	2,074,141	2,008,489	97%
1997		1,602,284	2,262,548	2,087,267	2,437,829	1,991,377	114%
1998		1,882,191	2,199,661	2,037,152	2,362,171	2,471,127	89%
1999		1,711,260	2,075,601	1,919,739	2,231,464	2,616,037	79%
2000		1,767,749	2,147,672	1,980,115	2,315,228	2,601,106	83%
1993	Jumbo	105,726	120,346	85,446	155,247	154,672	78%
1994		142,019	148,861	111,737	185,985	202,714	73%
1995		107,075	130,029	90,396	169,662	173,964	75%
1996		169,193	178,748	134,837	222,659	204,781	87%
1997		111,559	153,438	112,001	194,875	265,673	58%
1998		187,820	213,310	167,039	259,581	275,392	77%
1999		182,422	215,388	164,885	265,892	297,886	72%
2000		156,457	185,787	139,051	232,524	349,459	53%

Refinances

Origination Year	Market Segment	AHS				HMDA	AHS Coverage
		Unadjusted	Adjusted	95% Conf. Interval			
1993	FHA	240,735	273,891	221,279	326,503	474,782	58%
1994		155,136	162,376	122,361	202,392	163,774	99%
1995		13,891	16,843	3,303	30,383	52,960	32%
1996		18,054	19,026	4,799	33,253	94,352	20%
1997		80,731	110,885	74,157	147,613	93,641	118%
1998		189,328	214,538	167,551	261,525	303,622	71%
1999		155,808	183,550	135,871	231,230	163,729	112%
2000		64,698	76,714	46,921	106,508	52,680	146%
1993	VA/FmHA/RHS	104,678	118,963	83,913	154,014	204,789	58%
1994		48,483	50,694	28,707	72,680	90,655	56%
1995		5,264	6,380	(2,470)	15,229	32,229	20%
1996		2,312	2,436	(2,342)	7,215	51,112	5%
1997		34,701	47,659	23,707	71,610	49,639	96%
1998		88,803	100,617	62,906	138,327	186,013	54%
1999		53,671	63,222	36,521	89,924	65,775	96%
2000		20,642	24,473	7,374	41,572	6,243	392%
1993	Conforming	2,154,904	2,493,776	2,354,688	2,632,864	3,901,853	64%
1994		1,206,801	1,280,148	1,176,613	1,383,683	1,498,767	85%
1995		186,321	227,406	176,111	278,701	1,013,623	22%
1996		229,641	242,005	193,668	290,341	1,690,654	14%
1997		847,380	1,165,616	1,044,909	1,286,324	1,774,887	66%
1998		1,685,503	1,913,978	1,743,687	2,084,269	4,550,599	42%
1999		1,098,509	1,295,372	1,174,143	1,416,601	2,967,123	44%
2000		548,159	651,217	563,616	738,818	1,664,685	39%
1993	Jumbo	114,767	130,430	94,891	165,969	323,164	40%
1994		60,946	63,728	39,041	88,415	151,738	42%
1995		8,141	9,868	(1,429)	21,165	97,877	10%
1996		22,114	23,305	8,040	38,569	134,098	17%
1997		49,949	68,603	40,198	97,007	189,334	36%
1998		154,897	175,517	133,951	217,083	388,342	45%
1999		70,712	83,297	53,084	113,510	223,128	37%
2000		59,407	70,440	41,982	98,897	121,633	58%

Exhibit D-10: Shares of Loans by Mortgage Market Segment - AHS vs. RFS

All Loans

Origination Year	Market segment	AHS			RFS		
		Estimate	95% Conf. Interval		Estimate	95% Conf. Interval	
1989-91	FHA	17%	15%	18%	20%	19%	22%
	VA/FmHA	6%	5%	7%	7%	6%	8%
	Conforming	73%	72%	75%	69%	67%	70%
	Jumbo	4%	4%	5%	4%	3%	5%

Home Purchases

Origination Year	Market segment	AHS			RFS		
		Estimate	95% Conf. Interval		Estimate	95% Conf. Interval	
1989-91	FHA	20%	19%	22%	26%	24%	28%
	VA/FmHA	7%	6%	8%	9%	8%	10%
	Conforming	68%	66%	70%	61%	59%	63%
	Jumbo	4%	4%	5%	4%	3%	5%

Refinances

Origination Year	Market segment	AHS			RFS		
		Estimate	95% Conf. Interval		Estimate	95% Conf. Interval	
1989-91	FHA	4%	3%	5%	4%	3%	5%
	VA/FmHA	1%	1%	2%	1%	0%	2%
	Conforming	91%	89%	92%	91%	89%	93%
	Jumbo	4%	3%	5%	4%	3%	5%

Exhibit D-11-12: Shares of Loans by Mortgage Market Segment - AHS vs. HMDA
Loans originated in metro areas

Home Purchase

Origination Year	Market segment	AHS			HMDA
		Estimate	95% Conf. Interval		
1993	FHA	19%	17%	22%	20%
	VA/FmHA	8%	6%	10%	8%
	Conforming	68%	65%	71%	65%
	Jumbo	4%	3%	6%	6%
1994	FHA	17%	15%	20%	17%
	VA/FmHA	9%	7%	10%	7%
	Conforming	69%	66%	71%	69%
	Jumbo	5%	4%	7%	8%
1995	FHA	16%	14%	19%	18%
	VA/FmHA	7%	6%	9%	7%
	Conforming	72%	68%	75%	68%
	Jumbo	5%	3%	6%	7%
1996	FHA	21%	18%	23%	20%
	VA/FmHA	6%	5%	8%	6%
	Conforming	67%	64%	70%	67%
	Jumbo	6%	5%	8%	7%
1997	FHA	23%	20%	26%	21%
	VA/FmHA	6%	4%	7%	6%
	Conforming	66%	63%	69%	65%
	Jumbo	5%	3%	6%	9%
1998	FHA	18%	16%	20%	18%
	VA/FmHA	7%	6%	9%	6%
	Conforming	68%	65%	71%	68%
	Jumbo	7%	5%	8%	8%
1999	FHA	25%	22%	27%	20%
	VA/FmHA	6%	5%	8%	5%
	Conforming	62%	59%	65%	68%
	Jumbo	7%	5%	8%	8%
2000	FHA	21%	18%	23%	19%
	VA/FmHA	6%	4%	7%	4%
	Conforming	67%	64%	70%	68%
	Jumbo	6%	5%	7%	9%

Refinances

Origination Year	Market segment	AHS			HMDA
		Estimate	95% Conf. Interval		
1993	FHA	9%	8%	11%	10%
	VA/FmHA	4%	3%	5%	4%
	Conforming	82%	80%	85%	80%
	Jumbo	4%	3%	6%	7%
1994	FHA	11%	8%	13%	9%
	VA/FmHA	3%	2%	5%	5%
	Conforming	82%	79%	85%	79%
	Jumbo	4%	3%	6%	8%
1995	FHA	7%	1%	12%	4%
	VA/FmHA	2%	-1%	6%	3%
	Conforming	87%	80%	94%	85%
	Jumbo	4%	0%	8%	8%
1996	FHA	7%	2%	11%	5%
	VA/FmHA	1%	-1%	3%	3%
	Conforming	84%	78%	91%	86%
	Jumbo	8%	3%	13%	7%
1997	FHA	8%	5%	11%	4%
	VA/FmHA	3%	2%	5%	2%
	Conforming	84%	80%	87%	84%
	Jumbo	5%	3%	7%	9%
1998	FHA	9%	7%	11%	6%
	VA/FmHA	4%	3%	6%	3%
	Conforming	80%	77%	82%	84%
	Jumbo	7%	6%	9%	7%
1999	FHA	11%	9%	14%	5%
	VA/FmHA	4%	2%	6%	2%
	Conforming	80%	76%	83%	87%
	Jumbo	5%	3%	7%	7%
2000	FHA	9%	6%	13%	3%
	VA/FmHA	3%	1%	5%	0%
	Conforming	79%	74%	84%	90%
	Jumbo	9%	5%	12%	7%

Exhibit D-11-12a : Shares of Loans by Mortgage Market Segment - AHS vs. SCF

All Loans

Origination Year	Market segment	AHS			SCF		
		Estimate	95% Conf. Interval		Estimate	95% Conf. Interval	
1993	FHA	13%	11%	14%	18%	15%	20%
	VA/FmHA	6%	5%	7%	6%	5%	8%
	Conforming	78%	76%	79%	73%	70%	75%
	Jumbo	4%	3%	5%	3%	3%	4%
1994	FHA	13%	11%	14%	18%	15%	21%
	VA/Fmha	6%	5%	7%	12%	10%	15%
	Conforming	77%	75%	79%	65%	62%	69%
	Jumbo	4%	3%	5%	4%	4%	5%

Home Purchases

Origination Year	Market Segment	AHS			SCF		
		Estimate	95% Conf. Interval		Estimate	95% Conf. Interval	
1993	FHA	17%	15%	19%	25%	21%	29%
	VA/FmHA	8%	6%	10%	10%	7%	12%
	Conforming	71%	69%	74%	63%	59%	67%
	Jumbo	4%	3%	5%	2%	1%	3%
1994	FHA	15%	13%	17%	21%	17%	25%
	VA/FmHA	8%	6%	9%	15%	11%	18%
	Conforming	73%	71%	76%	60%	55%	65%
	Jumbo	4%	3%	5%	4%	3%	5%

Refinances

Origination Year	Market Segment	AHS			SCF		
		Estimate	95% Conf. Interval		Estimate	95% Conf. Interval	
1993	FHA	9%	7%	10%	13%	10%	15%
	VA/FmHA	4%	3%	5%	5%	3%	6%
	Conforming	84%	82%	86%	79%	76%	82%
	Jumbo	4%	3%	5%	4%	3%	5%
1994	FHA	9%	7%	11%	14%	11%	18%
	VA/FmHA	3%	2%	4%	9%	6%	13%
	Conforming	84%	82%	87%	72%	67%	76%
	Jumbo	4%	2%	5%	5%	4%	6%

**Exhibit D-11-12b:
Shares of Loans by Mortgage Market Segment - AHS vs. SCF
1993 and 1994 Origination Cohorts**

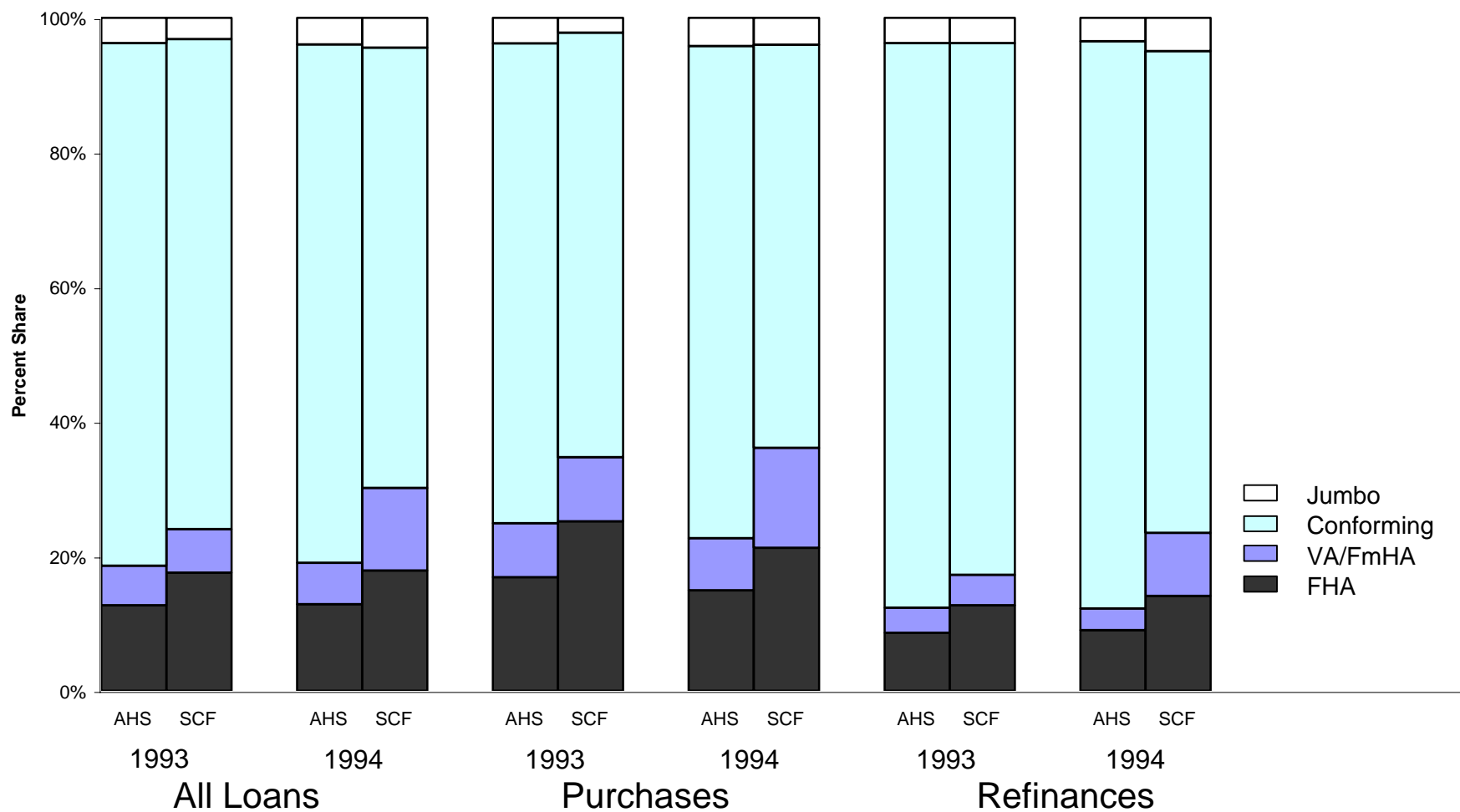


Exhibit D-13-14 : Original Mortgage Amount of FHA-Insured Loans - AHS vs. FHA (In thousands of dollars)

Home Purchases

Origination Year	AHS						FHA			
	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl
1989	\$45	\$66	\$64	\$83	\$62	\$71	\$47	\$64	\$62	\$80
1990	\$49	\$69	\$67	\$85	\$64	\$73	\$48	\$67	\$64	\$84
1991	\$49	\$73	\$68	\$87	\$68	\$79	\$50	\$70	\$66	\$86
1992	\$50	\$76	\$70	\$91	\$70	\$82	\$52	\$72	\$69	\$88
1993	\$53	\$75	\$69	\$93	\$70	\$79	\$56	\$77	\$74	\$94
1994	\$50	\$78	\$72	\$92	\$71	\$84	\$58	\$79	\$76	\$97
1995	\$55	\$85	\$79	\$110	\$77	\$93	\$59	\$81	\$77	\$100
1996	\$60	\$84	\$80	\$108	\$79	\$89	\$63	\$85	\$82	\$105
1997	\$62	\$90	\$80	\$105	\$84	\$95	\$65	\$88	\$84	\$109
1998	\$60	\$90	\$85	\$116	\$84	\$96	\$67	\$91	\$87	\$112
1999	\$66	\$99	\$91	\$119	\$94	\$105	\$72	\$100	\$94	\$122
2000	\$68	\$105	\$96	\$134	\$97	\$112	\$75	\$105	\$99	\$128

