

Appendix A: Description of the Data

This data release presents information by year of origination on the dollar amounts, loan counts, and delinquency experience through year-end 2009 of single-family mortgages originated between 2001 and 2008 and financed through the secondary mortgage market. The release presents information separately on loans acquired by Fannie Mae and Freddie Mac (the Enterprises) and mortgages financed through the issuance of private-label mortgage-backed securities and asset-backed securities (collectively, private-label MBS). The data used to prepare the analysis include conventional fixed- and adjustable-rate mortgages secured by first liens on owner-occupied and investor-owned one-to-four-family properties.

The data come from two distinct sources. FHFA's Historical Loan Performance dataset contains loan-level information on the characteristics and performance of all single-family mortgages acquired (that is, purchased for cash or financed with mortgage-backed securities guaranteed) by the Enterprises.⁴ FHFA updates the Historical Loan Performance dataset quarterly with information from each Enterprise. The data summarized in this release reflects the information provided by the Enterprises for the first quarter of 2010. The loans in the Historical Loan Performance dataset originated from 2001 through 2008 and had an aggregate unpaid principal balance at origination of \$8.6 trillion.

The LoanPerformance mortgage securities dataset, which FHFA leases from CoreLogic, Inc., contains loan-level information on the characteristics and performance of single-family mortgages financed with private-label MBS.⁵ CoreLogic updates the LoanPerformance dataset monthly. The data summarized in this release reflects information provided by CoreLogic in April 2010. The mortgages in the LoanPerformance dataset originated from 2001 through 2008 and had an aggregate unpaid principal balance at origination of \$1.8 trillion before weighting and \$2 trillion after weighting.

Each dataset contains information on factors that lenders use to assess mortgage credit risk at origination and information on subsequent loan performance. Risk factors include payment type (that is, fixed rate or adjustable rate), initial loan-to-value (LTV) ratio based on the purchase price

⁴ The Historical Loan Performance dataset does not include loans backing private-label MBS bought by the Enterprises.

⁵ FHFA leases two LoanPerformance securities databases that contain loan-level information on mortgages backing private-label MBS. The LoanPerformance dataset used to prepare this release includes data derived from both of those databases.



or appraised property value and the first-lien balance,⁶ and credit score(s) for the borrower(s) calculated using models developed by Fair Isaac Corporation (FICO). The analysis summarized in this release targets first-lien mortgages on one-unit properties not insured or guaranteed by an agency of the federal government that were originated from 2001 through 2008 and subsequently acquired by the Enterprises or financed by the issuance by private securitizers of private-label MBS.⁷ The sample excludes balloon and reverse mortgages as well as all loans acquired by the Enterprises or financed with private-label MBS in the period but originated before 2001.

Missing Data

Mortgages missing either FICO scores or LTV ratios comprise 0.7 percent of the principal of all mortgages in the datasets. The percentage of loans by principal balance with missing data is about 15 percent higher for Enterprise-acquired loans in the Historical Loan Performance dataset than for private-label MBS-financed loans in the LoanPerformance dataset. In the Historical Loan Performance dataset, 0.7 percent of adjustable-rate mortgages (ARMs) and 1.22 percent of fixed-rate mortgages (FRMs) are missing data, but in the LoanPerformance dataset, 0.9 percent of ARMs and 0.