Refinances

Origination Year	AHS						FHA			
	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl
1989	\$37	\$64	\$52	\$60	\$33	\$95	\$45	\$60	\$58	\$73
1990	\$10	\$54	\$32	\$93	\$20	\$87	\$46	\$62	\$59	\$75
1991	\$33	\$61	\$60	\$75	\$48	\$74	\$50	\$65	\$63	\$78
1992	\$48	\$73	\$70	\$91	\$65	\$81	\$55	\$71	\$69	\$85
1993	\$46	\$73	\$66	\$86	\$64	\$81	\$56	\$73	\$71	\$87
1994	\$40	\$67	\$65	\$90	\$58	\$75	\$51	\$68	\$65	\$82
1995	\$20	\$44	\$33	\$60	\$25	\$64	\$56	\$78	\$74	\$97
1996	\$20	\$53	\$39	\$104	\$21	\$86	\$59	\$82	\$78	\$103
1997	\$45	\$84	\$85	\$110	\$68	\$101	\$67	\$93	\$90	\$117
1998	\$69	\$100	\$93	\$118	\$89	\$111	\$74	\$98	\$95	\$119
1999	\$72	\$100	\$96	\$125	\$87	\$113	\$70	\$96	\$92	\$118
2000	\$69	\$117	\$120	\$150	\$94	\$140	\$75	\$106	\$101	\$131

**Exhibit D-15-16: Original Mortgage Amount by Mortgage Market Segment - AHS vs. HMDA and FHA (In Thousands of Dollars)
Loans Originated in Metro Areas**

Home Purchases		AHS						HMDA				FHA			
Origination Year	Market Segment	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl	25th Pctl	Mean	Median	75th Pctl
1993	FHA	\$54	\$76	\$72	\$95	\$71	\$80	\$58	\$80	\$76	\$97	\$58	\$79	\$76	\$96
1994		\$52	\$80	\$74	\$97	\$74	\$87	\$59	\$81	\$77	\$99	\$60	\$81	\$78	\$99
1995		\$60	\$88	\$80	\$110	\$80	\$96	\$60	\$82	\$79	\$101	\$61	\$83	\$79	\$102
1996		\$64	\$87	\$82	\$109	\$82	\$92	\$64	\$87	\$83	\$107	\$65	\$87	\$84	\$107
1997		\$66	\$93	\$85	\$107	\$86	\$99	\$66	\$90	\$86	\$110	\$67	\$90	\$86	\$111
1998		\$65	\$93	\$90	\$118	\$87	\$100	\$69	\$93	\$88	\$114	\$69	\$93	\$89	\$114
1999		\$69	\$100	\$92	\$119	\$94	\$106	\$74	\$102	\$96	\$124	\$74	\$102	\$96	\$124
2000		\$72	\$109	\$101	\$136	\$102	\$117	\$77	\$107	\$102	\$130	\$78	\$107	\$102	\$131
1993	VA/FmHA	\$60	\$87	\$82	\$105	\$78	\$95	\$69	\$97	\$90	\$122				
1994		\$60	\$83	\$84	\$110	\$75	\$91	\$69	\$98	\$91	\$124				
1995		\$54	\$86	\$91	\$112	\$76	\$96	\$71	\$101	\$93	\$126				
1996		\$65	\$103	\$90	\$133	\$90	\$116	\$75	\$105	\$97	\$130				
1997		\$62	\$95	\$90	\$125	\$84	\$107	\$77	\$108	\$101	\$134				
1998		\$73	\$106	\$106	\$140	\$96	\$115	\$83	\$114	\$109	\$141				
1999		\$63	\$111	\$98	\$154	\$95	\$127	\$86	\$118	\$114	\$146				
2000		\$85	\$116	\$118	\$146	\$104	\$129	\$89	\$121	\$118	\$150				
1993	Conforming	\$50	\$86	\$80	\$117	\$82	\$90	\$60	\$96	\$92	\$130				
1994		\$50	\$87	\$82	\$118	\$83	\$90	\$59	\$96	\$92	\$130				
1995		\$52	\$91	\$87	\$130	\$87	\$95	\$55	\$94	\$90	\$129				
1996		\$48	\$88	\$84	\$125	\$84	\$93	\$56	\$97	\$93	\$134				
1997		\$56	\$97	\$95	\$130	\$93	\$101	\$57	\$99	\$96	\$137				
1998		\$62	\$103	\$100	\$140	\$99	\$107	\$61	\$106	\$102	\$146				
1999		\$59	\$107	\$100	\$150	\$103	\$112	\$62	\$111	\$105	\$152				
2000		\$64	\$112	\$105	\$150	\$107	\$117	\$61	\$112	\$107	\$156				
1993	Jumbo	\$240	\$292	\$288	\$340	\$276	\$309	\$236	\$317	\$271	\$338				
1994		\$240	\$285	\$270	\$350	\$270	\$300	\$232	\$311	\$268	\$333				
1995		\$225	\$275	\$276	\$310	\$259	\$291	\$234	\$313	\$270	\$339				
1996		\$239	\$283	\$270	\$350	\$270	\$295	\$240	\$323	\$277	\$350				
1997		\$246	\$290	\$295	\$350	\$276	\$303	\$240	\$328	\$279	\$352				
1998		\$254	\$294	\$290	\$348	\$284	\$303	\$264	\$363	\$304	\$394				
1999		\$268	\$312	\$332	\$350	\$302	\$322	\$280	\$398	\$328	\$420				
2000		\$295	\$322	\$342	\$350	\$313	\$330	\$284	\$403	\$336	\$435				

Refinances		AHS						HMDA				FHA			
Origination Year	Market Segment	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl	25th Pctl	Mean	Median	75th Pctl
1993	FHA	\$49	\$74	\$70	\$87	\$66	\$83	\$58	\$76	\$73	\$89	\$57	\$74	\$72	\$88
1994		\$40	\$68	\$66	\$90	\$59	\$77	\$51	\$69	\$67	\$83	\$52	\$69	\$66	\$83
1995		\$25	\$45	\$55	\$60	\$26	\$65	\$58	\$80	\$77	\$99	\$57	\$79	\$75	\$98
1996		\$20	\$39	\$20	\$63	\$17	\$61	\$59	\$83	\$79	\$104	\$61	\$84	\$80	\$104
1997		\$60	\$89	\$85	\$110	\$74	\$105	\$70	\$95	\$92	\$119	\$69	\$94	\$91	\$118
1998		\$66	\$101	\$95	\$120	\$89	\$113	\$75	\$99	\$96	\$120	\$75	\$99	\$96	\$120
1999		\$72	\$101	\$97	\$125	\$87	\$114	\$70	\$97	\$92	\$119	\$72	\$97	\$93	\$120
2000		\$69	\$118	\$128	\$150	\$95	\$142	\$79	\$110	\$105	\$135	\$77	\$108	\$103	\$133
1993	VA/FmHA	\$53	\$83	\$80	\$105	\$72	\$95	\$65	\$94	\$86	\$118				
1994		\$45	\$58	\$58	\$68	\$48	\$68	\$55	\$82	\$73	\$102				
1995		\$67	\$122	\$67	\$181	\$43	\$201	\$67	\$101	\$94	\$133				
1996		\$135	\$135	\$135	\$135	\$135	\$135	\$65	\$98	\$91	\$127				
1997		\$60	\$110	\$98	\$155	\$84	\$136	\$75	\$108	\$103	\$139				
1998		\$70	\$97	\$97	\$117	\$83	\$112	\$80	\$112	\$108	\$141				
1999		\$45	\$112	\$90	\$162	\$81	\$144	\$71	\$102	\$96	\$130				
2000		\$100	\$152	\$132	\$157	\$90	\$214	\$84	\$119	\$119	\$153				
1993	Conforming	\$35	\$71	\$60	\$100	\$68	\$74	\$57	\$93	\$85	\$123				
1994		\$26	\$65	\$55	\$95	\$61	\$69	\$48	\$84	\$75	\$114				
1995		\$30	\$74	\$70	\$100	\$62	\$87	\$42	\$82	\$72	\$114				
1996		\$28	\$67	\$60	\$91	\$57	\$77	\$40	\$80	\$70	\$111				
1997		\$45	\$82	\$72	\$110	\$77	\$87	\$42	\$83	\$74	\$116				
1998		\$60	\$101	\$96	\$137	\$95	\$106	\$59	\$101	\$94	\$138				
1999		\$64	\$110	\$100	\$150	\$105	\$116	\$50	\$94	\$84	\$130				
2000		\$55	\$102	\$100	\$139	\$95	\$110	\$39	\$86	\$71	\$120				
1993	Jumbo	\$228	\$286	\$268	\$350	\$270	\$302	\$244	\$334	\$285	\$362				
1994		\$250	\$304	\$325	\$375	\$279	\$329	\$240	\$335	\$283	\$365				
1995		\$280	\$321	\$349	\$350	\$281	\$362	\$244	\$353	\$293	\$390				
1996		\$260	\$294	\$306	\$349	\$258	\$330	\$248	\$354	\$296	\$390				
1997		\$250	\$288	\$280	\$349	\$269	\$307	\$245	\$356	\$293	\$386				
1998		\$263	\$304	\$310	\$350	\$293	\$314	\$269	\$377	\$318	\$412				
1999		\$270	\$307	\$300	\$350	\$293	\$321	\$284	\$411	\$338	\$445				
2000		\$260	\$301	\$293	\$350	\$285	\$316	\$283	\$421	\$340	\$450				

Note: No attempt is made to exclude manufactured homes from the three data sets.

Exhibit D-17-18: Distribution of Interest Rates of Fixed-Rate FHA-Insured Mortgages by Loan Purpose - AHS vs. FHA

Purchases

Origination Year	AHS						FHA					
	Standard Deviation	25th	Mean	Median	75th	95% Conf. Interval of Mean		Standard Deviation	25th	Mean	Median	75th
1989	0.9%	9.3%	9.8%	9.8%	10.0%	9.6%	9.9%	0.8%	9.5%	9.9%	10.0%	10.5%
1990	1.0%	9.3%	9.7%	10.0%	10.0%	9.5%	9.8%	0.8%	9.5%	9.7%	10.0%	10.0%
1991	1.1%	8.3%	8.9%	9.0%	9.5%	8.7%	9.1%	0.7%	9.0%	9.2%	9.5%	9.5%
1992	0.9%	7.8%	8.2%	8.0%	8.5%	8.0%	8.3%	0.7%	8.0%	8.3%	8.5%	8.5%
1993	1.0%	7.3%	7.8%	7.5%	8.0%	7.6%	7.9%	0.7%	7.0%	7.5%	7.5%	8.0%
1994	1.3%	7.3%	8.0%	8.0%	8.8%	7.8%	8.2%	1.0%	7.4%	7.9%	8.0%	8.5%
1995	1.1%	7.5%	8.0%	8.0%	8.5%	7.8%	8.2%	0.9%	7.5%	8.2%	8.0%	8.8%
1996	1.0%	7.3%	7.9%	8.0%	8.5%	7.7%	8.1%	0.8%	7.5%	7.9%	8.0%	8.5%
1997	1.2%	7.0%	7.6%	7.5%	8.0%	7.4%	7.7%	0.8%	7.5%	7.8%	8.0%	8.5%
1998	1.0%	6.9%	7.3%	7.0%	7.5%	7.1%	7.4%	0.6%	7.0%	7.2%	7.3%	7.5%
1999	1.0%	7.0%	7.5%	7.5%	8.0%	7.4%	7.6%	0.8%	6.1%	7.5%	7.5%	8.0%
2000	1.2%	7.3%	8.0%	8.0%	8.6%	7.8%	8.2%	0.7%	8.0%	8.3%	8.5%	8.8%

Refinances

Origination Year	AHS						FHA					
	Standard Deviation	25th	Mean	Median	75th	95% Conf. Interval of Mean		Standard Deviation	25th	Mean	Median	75th
1989	2.4%	9.0%	10.4%	9.8%	10.5%	8.9%	11.9%	0.9%	10.0%	10.4%	10.5%	11.0%
1990	1.2%	9.3%	10.1%	10.0%	11.0%	9.4%	10.8%	0.8%	9.5%	10.2%	10.0%	10.5%
1991	0.8%	7.5%	8.1%	8.3%	8.8%	7.8%	8.5%	0.7%	9.0%	9.6%	9.5%	10.0%
1992	1.4%	8.0%	8.3%	8.3%	8.8%	7.9%	8.6%	0.6%	8.0%	8.5%	8.5%	9.0%
1993	0.8%	7.3%	7.8%	7.8%	8.5%	7.6%	7.9%	0.6%	7.5%	7.7%	7.9%	8.0%
1994	1.1%	7.0%	7.8%	7.5%	8.0%	7.5%	8.1%	0.7%	7.5%	7.8%	8.0%	8.0%
1995	0.6%	8.0%	8.3%	8.5%	8.5%	7.9%	8.7%	0.7%	8.0%	8.3%	8.4%	8.5%
1996	1.0%	7.3%	8.1%	8.0%	9.0%	7.4%	8.8%	0.7%	7.5%	7.9%	8.0%	8.5%
1997	2.2%	7.0%	8.4%	7.6%	8.9%	7.3%	9.4%	0.6%	7.5%	8.0%	8.0%	8.5%
1998	1.5%	6.9%	7.4%	7.0%	7.5%	7.1%	7.7%	0.5%	7.0%	7.3%	7.5%	7.5%
1999	0.0%	6.9%	7.4%	7.1%	7.8%	7.2%	7.7%	0.6%	6.5%	7.3%	7.3%	7.5%
2000	1.2%	7.0%	7.6%	7.4%	8.0%	7.2%	8.0%	0.5%	8.0%	8.4%	8.5%	8.8%

Purchases

Origination Year	AHS							FHA						
	Interest Rate Category							Interest Rate Category						
	<=6	6-6.9	7-7.9	8-9.9	10-11.9	12-13.9	>=14	<=6	6-6.9	7-7.9	8-9.9	10-11.9	12-13.9	>=14
1989	0%	0%	1%	49%	48%	2%	0%	0%	0%	0%	37%	62%	0%	0%
1990	1%	0%	3%	42%	53%	1%	0%	0%	0%	1%	38%	61%	0%	0%
1991	0%	1%	13%	77%	8%	1%	2%	0%	0%	4%	79%	17%	0%	0%
1992	4%	3%	19%	69%	5%	0%	0%	2%	2%	12%	84%	0%	0%	0%
1993	2%	4%	55%	35%	2%	2%	0%	4%	6%	58%	33%	0%	0%	0%
1994	4%	6%	35%	49%	6%	0%	0%	4%	6%	38%	51%	1%	0%	0%
1995	1%	4%	34%	57%	3%	0%	0%	2%	4%	27%	65%	3%	0%	0%
1996	2%	6%	37%	52%	3%	0%	0%	3%	8%	31%	58%	0%	0%	0%
1997	8%	10%	49%	31%	2%	0%	1%	4%	8%	32%	56%	0%	0%	0%
1998	5%	23%	57%	11%	4%	0%	0%	4%	16%	70%	11%	0%	0%	0%
1999	6%	17%	51%	24%	2%	1%	0%	5%	13%	49%	32%	0%	0%	0%
2000	5%	8%	34%	47%	6%	1%	0%	1%	3%	14%	82%	0%	0%	0%

Refinances

Origination Year	AHS							FHA						
	Interest Rate Category							Interest Rate Category						
	<=6	6-6.9	7-7.9	8-9.9	10-11.9	12-13.9	>=14	<=6	6-6.9	7-7.9	8-9.9	10-11.9	12-13.9	>=14
1989	0%	0%	0%	53%	35%	0%	12%	0%	0%	0%	24%	69%	7%	0%
1990	0%	0%	0%	31%	58%	11%	0%	0%	0%	0%	28%	70%	2%	0%
1991	0%	5%	28%	67%	0%	0%	0%	0%	0%	0%	61%	39%	0%	0%
1992	6%	1%	17%	66%	6%	3%	0%	1%	1%	9%	88%	1%	0%	0%
1993	4%	7%	43%	45%	1%	0%	0%	1%	4%	45%	50%	0%	0%	0%
1994	2%	10%	50%	33%	3%	2%	0%	1%	4%	42%	51%	1%	0%	0%
1995	0%	0%	23%	77%	0%	0%	0%	0%	1%	18%	77%	3%	0%	0%
1996	0%	13%	32%	56%	0%	0%	0%	1%	4%	37%	58%	0%	0%	0%
1997	3%	14%	36%	29%	4%	13%	0%	1%	2%	31%	66%	0%	0%	0%
1998	4%	24%	59%	8%	2%	2%	1%	1%	13%	75%	11%	0%	0%	0%
1999	0%	31%	45%	21%	2%	2%	0%	2%	16%	66%	16%	0%	0%	0%
2000	4%	21%	45%	23%	4%	3%	0%	0%	0%	11%	88%	0%	0%	0%

Exhibit D-19-20: Distributon of Interest Rates of Fixed-Rate Conventional Home-Purchase Mortgages - AHS vs. MIRS

Conforming

Origination Year	AHS							MIRS						
	Standard Deviation	25th	Mean	Median	75th	95% Conf. Interval of Mean		Standard Deviation	25th	Mean	Median	75th	95% Conf. Interval of Mean	
1992	1.1%	7.8%	8.2%	8.3%	8.8%	8.1%	8.3%	0.6%	7.9%	8.2%	8.3%	8.6%	8.2%	8.2%
1993	1.5%	7.0%	7.7%	7.5%	8.0%	7.6%	7.8%	0.6%	6.9%	7.3%	7.3%	7.6%	7.3%	7.3%
1994	1.3%	7.3%	8.1%	8.0%	8.8%	7.9%	8.2%	0.9%	7.4%	8.0%	8.0%	8.6%	8.0%	8.0%
1995	1.3%	7.5%	8.0%	7.8%	8.5%	7.9%	8.2%	0.7%	7.5%	8.0%	7.9%	8.5%	8.0%	8.0%
1996	1.4%	7.5%	8.0%	8.0%	8.5%	7.9%	8.1%	0.7%	7.4%	7.8%	7.9%	8.3%	7.8%	7.8%
1997	1.2%	7.0%	7.6%	7.5%	8.0%	7.5%	7.7%	0.5%	7.4%	7.7%	7.8%	8.0%	7.7%	7.7%
1998	1.2%	6.8%	7.3%	7.1%	7.5%	7.2%	7.4%	0.5%	6.8%	7.0%	7.0%	7.3%	7.0%	7.0%
1999	1.2%	6.9%	7.4%	7.3%	7.9%	7.3%	7.5%	0.6%	6.9%	7.3%	7.3%	7.8%	7.3%	7.3%
2000	1.2%	7.0%	7.8%	7.8%	8.3%	7.7%	7.9%	0.5%	7.9%	8.1%	8.1%	8.4%	8.1%	8.1%

Jumbo

Origination Year	Standard Deviation	25th	Mean	Median	75th	95% Conf. Interval of Mean		Standard Deviation	25th	Mean	Median	75th	95% Conf. Interval of Mean	
1992	0.7%	7.5%	8.1%	8.3%	8.8%	7.8%	8.4%	0.7%	8.0%	8.3%	8.4%	8.9%	8.3%	8.4%
1993	0.8%	7.0%	7.5%	7.5%	7.8%	7.3%	7.8%	0.6%	6.9%	7.3%	7.3%	7.6%	7.2%	7.3%
1994	0.8%	7.3%	7.7%	7.5%	8.0%	7.4%	8.0%	1.0%	7.0%	7.6%	7.5%	8.4%	7.6%	7.7%
1995	0.7%	7.0%	7.6%	7.8%	8.3%	7.4%	7.9%	0.6%	7.6%	8.0%	7.9%	8.4%	7.9%	8.0%
1996	0.9%	7.3%	7.7%	7.5%	8.3%	7.4%	8.0%	0.7%	7.5%	7.9%	7.9%	8.4%	7.9%	7.9%
1997	0.7%	7.0%	7.4%	7.4%	7.6%	7.2%	7.6%	0.4%	7.5%	7.7%	7.8%	8.0%	7.7%	7.8%
1998	1.0%	6.9%	7.2%	7.0%	7.4%	7.0%	7.5%	0.4%	7.0%	7.2%	7.3%	7.5%	7.2%	7.2%
1999	0.9%	6.8%	7.2%	7.1%	7.5%	7.0%	7.5%	0.6%	6.9%	7.3%	7.3%	7.6%	7.3%	7.3%
2000	0.9%	7.0%	7.5%	7.8%	8.0%	7.2%	7.8%	0.5%	7.9%	8.1%	8.1%	8.5%	8.1%	8.2%

Conforming

Interest Rate Category	AHS										MIRS								
	Origination Year										Origination Year								
	1992	1993	1994	1995	1996	1997	1998	1999	2000		1992	1993	1994	1995	1996	1997	1998	1999	2000
<=6	2.6%	5.1%	3.9%	2.4%	3.5%	4.0%	5.9%	6.0%	4.0%	0.5%	2.6%	1.8%	0.6%	1.0%	0.4%	1.0%	0.5%	0.1%	
6-6.9	3.4%	13.0%	7.5%	4.8%	4.4%	10.9%	27.5%	20.0%	11.2%	2.7%	22.9%	8.6%	4.4%	7.6%	4.0%	40.3%	29.4%	0.7%	
7-7.9	25.5%	49.6%	34.0%	43.8%	38.9%	54.3%	50.7%	51.2%	44.7%	25.6%	63.3%	35.4%	46.1%	47.3%	65.7%	55.7%	56.0%	36.4%	
8-9.9	62.4%	26.4%	47.8%	41.9%	48.5%	25.9%	12.1%	18.6%	34.3%	70.7%	11.1%	53.8%	47.9%	43.7%	29.4%	2.9%	14.0%	61.7%	
10-11.9	5.3%	3.8%	5.3%	4.7%	3.0%	3.8%	2.4%	3.2%	4.1%	0.6%	0.1%	0.3%	0.9%	0.5%	0.4%	0.2%	0.2%	1.2%	
12-13.9	0.7%	1.5%	1.1%	2.1%	0.7%	0.9%	1.3%	0.7%	1.4%	0%	0%	0%	0%	0%	0%	0%	0.0%	0.0%	
>=14	0.2%	0.5%	0.4%	0.4%	1.0%	0.2%	0.3%	0.3%	0.3%	0%	0%	0%	0%	0%	0%	0%	0.0%	0.0%	

Jumbo

Interest Rate Category	AHS										MIRS								
	Origination Year										Origination Year								
	1992	1993	1994	1995	1996	1997	1998	1999	2000		1992	1993	1994	1995	1996	1997	1998	1999	2000
<=6	0.0%	3.4%	4.0%	4.6%	4.7%	4.3%	1.4%	7.0%	7.7%	2.1%	3.0%	7.5%	0.4%	1.5%	0.2%	0.8%	0.2%	0.1%	
6-6.9	4.6%	13.4%	11.4%	3.8%	8.2%	14.0%	27.8%	25.5%	9.9%	2.1%	22.4%	11.7%	4.2%	5.2%	2.3%	22.8%	28.0%	0.6%	
7-7.9	26.3%	59.4%	45.3%	56.9%	46.8%	65.1%	61.9%	50.7%	50.3%	17.9%	63.5%	43.6%	45.7%	43.7%	67.4%	73.2%	59.7%	33.1%	
8-9.9	69.1%	23.9%	39.2%	34.8%	37.8%	14.2%	6.7%	15.4%	32.1%	77.6%	11.0%	37.1%	49.5%	49.6%	30.0%	3.2%	12.0%	65.4%	
10-11.9	0.0%	0.0%	0.0%	0.0%	2.4%	2.3%	0.6%	1.5%	0.0%	0.2%	0.0%	0.1%	0.2%	0.1%	0.0%	0.1%	0.1%	0.8%	
12-13.9	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	
>=14	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	

Note: Comparisons are restricted to mortgages for non-mobile homes with 1-unit in structure.

Exhibit D-21-22 : Distribution of Loan-to-Value (LTV) Ratio of FHA-Insured Mortgages at Origination by Loan Purpose - AHS vs.FHA

Purchases Origination Year	AHS						FHA			
	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl
1989	89%	90%	95%	100%	87%	92%	90%	91%	94%	96%
1990	88%	89%	95%	100%	86%	91%	91%	92%	94%	96%
1991	90%	93%	96%	100%	90%	95%	92%	93%	95%	97%
1992	90%	93%	95%	99%	88%	99%	93%	94%	96%	97%
1993	87%	89%	95%	98%	87%	91%	94%	94%	96%	97%
1994	89%	88%	95%	99%	85%	91%	94%	94%	96%	97%
1995	85%	89%	95%	99%	86%	92%	94%	94%	96%	97%
1996	92%	92%	97%	100%	90%	94%	94%	94%	96%	97%
1997	90%	95%	98%	100%	91%	99%	94%	94%	96%	97%
1998	91%	94%	97%	100%	92%	95%	94%	94%	96%	97%
1999	87%	91%	97%	100%	87%	94%	97%	96%	97%	98%
2000	87%	92%	97%	100%	89%	96%	97%	96%	97%	98%

Refinances Origination Year	AHS						FHA			
	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl
1989	48%	60%	65%	71%	49%	72%	75%	81%	85%	95%
1990	16%	73%	33%	73%	15%	131%	75%	80%	85%	91%
1991	38%	59%	56%	79%	47%	71%	77%	81%	85%	90%
1992	57%	76%	77%	91%	68%	83%	82%	85%	87%	93%
1993	53%	69%	72%	86%	62%	75%	81%	85%	88%	93%
1994	53%	66%	73%	83%	60%	72%	78%	83%	86%	92%
1995	20%	52%	47%	79%	30%	75%	79%	83%	85%	91%
1996	23%	49%	43%	86%	28%	69%	81%	84%	86%	93%
1997	57%	68%	69%	79%	59%	78%	81%	84%	85%	92%
1998	56%	70%	73%	86%	65%	75%	83%	86%	88%	93%
1999	50%	66%	67%	77%	58%	74%	81%	85%	86%	92%
2000	63%	73%	76%	88%	66%	81%	80%	84%	85%	92%

Purchases Origination Year	AHS LTV Category				FHA LTV Category			
	<80%	80.1-90%	90.1-95%	>95%	<80%	80.1-90%	90.1-95%	>95%
1989	13%	14%	23%	49%	8%	17%	35%	40%
1990	15%	16%	17%	52%	6%	16%	35%	42%
1991	7%	18%	17%	58%	6%	13%	27%	54%
1992	10%	16%	25%	50%	4%	12%	27%	57%
1993	14%	17%	21%	48%	3%	10%	23%	64%
1994	17%	12%	22%	50%	3%	10%	23%	65%
1995	16%	17%	19%	49%	3%	9%	23%	64%
1996	8%	11%	17%	64%	3%	9%	24%	63%
1997	13%	13%	12%	63%	3%	10%	25%	62%
1998	9%	13%	17%	61%	3%	9%	25%	63%
1999	18%	12%	15%	56%	2%	5%	8%	85%
2000	18%	10%	9%	63%	2%	4%	6%	88%

Refinances Origination Year	AHS LTV Category				FHA LTV Category			
	<80%	80.1-90%	90.1-95%	>95%	<80%	80.1-90%	90.1-95%	>95%
1989	83%	9%	8%	0%	33%	33%	16%	18%
1990	76%	19%	0%	6%	35%	39%	15%	12%
1991	78%	11%	6%	5%	30%	45%	14%	10%
1992	53%	21%	14%	12%	41%	21%	22%	16%
1993	63%	22%	10%	5%	22%	38%	24%	16%
1994	68%	17%	5%	10%	28%	38%	19%	14%
1995	76%	0%	0%	24%	27%	45%	15%	13%
1996	74%	12%	14%	0%	23%	43%	19%	15%
1997	78%	10%	1%	0%	11%	24%	17%	13%
1998	68%	17%	4%	11%	19%	41%	25%	15%
1999	78%	17%	1%	4%	23%	42%	21%	13%
2000	63%	14%	12%	11%	25%	46%	16%	14%

Note: LTV estimates from the FHA administrative data exclude any financed upfront MIP.

**Exhibit D-21-22a: Comparison of Purchase House Price in AHS and Appraised House Value in FHA Data --
Home-Purchase Loans with FHA Insurance (in thousands of dollars)**

Origination Year	AHS (Purchase House Price)						FHA (Appraised House Value)			
	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl
1989	\$50	\$91	\$68	\$92	\$75	\$107	\$48	\$71	\$64	\$82
1990	\$53	\$111	\$74	\$95	\$91	\$132	\$50	\$73	\$66	\$85
1991	\$52	\$89	\$73	\$95	\$75	\$102	\$51	\$76	\$69	\$89
1992	\$57	\$91	\$77	\$96	\$78	\$105	\$55	\$78	\$72	\$92
1993	\$62	\$100	\$78	\$104	\$83	\$117	\$58	\$80	\$76	\$98
1994	\$60	\$123	\$82	\$110	\$98	\$147	\$60	\$83	\$78	\$101
1995	\$61	\$96	\$85	\$125	\$86	\$106	\$61	\$85	\$80	\$104
1996	\$65	\$94	\$87	\$115	\$88	\$100	\$65	\$89	\$85	\$110
1997	\$68	\$96	\$86	\$115	\$90	\$103	\$68	\$92	\$88	\$114
1998	\$68	\$99	\$90	\$125	\$92	\$106	\$70	\$95	\$90	\$116
1999	\$74	\$112	\$97	\$126	\$105	\$119	\$75	\$103	\$97	\$126
2000	\$75	\$119	\$103	\$145	\$109	\$129	\$79	\$109	\$103	\$133

Exhibit D-23-24 : Distribution of Loan-to-Value (LTV) Ratio at Origination for Conventional Purchases - AHS vs. MIRS

Conforming												
Origination Year	AHS						MIRS					
	25th	Mean	Median	75th	95% Conf. Interval of		25th	Mean	Median	75th	95% Conf. Interval of	
					Mean	77%					Mean	75%
1989	67%	79%	79%	90%	74%	77%	69%	75%	79%	85%	74%	75%
1990	63%	74%	79%	89%	72%	76%	68%	74%	79%	84%	74%	75%
1991	70%	79%	81%	93%	77%	81%	67%	74%	78%	86%	74%	74%
1992	71%	78%	82%	92%	77%	80%	70%	76%	80%	90%	76%	77%
1993	70%	78%	82%	93%	76%	79%	70%	77%	80%	90%	77%	77%
1994	74%	80%	86%	95%	78%	81%	74%	80%	80%	93%	80%	80%
1995	74%	81%	87%	95%	79%	82%	74%	80%	80%	95%	80%	80%
1996	74%	80%	86%	95%	79%	82%	73%	79%	80%	93%	79%	79%
1997	73%	82%	85%	95%	80%	84%	74%	79%	80%	93%	79%	79%
1998	72%	83%	82%	95%	79%	87%	73%	79%	80%	93%	79%	79%
1999	70%	87%	86%	95%	83%	91%	74%	78%	80%	90%	78%	79%
2000	71%	83%	83%	95%	80%	87%	74%	78%	80%	90%	78%	78%

Jumbo												
Origination Year	AHS						MIRS					
	25th	Mean	Median	75th	95% Conf. Interval of		25th	Mean	Median	75th	95% Conf. Interval of	
					Mean	90%					Mean	80%
1989	80%	87%	87%	100%	84%	90%	75%	78%	80%	80%	77%	78%
1990	76%	83%	80%	100%	80%	87%	75%	78%	80%	80%	78%	78%
1991	74%	82%	80%	89%	79%	86%	74%	78%	80%	80%	77%	78%
1992	80%	88%	83%	100%	84%	91%	75%	79%	80%	89%	79%	79%
1993	80%	89%	90%	100%	86%	92%	75%	79%	80%	90%	79%	80%
1994	79%	87%	88%	98%	84%	90%	77%	81%	80%	90%	80%	81%
1995	79%	87%	89%	96%	83%	91%	78%	81%	80%	90%	81%	82%
1996	80%	89%	90%	100%	86%	92%	78%	81%	80%	90%	81%	81%
1997	61%	97%	77%	94%	66%	128%	77%	80%	80%	90%	80%	80%
1998	60%	72%	69%	81%	68%	75%	75%	79%	80%	80%	79%	79%
1999	64%	90%	64%	86%	59%	120%	75%	78%	80%	80%	78%	79%
2000	64%	79%	64%	88%	68%	90%	75%	78%	80%	80%	77%	78%

Conforming								
Origination Year	AHS				MIRS			
	LTV Category				LTV Category			
	<80%	80.1-90%	90.1-95%	>95%	<80%	80.1-90%	90.1-95%	>95%
1989	52%	26%	10%	13%	73%	20%	5%	2%
1990	54%	24%	9%	13%	74%	18%	6%	3%
1991	44%	24%	13%	19%	72%	18%	8%	2%
1992	43%	27%	15%	16%	64%	21%	13%	2%
1993	42%	26%	14%	18%	62%	20%	17%	2%
1994	37%	26%	15%	22%	53%	20%	23%	4%
1995	37%	22%	18%	23%	54%	18%	24%	4%
1996	37%	23%	18%	22%	57%	17%	21%	5%
1997	38%	22%	15%	25%	57%	17%	20%	6%
1998	42%	22%	14%	22%	58%	16%	19%	7%
1999	39%	19%	16%	26%	60%	15%	18%	6%
2000	41%	23%	13%	23%	63%	14%	16%	7%

Jumbo								
Origination Year	AHS				MIRS			
	LTV Category				LTV Category			
	<80%	80.1-90%	90.1-95%	>95%	<80%	80.1-90%	90.1-95%	>95%
1989	26%	32%	6%	36%	80%	18%	1%	0%
1990	50%	24%	0%	26%	79%	18%	2%	1%
1991	45%	31%	6%	18%	77%	21%	2%	1%
1992	26%	36%	8%	30%	69%	29%	2%	1%
1993	13%	43%	15%	28%	68%	29%	3%	0%
1994	28%	36%	9%	27%	61%	31%	7%	1%
1995	27%	37%	8%	28%	63%	24%	12%	1%
1996	22%	29%	17%	31%	66%	21%	12%	1%
1997	53%	18%	11%	17%	70%	18%	11%	1%
1998	71%	15%	10%	4%	76%	15%	7%	1%
1999	71%	7%	3%	18%	79%	13%	7%	2%
2000	63%	14%	12%	11%	83%	10%	5%	2%

Note: Comparisons are restricted to mortgages for non-mobile homes with 1-unit in structure.

Exhibit D-23-24a: Purchase Price of Conventional Conforming Home-Purchase Loans (in thousands of dollars) - AHS vs. MIRS

Origination Year	AHS						MIRS					
	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean	
1989	\$62	\$119	\$90	\$144	\$110	\$129	\$75	\$124	\$117	\$163	\$123	\$125
1990	\$62	\$129	\$103	\$158	\$115	\$144	\$76	\$126	\$119	\$165	\$126	\$127
1991	\$59	\$118	\$90	\$142	\$106	\$131	\$78	\$131	\$121	\$171	\$130	\$132
1992	\$68	\$128	\$100	\$157	\$119	\$137	\$82	\$130	\$120	\$165	\$129	\$131
1993	\$70	\$177	\$108	\$165	\$161	\$194	\$80	\$127	\$116	\$162	\$126	\$128
1994	\$65	\$162	\$103	\$159	\$148	\$176	\$76	\$123	\$114	\$160	\$123	\$124
1995	\$68	\$116	\$102	\$155	\$110	\$121	\$79	\$125	\$117	\$160	\$124	\$126
1996	\$72	\$121	\$110	\$159	\$116	\$126	\$85	\$133	\$124	\$170	\$133	\$134
1997	\$75	\$123	\$114	\$160	\$118	\$129	\$92	\$142	\$132	\$179	\$141	\$142
1998	\$84	\$140	\$127	\$184	\$134	\$146	\$95	\$150	\$136	\$188	\$149	\$150
1999	\$74	\$136	\$125	\$177	\$129	\$143	\$99	\$157	\$142	\$198	\$156	\$157
2000	\$80	\$144	\$130	\$182	\$136	\$151	\$105	\$168	\$151	\$215	\$167	\$168

Note: Comparisons are restricted to mortgages for non-mobile homes with 1-unit in structure.

Exhibit D-25-26: Shares of FHA-Insured Mortgages by Payment Plan - Stratified by Loan Purpose - AHS vs. FHA

Home-Purchases

FRMs	AHS			FHA
	Estimate	95% Conf. Interval		
1989	92%	89%	96%	98%
1990	94%	91%	97%	98%
1991	89%	84%	93%	92%
1992	81%	75%	86%	81%
1993	83%	78%	88%	81%
1994	79%	73%	85%	73%
1995	84%	78%	90%	74%
1996	85%	80%	89%	72%
1997	82%	77%	87%	67%
1998	92%	87%	97%	93%
1999	93%	90%	96%	93%
2000	93%	90%	96%	92%

ARMs	AHS			FHA
	Estimate	95% Conf. Interval		
1989	4%	1%	6%	1%
1990	2%	0%	3%	1%
1991	8%	4%	12%	7%
1992	17%	12%	22%	18%
1993	14%	9%	18%	19%
1994	17%	12%	23%	27%
1995	11%	6%	16%	26%
1996	13%	9%	17%	28%
1997	12%	8%	16%	33%
1998	3%	1%	6%	7%
1999	6%	3%	8%	7%
2000	4%	1%	7%	8%

Other	AHS			FHA
	Estimate	95% Conf. Interval		
1989	4%	2%	5%	0.82%
1990	5%	3%	6%	0.90%
1991	3%	0%	5%	0.86%
1992	3%	0%	5%	0.31%
1993	3%	1%	6%	0.18%
1994	4%	1%	7%	0.17%
1995	5%	2%	8%	0.09%
1996	2%	0%	5%	0.03%
1997	6%	3%	9%	0.01%
1998	5%	0%	9%	0.02%
1999	2%	0%	3%	0.02%
2000	3%	1%	5%	0.01%

Refinances

FRMs	AHS			FHA
	Estimate	95% Conf. Interval		
1989	92%	78%	107%	100%
1990	91%	73%	108%	100%
1991	100%	100%	100%	97%
1992	88%	80%	95%	91%
1993	91%	85%	96%	94%
1994	87%	79%	95%	92%
1995	89%	69%	109%	87%
1996	88%	66%	110%	85%
1997	86%	75%	97%	72%
1998	96%	91%	100%	95%
1999	95%	89%	101%	94%
2000	100%	100%	100%	90%

ARMs	AHS			FHA
	Estimate	95% Conf. Interval		
1989	8%	-7%	22%	0%
1990	9%	-8%	27%	0%
1991	0%	0%	0%	3%
1992	11%	4%	18%	9%
1993	9%	4%	15%	6%
1994	11%	3%	19%	8%
1995	0%	0%	0%	13%
1996	12%	-10%	34%	15%
1997	11%	1%	22%	28%
1998	4%	0%	9%	5%
1999	3%	-1%	8%	6%
2000	0%	0%	0%	10%

Other	AHS			FHA
	Estimate	95% Conf. Interval		
1989	0%	0%	0%	0.01%
1990	0%	0%	0%	0.02%
1991	0%	0%	0%	0.02%
1992	1%	-1%	4%	0.01%
1993	0%	0%	0%	0.00%
1994	2%	-1%	5%	0.00%
1995	11%	-9%	31%	0.00%
1996	0%	0%	0%	0.00%
1997	2%	-2%	7%	0.00%
1998	0%	0%	0%	0.00%
1999	2%	-1%	5%	0.00%
2000	0%	0%	0%	0.00%

Exhibit D-27-28 : Shares of Conventional Purchase Mortgages by Payment Plan - AHS vs. MIRS

Conforming

FRMs	AHS			MIRS		
	Origination Year	Estimate	95% Conf. Interval	Estimate	95% Conf. Interval	
1992	88%	85%	91%	81%	81%	82%
1993	88%	85%	90%	82%	81%	83%
1994	83%	80%	85%	64%	63%	64%
1995	91%	89%	94%	71%	71%	72%
1996	91%	89%	94%	77%	77%	78%
1997	95%	93%	96%	81%	81%	82%
1998	97%	96%	98%	90%	90%	91%
1999	95%	94%	97%	83%	83%	83%
2000	94%	92%	96%	81%	81%	81%

ARMs						
Origination Year	Estimate	95% Conf. Interval		Estimate	95% Conf. Interval	
1992	12%	10%	14%	19%	18%	19%
1993	12%	10%	15%	18%	17%	19%
1994	17%	15%	20%	36%	36%	37%
1995	9%	7%	10%	29%	28%	29%
1996	9%	7%	11%	23%	22%	23%
1997	5%	4%	7%	19%	18%	19%
1998	3%	2%	4%	10%	9%	10%
1999	5%	3%	6%	17%	17%	17%
2000	6%	4%	8%	19%	19%	19%

Jumbo

FRMs	AHS			MIRS		
	Origination Year	Estimate	95% Conf. Interval	Estimate	95% Conf. Interval	
1992	66%	49%	84%	57%	55%	59%
1993	80%	67%	93%	53%	51%	55%
1994	60%	47%	73%	28%	27%	29%
1995	80%	66%	95%	33%	32%	35%
1996	75%	64%	87%	35%	34%	37%
1997	91%	83%	99%	53%	52%	54%
1998	93%	87%	99%	60%	59%	61%
1999	94%	87%	100%	42%	41%	43%
2000	86%	77%	96%	29%	28%	29%

ARMs						
Origination Year	Estimate	95% Conf. Interval		Estimate	95% Conf. Interval	
1992	34%	17%	50%	43%	41%	45%
1993	20%	9%	31%	47%	45%	49%
1994	40%	28%	52%	72%	71%	73%
1995	20%	8%	31%	67%	65%	68%
1996	25%	14%	35%	65%	63%	66%
1997	9%	1%	17%	47%	46%	48%
1998	7%	1%	13%	40%	39%	41%
1999	6%	0%	13%	58%	57%	59%
2000	14%	4%	23%	71%	71%	72%

Note: Comparisons are restricted to mortgages for non-mobile homes with 1-unit in structure.

Exhibit D-29: Share of First-time Buyers - AHS vs. Chicago Title Data
Home Purchase Loans Originated in Metropolitan Areas

Origination Year	AHS			Chicago Title
	Estimate	95% Conf. Interval		
1995	44%	41%	47%	42%
1996	46%	43%	49%	45%
1997	45%	42%	49%	47%
1998	44%	41%	47%	46%
1999	45%	42%	48%	45%

Exhibit D-30: Share of First-time Buyers by Mortgage Market Segment - AHS vs. RFS

Origination Year: 1989-1991

Mortgage Market Segment	AHS			RFS		
	Estimate	95% Conf. Interval		Estimate	95% Conf. Interval	
FHA	56%	52%	60%	63%	59%	67%
VA/FmHA	49%	43%	55%	49%	42%	56%
Conforming	38%	36%	40%	41%	39%	43%
Jumbo	15%	9%	21%	19%	12%	26%
All	43%	41%	44%	45%	44%	47%

Exhibit D-31: Share of First-time Buyer - AHS vs. FHA
FHA Home-Purchase Loans

Origination Year	AHS			FHA
	Estimate	95% Conf. Interval		
1991	60%	53%	67%	63%
1992	61%	55%	68%	66%
1993	53%	46%	60%	66%
1994	66%	60%	73%	68%
1995	55%	48%	63%	71%
1996	63%	56%	69%	75%
1997	68%	62%	74%	80%
1998	65%	59%	72%	81%
1999	60%	54%	65%	81%
2000	64%	57%	71%	81%

Exhibit D-32-37: Distribution of FHA-Insured Mortgages by Borrower Race - AHS vs. FHA

Home-Purchases

Origination Year	Race	AHS			FHA
		Estimate	95% Conf. Interval		
1989	White	81%	76%	86%	80%
	Black	11%	7%	15%	9%
	Hispanic	6%	3%	9%	8%
	Other	2%	0%	4%	3%
1990	White	78%	72%	83%	80%
	Black	13%	9%	17%	9%
	Hispanic	6%	3%	9%	9%
	Other	3%	1%	5%	3%
1991	White	79%	74%	85%	79%
	Black	11%	6%	15%	10%
	Hispanic	7%	4%	11%	9%
	Other	2%	0%	5%	2%
1992	White	84%	79%	89%	78%
	Black	5%	2%	8%	9%
	Hispanic	9%	5%	13%	10%
	Other	2%	0%	4%	3%
1993	White	72%	65%	78%	76%
	Black	15%	10%	20%	10%
	Hispanic	11%	7%	16%	11%
	Other	2%	0%	5%	3%
1994	White	75%	68%	81%	72%
	Black	10%	6%	15%	12%
	Hispanic	12%	8%	17%	13%
	Other	3%	1%	5%	3%
1995	White	75%	69%	82%	68%
	Black	10%	5%	15%	12%
	Hispanic	12%	7%	17%	15%
	Other	3%	0%	5%	4%
1996	White	73%	67%	79%	67%
	Black	13%	9%	18%	12%
	Hispanic	11%	7%	15%	17%
	Other	3%	0%	5%	5%
1997	White	64%	57%	70%	65%
	Black	21%	15%	26%	12%
	Hispanic	12%	8%	16%	18%
	Other	4%	1%	6%	5%
1998	White	64%	57%	70%	63%
	Black	17%	12%	22%	13%
	Hispanic	15%	10%	19%	18%
	Other	5%	2%	7%	5%
1999	White	68%	63%	74%	61%
	Black	15%	11%	19%	14%
	Hispanic	13%	9%	17%	19%
	Other	4%	2%	6%	7%
2000	White	67%	60%	73%	59%
	Black	14%	9%	18%	15%
	Hispanic	16%	11%	21%	20%
	Other	4%	1%	6%	7%

Refinances

Origination Year	Race	AHS			FHA
		Estimate	95% Conf. Interval		
1989	White	79%	57%	100%	78%
	Black	21%	0%	43%	7%
	Hispanic	0%	0%	0%	11%
	Other	0%	0%	0%	5%
1990	White	61%	31%	90%	78%
	Black	18%	-5%	41%	7%
	Hispanic	21%	-4%	47%	12%
	Other	0%	0%	0%	3%
1991	White	95%	85%	105%	82%
	Black	5%	-5%	15%	7%
	Hispanic	0%	0%	0%	8%
	Other	0%	0%	0%	2%
1992	White	85%	77%	93%	86%
	Black	5%	0%	10%	5%
	Hispanic	6%	1%	12%	6%
	Other	4%	0%	8%	3%
1993	White	85%	78%	91%	82%
	Black	5%	1%	9%	5%
	Hispanic	5%	1%	9%	9%
	Other	5%	1%	9%	3%
1994	White	82%	72%	91%	76%
	Black	13%	5%	22%	9%
	Hispanic	4%	-1%	9%	12%
	Other	1%	-1%	2%	4%
1995	White	79%	53%	105%	70%
	Black	10%	-9%	28%	10%
	Hispanic	11%	-9%	31%	15%
	Other	0%	0%	0%	5%
1996	White	74%	43%	105%	71%
	Black	26%	-5%	57%	10%
	Hispanic	0%	0%	0%	13%
	Other	0%	0%	0%	6%
1997	White	73%	59%	87%	65%
	Black	15%	3%	26%	11%
	Hispanic	12%	3%	22%	18%
	Other	0%	0%	0%	7%
1998	White	78%	69%	86%	67%
	Black	11%	4%	17%	11%
	Hispanic	9%	4%	15%	14%
	Other	2%	-1%	5%	8%
1999	White	70%	59%	82%	63%
	Black	23%	12%	33%	13%
	Hispanic	7%	1%	12%	15%
	Other	0%	0%	0%	9%
2000	White	69%	52%	87%	60%
	Black	12%	-1%	25%	16%
	Hispanic	15%	1%	28%	15%
	Other	4%	-4%	12%	8%

Exhibit D-33-40: Share of Loans by Race and Mortgage Market Segment - Metropolitan Areas Only

Home-Purchases

Origination Year	Race	FHA					VA/FmHA			
		AHS			HMDA	FHA	AHS			HMDA
		Estimate	95% Conf. Interval				Estimate	95% Conf. Interval		
1993	White	73%	66%	79%	75%	75%	76%	66%	86%	81%
	Black	14%	9%	20%	11%	10%	11%	3%	19%	12%
	Hispanic	12%	7%	16%	11%	11%	8%	2%	15%	5%
	Other	1%	0%	2%	3%	3%	5%	0%	9%	2%
1994	White	73%	67%	80%	70%	71%	77%	68%	86%	77%
	Black	11%	6%	16%	14%	12%	18%	9%	26%	14%
	Hispanic	13%	8%	17%	14%	14%	3%	0%	6%	6%
	Other	3%	1%	5%	3%	3%	2%	-1%	6%	3%
1995	White	75%	68%	82%	67%	67%	70%	58%	81%	76%
	Black	9%	4%	14%	15%	13%	21%	11%	31%	16%
	Hispanic	13%	7%	18%	15%	16%	8%	2%	14%	6%
	Other	3%	0%	6%	3%	4%	1%	-1%	4%	2%
1996	White	71%	64%	77%	66%	66%	75%	64%	86%	77%
	Black	14%	9%	19%	14%	12%	13%	5%	22%	14%
	Hispanic	12%	8%	17%	17%	17%	6%	0%	12%	6%
	Other	3%	1%	5%	3%	5%	6%	0%	11%	2%
1997	White	61%	55%	68%	64%	64%	73%	61%	84%	77%
	Black	22%	16%	28%	14%	13%	17%	7%	27%	14%
	Hispanic	12%	8%	17%	18%	18%	9%	2%	16%	6%
	Other	4%	1%	6%	4%	5%	2%	-1%	4%	2%
1998	White	62%	55%	69%	63%	62%	77%	67%	86%	77%
	Black	18%	12%	24%	14%	14%	10%	3%	18%	14%
	Hispanic	15%	11%	20%	19%	19%	9%	3%	16%	6%
	Other	5%	2%	8%	4%	5%	4%	0%	8%	3%
1999	White	65%	59%	71%	62%	60%	77%	67%	87%	76%
	Black	17%	12%	21%	15%	14%	7%	1%	13%	14%
	Hispanic	14%	10%	18%	19%	20%	9%	3%	16%	7%
	Other	4%	2%	6%	4%	7%	6%	1%	12%	3%
2000	White	63%	56%	70%	60%	58%	74%	63%	86%	75%
	Black	15%	10%	20%	16%	16%	10%	2%	18%	15%
	Hispanic	18%	13%	23%	21%	20%	13%	4%	21%	7%
	Other	4%	2%	7%	4%	7%	4%	-1%	8%	3%

Origination Year	Race	Conforming			
		AHS			HMDA
		Estimate	95% Conf. Interval		
1993	White	83%	80%	86%	87%
	Black	6%	4%	8%	4%
	Hispanic	6%	4%	8%	5%
	Other	5%	3%	7%	5%
1994	White	82%	79%	85%	84%
	Black	7%	5%	9%	5%
	Hispanic	7%	5%	9%	6%
	Other	4%	3%	6%	5%
1995	White	83%	80%	86%	83%
	Black	5%	3%	8%	6%
	Hispanic	7%	5%	9%	7%
	Other	4%	3%	6%	5%
1996	White	82%	79%	86%	84%
	Black	6%	4%	9%	5%
	Hispanic	8%	5%	10%	6%
	Other	4%	2%	5%	5%
1997	White	79%	75%	82%	83%
	Black	8%	5%	10%	5%
	Hispanic	9%	7%	11%	6%
	Other	5%	3%	6%	6%
1998	White	81%	78%	84%	83%
	Black	7%	4%	9%	5%
	Hispanic	7%	5%	9%	6%
	Other	6%	4%	8%	6%
1999	White	79%	76%	82%	81%
	Black	9%	6%	11%	5%
	Hispanic	8%	6%	10%	7%
	Other	5%	3%	6%	7%
2000	White	78%	75%	81%	79%
	Black	7%	5%	9%	6%
	Hispanic	9%	6%	11%	8%
	Other	6%	4%	8%	7%

Origination Year	Race	Jumbo			
		AHS			HMDA
		Estimate	95% Conf. Interval		
1993	White	85%	75%	95%	85%
	Black	3%	-2%	8%	2%
	Hispanic	3%	-2%	9%	3%
	Other	8%	0%	16%	10%
1994	White	86%	78%	94%	85%
	Black	3%	-1%	7%	3%
	Hispanic	4%	0%	8%	3%
	Other	7%	1%	13%	9%
1995	White	90%	81%	99%	85%
	Black	2%	-2%	7%	3%
	Hispanic	3%	-3%	8%	3%
	Other	5%	-2%	11%	9%
1996	White	87%	78%	95%	85%
	Black	6%	-1%	12%	3%
	Hispanic	3%	-1%	7%	3%
	Other	5%	-1%	10%	9%
1997	White	94%	87%	100%	85%
	Black	1%	-1%	4%	3%
	Hispanic	0%	0%	0%	3%
	Other	5%	-1%	11%	10%
1998	White	80%	71%	89%	85%
	Black	3%	-1%	6%	2%
	Hispanic	5%	0%	10%	3%
	Other	12%	4%	19%	10%
1999	White	81%	72%	90%	82%
	Black	2%	-1%	5%	3%
	Hispanic	8%	2%	13%	4%
	Other	9%	3%	15%	12%
2000	White	81%	71%	91%	80%
	Black	4%	-1%	9%	3%
	Hispanic	3%	-1%	8%	4%
	Other	12%	4%	20%	13%

Refinances

Origination Year	Race	FHA					VA/FmHA			
		AHS			HMDA	FHA	AHS			HMDA
		Estimate	95% Conf. Interval				Estimate	95% Conf. Interval		
1993	White	84%	77%	91%	81%	82%	75%	62%	89%	84%
	Black	6%	1%	10%	6%	6%	11%	1%	22%	9%
	Hispanic	5%	1%	9%	9%	9%	8%	0%	17%	4%
	Other	5%	1%	9%	3%	3%	5%	-2%	12%	3%
1994	White	82%	73%	92%	72%	75%	69%	49%	89%	77%
	Black	13%	4%	21%	11%	9%	21%	3%	39%	14%
	Hispanic	4%	-1%	10%	14%	12%	5%	-3%	14%	5%
	Other	1%	-1%	2%	4%	4%	5%	-5%	14%	3%
1995	White	100%	100%	100%	69%	70%	100%	100%	100%	75%
	Black	0%	0%	0%	13%	11%	0%	0%	0%	16%
	Hispanic	0%	0%	0%	14%	15%	0%	0%	0%	6%
	Other	0%	0%	0%	4%	5%	0%	0%	0%	3%
1996	White	70%	35%	105%	70%	70%	100%	100%	100%	76%
	Black	30%	-5%	65%	13%	11%	0%	0%	0%	16%
	Hispanic	0%	0%	0%	14%	13%	0%	0%	0%	5%
	Other	0%	0%	0%	3%	6%	0%	0%	0%	3%
1997	White	72%	57%	87%	66%	64%	76%	56%	97%	75%
	Black	14%	2%	26%	12%	11%	17%	-1%	35%	16%
	Hispanic	14%	3%	25%	18%	18%	7%	-6%	19%	5%
	Other	0%	0%	0%	4%	7%	0%	0%	0%	3%
1998	White	77%	68%	86%	69%	66%	89%	78%	99%	78%
	Black	10%	4%	17%	13%	11%	6%	-2%	13%	14%
	Hispanic	10%	4%	16%	14%	15%	5%	-2%	13%	5%
	Other	3%	-1%	6%	4%	8%	0%	0%	0%	3%
1999	White	71%	59%	82%	65%	62%	71%	52%	90%	75%
	Black	22%	12%	33%	15%	14%	8%	-3%	19%	15%
	Hispanic	7%	1%	13%	15%	15%	21%	4%	38%	6%
	Other	0%	0%	0%	4%	9%	0%	0%	0%	3%
2000	White	68%	50%	87%	63%	59%	65%	31%	98%	71%
	Black	13%	-1%	26%	17%	17%	0%	0%	0%	19%
	Hispanic	15%	1%	29%	16%	15%	24%	-6%	53%	6%
	Other	4%	-4%	12%	4%	8%	12%	-10%	34%	3%

Origination Year	Race	Conforming			
		AHS			HMDA
		Estimate	95% Conf. Interval		
1993	White	90%	88%	92%	89%
	Black	3%	2%	4%	2%
	Hispanic	4%	3%	5%	4%
	Other	3%	2%	4%	5%
1994	White	90%	88%	93%	84%
	Black	4%	2%	5%	5%
	Hispanic	4%	2%	5%	6%
	Other	2%	1%	4%	5%
1995	White	79%	69%	88%	83%
	Black	14%	6%	23%	7%
	Hispanic	3%	-1%	6%	5%
	Other	4%	0%	9%	4%
1996	White	88%	81%	94%	85%
	Black	7%	2%	13%	7%
	Hispanic	3%	0%	6%	5%
	Other	2%	-1%	5%	4%
1997	White	87%	84%	91%	83%
	Black	6%	4%	9%	8%
	Hispanic	4%	2%	6%	5%
	Other	2%	1%	4%	4%
1998	White	89%	86%	91%	86%
	Black	5%	3%	6%	5%
	Hispanic	5%	3%	7%	5%
	Other	2%	1%	3%	5%
1999	White	86%	82%	89%	81%
	Black	8%	5%	11%	8%
	Hispanic	4%	2%	5%	7%
	Other	3%	1%	4%	5%
2000	White	83%	78%	88%	78%
	Black	8%	4%	11%	10%
	Hispanic	6%	3%	8%	8%
	Other	4%	1%	6%	4%

Origination Year	Race	Jumbo			
		AHS			HMDA
		Estimate	95% Conf. Interval		
1993	White	88%	79%	97%	84%
	Black	0%	0%	0%	1%
	Hispanic	5%	-2%	11%	2%
	Other	7%	1%	14%	13%
1994	White	82%	66%	98%	84%
	Black	5%	-5%	15%	2%
	Hispanic	0%	0%	0%	3%
	Other	13%	-1%	26%	12%
1995	White	100%	100%	100%	86%
	Black	0%	0%	0%	2%
	Hispanic	0%	0%	0%	2%
	Other	0%	0%	0%	9%
1996	White	91%	75%	108%	87%
	Black	0%	0%	0%	2%
	Hispanic	0%	0%	0%	2%
	Other	9%	-8%	25%	9%
1997	White	90%	78%	102%	87%
	Black	6%	-5%	16%	2%
	Hispanic	2%	-2%	6%	2%
	Other	2%	-2%	6%	9%
1998	White	88%	80%	96%	86%
	Black	4%	-1%	8%	2%
	Hispanic	3%	-1%	7%	2%
	Other	6%	1%	12%	10%
1999	White	79%	64%	94%	84%
	Black	0%	0%	0%	2%
	Hispanic	6%	-2%	13%	3%
	Other	16%	2%	29%	10%
2000	White	78%	60%	95%	82%
	Black	0%	0%	0%	3%
	Hispanic	9%	-3%	21%	5%
	Other	13%	-1%	27%	10%

**Exhibit D-41-44: Borrower Income by Loan Purpose and Mortgage Market Segment - Metropolitan Areas
(In Thousands of Dollars)**

Home-Purchases

FHA Origination Year	AHS					HMDA				
	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl
1993	\$33	\$52	\$48	\$65	\$48	\$56	\$28	\$48	\$37	\$48
1994	\$34	\$49	\$47	\$56	\$45	\$52	\$28	\$41	\$37	\$48
1995	\$28	\$54	\$49	\$71	\$49	\$60	\$29	\$42	\$38	\$49
1996	\$35	\$52	\$50	\$62	\$48	\$55	\$29	\$44	\$39	\$51
1997	\$35	\$57	\$50	\$70	\$52	\$62	\$30	\$48	\$39	\$52
1998	\$39	\$58	\$52	\$70	\$53	\$63	\$30	\$44	\$40	\$53
1999	\$39	\$72	\$60	\$80	\$64	\$80	\$32	\$47	\$42	\$56
2000	\$40	\$74	\$62	\$90	\$65	\$82	\$33	\$49	\$44	\$59

VA/FmHA

Origination Year	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl
1993	\$36	\$55	\$50	\$67	\$48	\$62	\$31	\$52	\$41	\$54
1994	\$38	\$56	\$48	\$72	\$50	\$61	\$31	\$45	\$41	\$53
1995	\$35	\$56	\$56	\$71	\$48	\$64	\$32	\$46	\$42	\$55
1996	\$40	\$59	\$53	\$74	\$52	\$65	\$33	\$48	\$44	\$57
1997	\$33	\$59	\$54	\$67	\$49	\$69	\$34	\$54	\$45	\$59
1998	\$36	\$69	\$58	\$77	\$57	\$82	\$35	\$50	\$46	\$60
1999	\$40	\$83	\$66	\$93	\$64	\$101	\$37	\$52	\$48	\$62
2000	\$40	\$69	\$60	\$86	\$57	\$82	\$39	\$55	\$51	\$66

Conventional Conforming

Origination Year	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl
1993	\$35	\$60	\$53	\$76	\$57	\$63	\$35	\$61	\$49	\$69
1994	\$36	\$61	\$55	\$80	\$59	\$64	\$34	\$56	\$49	\$68
1995	\$39	\$67	\$60	\$89	\$63	\$70	\$34	\$57	\$49	\$69
1996	\$35	\$61	\$55	\$80	\$58	\$65	\$35	\$60	\$51	\$72
1997	\$39	\$73	\$62	\$90	\$69	\$77	\$37	\$62	\$54	\$75
1998	\$36	\$75	\$62	\$95	\$71	\$80	\$38	\$65	\$56	\$78
1999	\$40	\$88	\$66	\$97	\$81	\$95	\$39	\$67	\$57	\$81
2000	\$43	\$87	\$68	\$94	\$80	\$94	\$41	\$70	\$60	\$85

Conventional Jumbo

Origination Year	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl
1993	\$100	\$119	\$111	\$153	\$103	\$134	\$91	\$162	\$120	\$169
1994	\$90	\$112	\$107	\$145	\$99	\$125	\$91	\$157	\$119	\$166
1995	\$100	\$135	\$119	\$180	\$118	\$152	\$96	\$166	\$125	\$175
1996	\$100	\$118	\$110	\$145	\$106	\$129	\$98	\$171	\$127	\$180
1997	\$85	\$181	\$140	\$277	\$148	\$215	\$99	\$173	\$129	\$180
1998	\$100	\$171	\$161	\$205	\$150	\$192	\$103	\$184	\$136	\$192
1999	\$90	\$208	\$163	\$283	\$172	\$244	\$107	\$191	\$141	\$200
2000	\$90	\$209	\$141	\$263	\$167	\$250	\$111	\$197	\$146	\$205

Refinances

FHA Origination Year	AHS						HMDA			
	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl
1993	\$42	\$64	\$56	\$74	\$57	\$70	\$32	\$109	\$44	\$58
1994	\$33	\$52	\$48	\$64	\$45	\$58	\$28	\$44	\$40	\$54
1995	\$33	\$49	\$46	\$71	\$34	\$64	\$29	\$45	\$40	\$55
1996	\$30	\$40	\$33	\$41	\$22	\$59	\$30	\$49	\$43	\$55
1997	\$35	\$63	\$58	\$78	\$50	\$75	\$30	\$122	\$42	\$57
1998	\$42	\$73	\$64	\$83	\$61	\$85	\$30	\$48	\$43	\$59
1999	\$41	\$90	\$65	\$100	\$66	\$113	\$34	\$51	\$46	\$62
2000	\$50	\$86	\$63	\$90	\$57	\$115	\$35	\$53	\$48	\$63

VA/FmHA

Origination Year	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl
1993	\$27	\$56	\$56	\$72	\$47	\$64	\$36	\$90	\$50	\$66
1994	\$37	\$56	\$53	\$66	\$48	\$65	\$31	\$57	\$46	\$63
1995	\$35	\$64	\$35	\$94	\$23	\$104	\$32	\$49	\$45	\$60
1996	\$130	\$130	\$130	\$130	\$130	\$130	\$35	\$59	\$48	\$62
1997	\$48	\$113	\$73	\$142	\$65	\$161	\$36	\$135	\$47	\$63
1998	\$51	\$73	\$68	\$90	\$62	\$85	\$30	\$53	\$48	\$66
1999	\$30	\$80	\$45	\$95	\$44	\$116	\$38	\$58	\$53	\$70
2000	\$42	\$78	\$77	\$125	\$53	\$104	\$39	\$56	\$51	\$69

Conventional Conforming

Origination Year	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl
1993	\$43	\$69	\$64	\$90	\$67	\$72	\$41	\$69	\$57	\$78
1994	\$38	\$67	\$60	\$90	\$64	\$71	\$35	\$60	\$50	\$71
1995	\$33	\$69	\$66	\$93	\$59	\$79	\$34	\$58	\$50	\$70
1996	\$35	\$69	\$61	\$88	\$59	\$79	\$36	\$60	\$51	\$71
1997	\$45	\$81	\$67	\$100	\$75	\$87	\$36	\$60	\$52	\$72
1998	\$47	\$86	\$73	\$104	\$80	\$91	\$41	\$68	\$59	\$82
1999	\$52	\$108	\$76	\$118	\$99	\$118	\$38	\$65	\$55	\$79
2000	\$40	\$90	\$66	\$94	\$77	\$103	\$36	\$64	\$54	\$78

Conventional Jumbo

Origination Year	25th Pctl	Mean	Median	75th Pctl	95% Conf. Interval of Mean		25th Pctl	Mean	Median	75th Pctl
1993	\$87	\$108	\$105	\$129	\$94	\$122	\$100	\$182	\$134	\$198
1994	\$102	\$127	\$118	\$139	\$111	\$142	\$95	\$179	\$128	\$190
1995	\$117	\$228	\$260	\$300	\$145	\$310	\$98	\$187	\$132	\$200
1996	\$90	\$113	\$94	\$106	\$87	\$139	\$101	\$190	\$137	\$203
1997	\$97	\$181	\$189	\$250	\$138	\$223	\$101	\$188	\$136	\$198
1998	\$102	\$198	\$192	\$244	\$170	\$225	\$108	\$202	\$148	\$217
1999	\$124	\$252	\$253	\$289	\$190	\$315	\$108	\$205	\$149	\$220
2000	\$120	\$259	\$253	\$343	\$191	\$328	\$104	\$200	\$143	\$211

Appendix E

Procedures for Tracking Mortgage Records Across AHS Survey Years

This appendix describes how we linked observations on mortgages from multiple years of the AHS national survey. Archer, Ling and McGill (1996, 1997) employed a similar method to construct their longitudinal AHS files.⁷⁴ Our procedures have made several improvements to ensure that the same set of households and loans are being followed throughout the entire study period. We will use the construction of the 1991-1993-1995-1997-1999 study panel as an example. Panels consisting of other survey years can be created in a similar fashion.

To construct the 1991-1993-1995-1997-1999 panel, we first created the 1991-1993 matched file. The very first step was to select all the owner-occupied sample housing units with primary mortgages identified in the 1991 survey as having been originated in the 1989-1990 period.⁷⁵ Only loans originated in the 2-year period prior to the interview were included because these homeowners are expected to more accurately remember facts about their mortgages than those with mortgages originated in earlier years. We included both home-purchase and refinance mortgages. This resulted in a total of 3,528 records, representing 7,567,469 loans.

We then tracked the status of these loans in the 1993 survey. Since the AHS follows housing units rather than households or mortgages, a sequence of data steps was needed to ensure that we were following the same set of households and loans across survey years. The 3,528 records were first matched to the 1993 survey by the Unit Control Number provided in the public-use files. This match yielded a total of 2,824 records in the 1993 survey. The attrition of the observations at this step occurred largely because rural housing units were oversampled in the 1991 survey; none of these housing units were included in 1993 or any of the subsequent survey years.^{76,77}

⁷⁴ Wayne R. Archer, David C. Ling and Gary A. McGill, “The effect of income and collateral constraints on residential mortgage terminations.” *Regional Science and Urban Economics*, 26: 235-261 (1996). Wayne R. Archer, David C. Ling and Gary A. McGill, “Demographic versus option-driven mortgage terminations.” *Journal of Housing Economics*, 6: 137-163 (1997).

⁷⁵ We also used the *ISTATUS* (interview status) variable to exclude sample housing units occupied by people who all have usual residence elsewhere (URE). In addition, using the TEN variable, we only included records where the survey respondent is the owner of the house unit. These two screening procedures are intended to ensure that the survey respondents are in the best position to provide accurate information about their mortgage. Only a handful of records was filtered out by these two procedures.

⁷⁶ Similar sample changes occurred between the 1993 and 1995 surveys and between the 1995 and 1997 surveys. In 1993, there were a total of 10,279 cases in the “neighborhood” sample, none of which were included in 1995. In 1995, there were 5,697 cases in the 6-city metro supplement sample that were not included in 1993 and 1997. We thank the Census Bureau (Barbara Williams), HUD PD&R (Dave Vandenbroucke), and ICF Consulting (Gregory Watson) for providing guidance on this issue.

⁷⁷ A very small proportion of this sample attrition results from Type C non-interviews. These include units lost through demolitions or disasters and mobile homes that were moved or were abandoned. These units can be identified by the *ISTATUS* variable.

The use of control numbers alone does not guarantee that we are following the same housing units. In a few rare instances, observations across two surveys could have the same control number but in fact they represent different housing units. This could happen when a replacement unit was built in the same location or the wrong unit was enumerated in the first survey. These cases can be identified by the *SAMEDU* variable.⁷⁸ Seven observations were excluded from our 1991-1993 merge file as a result of this.

The next step was to check whether these matched housing units were occupied by the same households in the later survey year. This is a crucial step because, usually, when a household moves, the original mortgage attached to the house is prepaid (that is, terminated). For the purpose of this analysis, we considered a sample housing unit to be occupied by the same household if at least some of the household members remain the same from the last interview. This check was performed using the *SAMEHH* variable, which indicates whether some of the household members lived in the same unit in the last survey. We found that, of the 2,824 sample housing units in the 1991-1993 merged file, 235 were occupied by new households and, therefore, were excluded from the subsequent analysis.

This left us with a total of 2,582 observations. Among them, 191 were no longer associated with a mortgage in 1993, according to the survey respondents. The next challenge was to ensure that the remaining sample of housing units in the 1993 survey contained the same set of mortgages as the 1991 survey. For this analysis, we depended on a comparison of the mortgage origination date reported in the two survey years. A later mortgage origination date reported in the 1993 survey would be an indication that the homeowner had refinanced his/her mortgage between the 1991 and 1993 interviews. Our tabulation showed 857 cases in this category. However, if the origination date in the 1993 survey was earlier than the one reported in the 1991 survey, this indicated an error in reporting by the homeowner in one of the survey years (327 cases). Homeowners did not provide loan origination date information in 1993 for 65 cases of the matched file. All these records were excluded. As shown in Appendix F, only 1,143 of the 3,528 loans in the 1991 survey were still active in the 1993 survey.

Following similar data steps, these 1,143 records were then matched to the 1995, 1997 and 1999 survey data. By 1999, only 99 records (representing a total of 223,752 mortgages originated in the 1989-1990 period) were left in this study panel. Longitudinal panels containing records from other survey years were created in a similar fashion. Appendix F presents the sample sizes and loan attrition situation for each panel used in Chapter Five.

⁷⁸ See page 21 of the *Codebook for the American Housing Survey, Volume 3: 1997 SAS Files and Questionnaire*, draft, February 16, 1998.

Appendix F

Tabulations Supporting the Analysis in Chapter Five

Status of 1989-1990 Mortgage Originations from 1991 AHS in Subsequent Surveys (1991-1993-1995-1997-1999 Panel)

Number of Mortgages Originated in 1989-1990 from 1991 Survey		3,528
Housing unit not surveyed in 1993		
Not Surveyed	704	711
Not Same Housing Unit	7	
Different mortgage in 1993		
Household Moved	235	1,283
Paid Off	191	
Refinanced	857	
Incomplete or inconsistent origination year		
Inconsistent	327	392
Missing	65	
Number of Mortgages Originated in 1989-1990 in 1993 Survey		1,143
Housing unit not surveyed in 1995		
Not Surveyed	130	132
Not Same Housing Unit	2	
Different mortgage in 1995		
Household Moved	111	469
Paid Off	68	
Refinanced	290	
Incomplete or inconsistent origination year		
Inconsistent	55	56
Missing	1	
Number of Mortgages Originated in 1989-1990 in 1995 Survey		486
Housing unit not surveyed in 1997		
Not Surveyed	68	70
Not Same Housing Unit	2	
Different mortgage in 1997		
Household Moved	47	158
Paid Off	64	
Refinanced	47	
Incomplete or inconsistent origination year		
Inconsistent	39	39
Missing	0	

**Status of 1989-1990 Mortgage Originations from 1991 AHS in Subsequent Surveys
(Continued)
(1991-1993-1995-1997-1999 Panel)**

Number of Mortgages Originated in 1989-1990 in 1997 Survey		219
Housing unit not surveyed in 1999		17
Not Surveyed	16	
Not Same Housing Unit	1	
Different mortgage in 1999		100
Household Moved	20	
Paid Off	27	
Refinanced	53	
Incomplete or inconsistent origination year		3
Inconsistent	3	
Missing	0	
Number of Mortgages Originated in 1989-1990 in 1999 Survey		99

**Status of 1991-1992 Mortgage Originations from 1993 AHS in Subsequent Surveys
(1993-1995-1997-1999 Panel)**

Number of Mortgages Originated in 1991-1992 from 1993 Survey		4,809
Housing unit not surveyed in 1995		1,114
Not Surveyed	1,105	
Not Same Housing Unit	9	
Different mortgage in 1995		1,468
Household Moved	374	
Paid Off	256	
Refinanced	838	
Incomplete or inconsistent origination year		818
Inconsistent	809	
Missing	9	
Number of Mortgages Originated in 1991-1992 in 1995 Survey		1,409
Housing unit not surveyed in 1997		179
Not Surveyed	171	
Not Same Housing Unit	8	
Different mortgage in 1997		453
Household Moved	134	
Paid Off	198	
Refinanced	121	
Incomplete or inconsistent origination year		232
Inconsistent	232	
Missing	0	
Number of Mortgages Originated in 1991-1992 in 1997 Survey		545
Housing unit not surveyed in 1999		49
Not Surveyed	49	
Not Same Housing Unit	0	
Different mortgage in 1999		227
Household Moved	67	
Paid Off	51	
Refinanced	109	
Incomplete or inconsistent origination year		6
Inconsistent	6	
Missing	0	
Number of Mortgages Originated in 1991-1992 in 1999 Survey		263

**Status of 1993-1994 Mortgage Originations from 1995 AHS in Subsequent Surveys
(1995-1997-1999 Panel)**

Number of Mortgages Originated in 1993-1994 from 1995 Survey		5,407
Housing unit not surveyed in 1997		1,199
Not Surveyed	1,171	
Not Same Housing Unit	28	
Different mortgage in 1997		1,417
Household Moved	377	
Paid Off	721	
Refinanced	319	
Incomplete or inconsistent origination year		1,368
Inconsistent	1,366	
Missing	2	
Number of Mortgages Originated in 1993-1994 in 1997 Survey		1,423
Housing unit not surveyed in 1999		154
Not Surveyed	153	
Not Same Housing Unit	1	
Different mortgage in 1999		519
Household Moved	149	
Paid Off	147	
Refinanced	223	
Incomplete or inconsistent origination year		31
Inconsistent	31	
Missing	0	
Number of Mortgages Originated in 1993-1994 in 1999 Survey		719

**Status of 1995-1996 Mortgage Originations from 1997 AHS in 1999 Survey
(1997-1999 Panel)**

Number of Mortgages Originated in 1995-1996 from 1997 Survey		2,869
Housing unit not surveyed in 1999		412
Not Surveyed	406	
Not Same Housing Unit	6	
Different mortgage in 1999		1,021
Household Moved	295	
Paid Off	316	
Refinanced	410	
Incomplete or inconsistent origination year		105
Inconsistent	105	
Missing	0	
Number of Mortgages Originated in 1995-1996 in 1999 Survey		1,331

Source: Authors' calculations of the AHS National Surveys

Note: Unweighted counts

Appendix G

Mortgage and Purchase Modules in the 1997 AHS Questionnaire

Mortgage Module

The next questions are about mortgages or other loans that are secured by the property. You may check your records if you wish. If you filled out your letter you may wish to refer to it for these questions

- (C) Arrange Callback to Speak with Spouse and/or Co-Owner or to Provide Respondent Time to Research Information
- (P) Proceed

- RAMORT Some people take out a special mortgage called a Reverse Annuity Mortgage or Home Equity Conversion Mortgage that borrows against the equity in their homes to give them retirement money or income. Some of these loans do not have to be paid back during the owner's lifetime because it will be paid by the sale of the home when the estate is settled. Some provide monthly income over a specified period of time, after which it must be paid back. Have you heard of this type of mortgage?
- (1) Yes
 - (2) No
- RAMAP Have you applied for this kind of mortgage that would give you payments?
- (1) Yes
 - (2) No
- RAMAPP Were you approved for this loan?
- (1) Yes
 - (2) No
- RAMPAY Have you received any of these payments?
- (1) Yes
 - (2) No
- RAMTME Do you receive payments for your lifetime or only for a specified period of time?
- (1) Lifetime
 - (2) Specified
- RAMYRS How many years is the loan for?
- (1-30) Years
- RAMBOR How much will be borrowed?
- | | |
|--------------|-------------------|
| (\$1-999997) | \$1-999,997 |
| (\$999998) | \$999,998 or more |
| \$.00 | |
- How often do you receive the payments?
- (1) Monthly
 - (2) Quarterly
 - (3) Other regular payment (specify)
 - (4) Lump sum
- RAMSCH
RAMSCHS
- RAMAMT How much ... you receive in ... payments?
- | | |
|------------|-------------|
| (1-999997) | \$1-999,997 |
|------------|-------------|

(999,998 or more) \$999,998 or more
\$.00

RAMMO What month and year did you begin to receive payments?
RAMYR (Month)
 (Year)

RAMRAT Is the interest rate fixed or variable rate?
 (1) Fixed
 (2) Variable
 (3) Don't know

RAMFI Is your loan Federally insured?
 (1) Yes
 (2) No
 (3) Don't know

RAMEVC Does this loan have an eviction protection clause?
 (1) Yes
 (2) No
 (3) Don't know

 What is the current interest rate?
 (Rounded down to the nearest 1/4 percent
 (0 = no fraction, 1 = 1/4 percent
 2 = 1/2 percent, 3 = 3/4 percent)
 percent

RAMIRW
RAMIRF

MG Not counting Home Equity loans, is there a mortgage or any loans on this ...?
 (1) Yes
 (2) No

MGP1 **POSSIBLE PROBLEM**
Earlier you told me that someone not living here pays some of
the mortgage or utility costs. Was that for the utility costs
the mortgage, or both?
 (1) Previous answer to RET12b incorrect
 (2) Utility costs only
 (3) Mortgage costs only
 (4) Both utility and mortgage costs

MGP2 **POSSIBLE PROBLEM**
Then there is a mortgage or other loan on this ... is that correct?
 (1) Yes ? (CHANGE ANSWER IN MORT2A TO "YES")
 (2) No ? (PROBE AS NECESSARY TO DETERMINE IF THERE IS A MORTGAGE)

HEL Do you ... have a Home Equity loan or Home Equity line
of credit?
 (1) Yes
 (2) No

REGMOR ... many mortgages (or loans) are there now on the ...?
 (0) mortgages
 (1-7)

SUBMOR Did you get your mortgage through a State or local government
program that provides lower cost mortgages?
 (1) Yes
 (2) No

MATBUY Did you get the first mortgage the same year you bought your home?
 (1) Yes
 (2) No

NEWMOR With regard to the first mortgage, did you get a new mortgage or did you assume someone else's mortgage?
(1) New
(2) Assumed
(3) Wrap around

AMMORT1 How much was left to pay off when you assumed it?
(1-999997) \$1-999,997
(999,998) \$999,998 or more
\$.00

TERM1 How many years remained on the mortgage then?
(1-40) Years

YRMOR What year did you get the mortgage?
(00-97) 1 Year

When you first obtained THIS mortgage, how many years was it for?
(1-40) Years

TERM2 Can Vary
CANVAR Can Vary

AMRTZ At your current payments, how long would it take to pay off the loan?
(1-40) Years

AMMORT2 How much was borrowed?
(1-999997) \$1-999,997
(999,998) \$999,998 or more
\$.00

Does this mortgage cover?? (READ ANSWER CATEGORIES BELOW)
(1) Yes (2) No

PINCOP Other homes or apartments besides this one?
MFARM Farm land?
MCOM A business on this property?

RESMOR How much of the ... applies just to your home?
(1-999997) \$1-999,997
(999,998) \$999,998 or more
\$.00

What is the current interest rate on the mortgage?
(Rounded down to nearest 1/4 percent)
(0 = no fraction, 1 = 1/4 percent
2 = 1/2 percent, 3 = 3/4 percent)
percent

INTW
INTF

PMT What is the current monthly payment?
(Include as much of PITI as they pay)
(1-9997) \$1-9997
(9998) \$9998 or more
\$.00

Besides principal and interest, does the payment include:
(READ ANSWER CATEGORIES)
(1) Yes (2) No

TAXPMT Property taxes?
INSPMT Homeowner's insurance?
OTHMOT Anything else? (exclude anything already mentioned)

AMTM How much were the other charges last year?
(Exclude property tax and homeowner's insurance)
(1-999997) \$1-999,997
(999,998) \$999,998 or more

\$.00

MORTIN Is the mortgage an FHA, VA, Farmers Home Administration Mortgage or some other type?
(1) FHA
(2) VA
(3) Farmer's Home Administration Mortgage
(4) Some other type

BANK Did you borrow money from a bank or other organization OR did you borrow it from an individual?
(1) Bank or Organization
(2) Individual

SELL Was that the former owner of the home?
(1) Yes
(2) No

VARY Are the payments on this loan the same during the whole length of the mortgage?
(1) Yes
(2) No

How do they change?
Mark (x) all the apply

FIXED 1 Change in taxes or insurance, or due to decline in principle balance
ARM 2 Change based on interest rates
GPM 3 Rise at fixed schedule during part of loan
GPMW 4 Rise at fixed schedule during whole length of loan
BLOON 5 Last payment biggest
VARM 7 Other, specify

ARMASK Do they change for any other reason?
(1) Yes
(2) No

LOON Of the total amount you borrowed, what percentage will have to be paid off in this last payment?
(1) 1-25 percent
(2) 26-50 percent
(3) 51-75 percent
(4) 76-100 percent

MATBY2 Did you get the second mortgage the same year you bought your home?
(1) Yes
(2) No

NEWMR2 With regard to the second mortgage, did you get a new mortgage or did you assume someone else's mortgage?
(1) New
(2) Assumed
(3) Wrap around

AMMRT12 How much was left to pay off when you assumed it?
(1-999997) \$1-999,997
(999,998) \$999,998 or more
\$.00

TERM12 How many years remained on the mortgage then?
(1-40) Years

YRMOR2 What year did you get the mortgage?
(00-97) 1 Year

When you first obtained THIS mortgage, how many years was it for?
(1-40)

TERM2 Years
CANVAR2 Can Vary

AMRTZ2 At your current payments, how long would it take to pay off the loan?
(1-40) Years

AMMRT22 How much was borrowed?
(1-999997) \$1-999,997
(999,998) \$999,998 or more
\$.00

Does this mortgage cover?? (READ ANSWER CATEGORIES BELOW)
(1) Yes (2) No

PINCO2 Other homes or apartments besides this one?
MFARM2 Farm land?
MCOM2 A business on this property?

RESMR2 How much of the ... applies just to your home?
(1-999997) \$1-999,997
(999,998) \$999,998 or more
\$.00

What is the current interest rate on the mortgage?
(Rounded down to nearest 1/4 percent)
(0 = no fraction, 1 = 1/4 percent
2 = 1/2 percent, 3 = 3/4 percent)
percent

INTW2
INTF2

PMT2 What is the current monthly payment?
(Include as much of PITI as they pay)
(1-9997) \$1-9997
(9998) \$9998 or more
\$.00

Besides principal and interest, does the payment include
(READ ANSWER CATEGORIES)
(1) Yes (2) No

TXPMT2 Property taxes?
INPMT2 Homeowner's insurance?
OTPM2 Anything else? (exclude anything already mentioned)

AMTM2 How much were the other charges last year?
(Exclude property tax and homeowner's insurance)
(1-999997) \$1-999,997
(999,998) \$999,998 or more
\$.00

MORTN2 Is the mortgage an FHA, VA, Farmers Home Administration Mortgage
or some other type?
(1) FHA
(2) VA
(3) Farmer's Home Administration Mortgage
(4) Some other type

BANK2 Did you borrow money from a bank or other organization OR
did you borrow it from an individual?
(1) Bank or Organization
(2) Individual

SELL2 Was that the former owner of the home?
(1) Yes
(2) No

VARY2 Are the payments on this loan the same during the whole length of the mortgage?
(1) Yes
(2) No

How do they change?
Mark (x) all the apply

FIXED2 1 Change in taxes or insurance, or due to decline in principle balance
ARM2 2 Change based on interest rates
GPM2 3 Rise at fixed schedule during part of loan
GPMW2 4 Rise at fixed schedule during whole length of loan
BLOON2 5 Last payment biggest
VARM2 7 Other, specify

ARMASK2 Do they change for any other reason?
(1) Yes
(2) No

LOON2 Of the total amount you borrowed, what percentage will have to be paid off in this last payment?
(1) 1-25 percent
(2) 26-50 percent
(3) 51-75 percent
(4) 76-100 percent

AMMRT3 For the third mortgage, how much did you borrow?
(1-999997) \$1-999,997
(999,998) \$999,998 or more
\$.00

PMT3 What is your current monthly payment for the third mortgage?
(1-9997) \$1-9997
(9998) \$9998 or more
\$.00

AMMRT4 For the other mortgages, how much did you borrow?
(1-999997) \$1-999,997
(999,998) \$999,998 or more
\$.00

PMT4 What is your current monthly payment for the other mortgages?
(1-9997) \$1-9997
(9998) \$9998 or more
\$.00

HENUM How many home equity loans do you have?
(1-100) Loans

HETYP1 Some people have a home equity loan that allows them to borrow against it as often as they wish up to a fixed limit. Other loans are a one-time, lump sum payment, which must be repaid over a period of time
What kind of home equity loan ... Is it a line of credit, or lump sum?
(1) Line of credit
(2) Lump sum

HECR1 What is your total credit limit on your ... home equity loan?
(1-999997) \$1-999,997
(999,998) \$999,998 or more
\$.00

HELMP1 How much was the lump sum ...
(1-999997) \$1-999,997

(999,998) \$999,998 or more
\$.00

HEBAL1 Do you now have an outstanding loan borrowed against ...
(1) Yes
(2) No

HEBAM1 What is your current (outstanding) balance ...
(1-999997) \$1-999,997
(999,998) \$999,998 or more
\$.00

 What is the current interest rate ...
(Round down to nearest 1/4 percent)
(0 = no fraction, 1 = 1/4 percent
2 = 1/2 percent, 3 = 3/4 percent)
 percent

HEINW1
HEINF1

HEPMT1 What was your last monthly payment?
(1-9997) \$1-9997
(9998) \$9998 or more
\$.00

HETYP2 What kind of home equity loan is the second loan? Is it a line of
credit, or lump sum?
(1) Line of credit
(2) Lump sum

HECR2 What is your total credit limit on your second home equity loan?
(1-999997) \$1-999,997
(999,998) \$999,998 or more
\$.00

HELMP2 How much was the lump sum on the second loan?
(1-999997) \$1-999,997
(999,998) \$999,998 or more
\$.00

HEBAL2 Do you now have an outstanding loan borrowed against the second loan?
(1) Yes
(2) No

HEBAM2 What is your current (outstanding) balance on the second loan?
(1-1999997) \$1-1999,997
(999,998) \$999,998 or more
\$.00

 What is the current interest rate on the second loan?
(Round down to nearest 1/4 percent)
(0 = no fraction, 1 = 1/4 percent
2 = 1/2 percent, 3 = 3/4 percent)
 percent

HEINW2
HEINF2

HEPMT2 What was your last monthly payment on the second loan?
(1-9997) \$1-9997
(9998) \$9998 or more
\$.00

HETYP3 What kind of home equity loan is the third loan? Is it a line of
credit, or lump sum?
(1) Line of credit
(2) Lump sum

HECR3 What is your total credit limit on your third home equity loan?
 (1-999997) \$1-999,997
 (999,998) \$999,998 or more
 \$.00

HELMP3 How much was the lump sum on the third loan?
 (1-999997) \$1-999,997
 (999,998) \$999,998 or more
 \$.00

HEBAL3 Do you now have an outstanding loan borrowed against the third loan?
 (1) Yes
 (2) No

HEBAM3 What is your current (outstanding) balance on the third loan?
 (1-999997) \$1-999,997
 (999,998) \$999,998 or more
 \$.00

 What is the current interest rate on the third loan?
 (Round down to nearest 1/4 percent)
 (0 = no fraction, 1 = 1/4 percent
 2 = 1/2 percent, 3 = 3/4 percent)
 percent

HEINW3
 HEINF3

HEPMT3 What was your last monthly payment on the third loan?
 (1-9997) \$1-9997
 (9998) \$9998 or more
 \$.00

Purchase Module

Variable	Questionnaire Description
FRSTOCQ	Were you the first ... to occupy this home or did someone else live here before you? (1) First occupants (2) Previously occupied
CUSHOMQ	How did ... obtain ... home? Did ...? (READ CATEGORIES UNTIL A "YES" IS OBTAINED) (1) Buy the house already built? (2) Sign a sales agreement that included the land as well as the cost of building the house? (INCLUDE BOTH UNITS THAT WERE UNDER CONSTRUCTION AND THOSE NOT YET STARTED) (3) Have a general contractor build it on YOUR OWN LAND? (ALSO INCLUDES LEASED LAND) (4) Build it yourself on YOUR OWN LAND? (INCLUDE PERSON ACTING AS OWN GENERAL CONTRACTOR. ALSO INCLUDES LEASED LAND) (5) Receive it as a gift or inheritance?
CUSHOMR	Last time we recorded that you ... This time you reported that you Which one should I keep as the correct answer? (6) ...
PREOCCQ	Did ... receive the home as a gift or inheritance or did ... purchase it? (1) Gift or inheritance (2) Purchased
PREOCCR	Last time we recorded that ... This time you reported that ... Which one should I keep as the correct answer? (1) Gift or inheritance

(2) Purchased

WHNGETQ When did this household ... the ...?
(IF LAND AND BUILDING BOUGHT AT DIFFERENT TIMES, BUILDING ONLY)
(1900-1997)

WHNGETR Last time we recorded that this household ... the ... in ...
This time we recorded that you ... the ... in ...
Which year should I keep as the correct answer?
1900 - 1997

HOUSEHOLD ROSTER
LINE NAME

When SEE NAMES ACROSS bought this ..., did anyone make inquiries or have
this ... inspected or tested for any of the following?

(READ ALL CATEGORIES)

Enter (1) yes or (2) no

TASB	Asbestos
TPBPAI	Lead-based paint
TRADON	Radon
TPBPIP	Lead pipes
TPBSOL	Lead solder on copper pipes
TWATER	Water quality
TALWIR	Aluminum wiring

TREP Were any repairs or replacements made as a result of the
inquiry, inspection, or test?
(1) yes
(2) no

TADJ Were any adjustments made to the purchase price of the house as a
result of the inquiry, inspection, or test?
(1) yes
(2) no

LPRICEQ What was the price?
(EXCLUDE CLOSING COST)
(FOR MOBILE HOMES, EXCLUDE VALUE OF THE LAND)
(1-999997) \$1-\$999,997
(999998) \$999,998 or more \$

CPRICEQ What was the construction cost as well as the value of the
land at that time?
(1-999997) \$1-\$999,997
(999998) \$999,998 or more \$

PRICER Last time we recorded that the ...
This time you reported that the ...
Which price should I keep as the correct answer?
(1-999997) \$1-\$999,997
(999998) \$999,998 or more \$

DWNPAYQ Was the main source of down payment the sale of a previous home
savings, or something else?
(IF BOUGHT OUTRIGHT, ENTER MAIN SOURCE OF FULL PAYMENT)
(1) Sale of previous home if sold during 12 months
prior to purchase of new home
(2) Savings or cash on hand
(3) Sale of other investment
(4) Borrowing, other than a mortgage on this property
(5) Inheritance or gift
(6) Land where building was built used for financing
(7) Other (Specify)
(8) No down payment

DWNPAYR Last time we recorded that your main source of down payment was This time you reported that your main source of down payment was ... Which one should I keep as the correct answer?
(1) ...

FRSTHOQ ... ever owned a home before?
(1) yes
(2) no

FRSTHOR Last time we recorded that ... of the owners living here ... owned a home before This time you reported that ... of the owners living here ... owned a home before. Which one should I keep as the correct answer?
(1) yes, owned a home before
(2) no, never owned a home before

OWNLOT8A Do you own the land on which the mobile home sits?
(1) yes
(2) no

LVALUE How much do you think the land would sell for on today's market?
(1-999997) \$1-\$999,997
(999998) \$999,998 or more \$

MOVAC How many months has that ... been vacant?
(1-24) NUMBER OF MONTHS (IF 1 to 24)
(00) Less than 1 month
(25) Over 2 years
(26) NEVER OCCUPIED
(D) Don't know

VACANC1 Is INTENDED for year-round use, for occupancy only on a seasonal basis, or for use by migrant workers?
(1) Year round
(8) Seasonal-Summer only
(9) Seasonal-Winter only
(10) Other seasonal (SPECIFY)
(11) Migratory

MOPERM How many months has it been since was occupied as a permanent home?
(1-24) NUMBER OF MONTHS (IF 1 to 24)
(00) Less than 1 month
(25) Over 2 years
(26) NEVER OCCUPIED AS A PERMANENT HOME
(27) Don't know

YRRND Does the construction and heating of make it suitable for year-round use?
(1) Yes
(2) No

TIMSHR Is the ownership of time-shared?
(1) Yes
(2) No

VACANC2 Does the owner/manager intend for this ... to be Is that ...
(1) For rent only?
(2) For rent or for sale?
(3) For sale only?
(4) Rented, but not yet occupied?

- (5) Sold, but not yet occupied?
- (6) Held for occasional use throughout the year?
- (7) Other (specify)

PVALUE How much do you think the ... would sell for on today's market?
 (1-999997) \$1-\$999,997
 (999998) \$999,998 or more \$

VALUE1 (1-999997) \$1-\$999,997
 (999998) \$999,998 or more \$

Last time we interviewed your household the value of the property at which you live was reported to be This time you reported the value to be This represents a substantial change. What is the reason for the large difference?

(ENTER ALL THAT APPLY OR ENTER (N) IF NO OTHERS)

- (1) Major repairs, alterations, additions, or improvements have since been made
- (2) Unit involved in disaster or partial demolition which changed the value
- (3) Sold or purchased land
- (4) Area has been much more developed
- (5) Area has had major diaster
- (6) Changes in economy
- (7) Rezoning
- (8) Previous (1995) response was wrong
- (9) Current (1997) response was wrong
- (10) Other response (SPECIFY)

VALUE2R (1-999997) \$1-\$999,997
 (999998) \$999,998 or more \$

VALUE3 What was the purchase price?
 (1-999997) \$1-\$999,997
 (999998) \$999,998 or more \$

Is this home currently for rent or sale?
 (1) yes
 (2) no

MARKT Is it for (READ ALL CATEGORIES)
 (1) Rent only?
 (2) Rent or for sale?
 (3) Sale only?

MARKET Is the ownership of the ... shared with anyone NOT
 living here?
 (1) yes
 (2) no

NROWNR Does anyone not living here pay some of the mortgage or
 utility costs?
 (1) yes
 (2) no

NRPAYMB

Appendix H

Derivation of AHS Mortgage Market Attributes Tested in the Study

MORTGAGE MARKET MEASURE	AHS SURVEY YEAR	ORIGINATION COHORTS	VARIABLES USED AND SELECTION CRITERIA
Mortgage Origination Volume	1991-2001	1989-2000	Primary mortgages of owner-occupied 1-4 unit properties TENURE - Owner/renter status NUNITS - Number of units in property
	1991-1995	1989-1994	MORT - Whether there is a mortgage on property Mortgage origination year is based on a combination of information from MATBUY, YRMOR, BUILT, WHNRCV, and BUYYR Origination year = YRMOR if the mortgage was not placed/assumed at house acquisition Origination year = BUILT if the mortgage was placed/assumed at house acquisition and the unit was built by owner Origination year = WHNRCV if the mortgage was placed/assumed at acquisition and the house was a gift or an inheritance Origination year = BUYYR if the mortgage was placed/assumed at acquisition and the house was not a gift or built by owner MATBUY - Whether the mortgage was placed/assumed at house purchase YRMOR - Year the mortgage was obtained BUILT - Year the unit was built WHNRCV - Year the household received or inherited the home BUYYR - Year the household purchased this property
	1997-2001	1995-2000	MG - Whether there is a mortgage on property Mortgage origination year is based on a combination of information from MATBUY, YRMOR, and WHNGET Origination year = YRMOR if the mortgage was not placed or assumed at house acquisition Origination year = WHNGET if the mortgage was placed or assumed at house acquisition MATBUY - Whether the mortgage was placed/assumed at house purchase YRMOR - Year the mortgage was originated WHNGET - Year the housing unit was obtained, bought, or received as a gift or inheritance
Loan Purpose (Home Purchase vs. Refinance)	1991-1999	1989-1998	Determination is based on a comparison between mortgage origination year and house acquisition year Loan purpose = home purchase if the year the house acquisition year was the same as the mortgage origination year Loan purpose = refinance if the year the house acquisition year was the same as the mortgage origination year See descriptions above for mortgage origination volume for determination of mortgage origination year Determination of house acquisition year for the 1991, 1993, and 1995 surveys is based on a combination of from BUILT, WHNRCV, and BUYYR BUILT - Year the unit was built WHNRCV - Year the household received or inherited the home BUYYR - Year the household purchased this property Starting from the 1997 survey, year of house acquisition is available in WHNGET
	2001	1999-2000	REFI - Whether the mortgage is a refinance of a previous mortgage For records where REFI is missing, determination is based on a comparison of mortgage origination year and WHNGET

MORTGAGE MARKET MEASURE	AHS SURVEY YEAR	ORIGINATION COHORTS	VARIABLES USED AND SELECTION CRITERIA
Mortgage Market Segment (FHA, VA/FmHA/RHS, Conforming Conventional, Jumbo Conventional)	1991-2001	1989-2000	Combination of information from MORTINS, AMMORT, and GSE conforming loan limits MORTINS - Mortgage insurance type (FHA, VA, FmHA/RHS, or some other type) AMMORT - Amount of mortgage when acquired
Origination Loan Amount	1991-2001	1989-2000	AMMORT - Amount of mortgage when acquired Prior to 1995 survey, if AMMORT has missing values, they are imputed with the sample median
Mortgage Interest Rate	1991-1993 1995 1997-2001	1989-1992 1993-1994 1995-2000	INT - Current interest rate on mortgage Combination of information from INTW and INTF INTW - Current interest rate on mortgage - whole number portion INTF (1/4ths) - Current interest rate on mortgage - fraction number portion INTW - Current interest rate on mortgage - whole number portion INTF (1/8ths) - Current interest rate on mortgage - fraction number portion
LTV at Origination	1991-2001	1989-2000	Combination of information from AMMORT, VALUE, and LPRICE AMMORT - Amount of primary mortgage when acquired Prior to 1995 survey, if AMMORT has missing values, they are imputed with the sample median VALUE - Property value (owner's estimate) LPRICE - Purchase price of unit and land Only included records with VALUE or LPRICE that is at least \$10,000 For refinance mortgages, LTV = AMMORT/VALUE For home-purchase mortgage, LTV = AMMORT/LPRICE
Mortgage Payment Product Type	1991-2001	1989-2000	Combination of information from ARM, TERM, AMRTZ, VARY, FIXED, GPM, GPMW, BLOON, VARM, and a SAS computer program written by the Census Bureau: http://www.huduser.org/datasets/ahs/1999table_recoded.txt ARM - Mortgage payments change based on interest rates TERM - Original term of mortgage AMRTZ - Remaining term of mortgage VARY - Mortgage payments vary during the length of mortgage FIXED - Mortgage payments change due to decline in principal balance, or change in taxes or insurance GPM - Mortgage payments rise at fixed schedule during part of loan GPMW - Mortgage payment rise a fixed schedule during whole length of loan BLOON - Last mortgage payment is the biggest among all payments VARM - Mortgage payments change due to other reasons

MORTGAGE MARKET MEASURE	AHS SURVEY YEAR	ORIGINATION COHORTS	VARIABLES USED AND SELECTION CRITERIA
First-time Homebuyer Status	1991-2001	1989-2000	FRSTHO - Owner of this unit ever owned a home before
Borrower Race/Ethnicity	1991-2001	1989-2000	Combination of information from RACE1 and SPAN1 for the reference person RACE1 - Race of household head/reference person SPAN1 - Hispanic Origin of household head/reference person
Borrower Income	1991-2001	1989-2000	ZINC2 - Total income of household members - the sum of the wage & salary income of all household members age 14+ and all other reported income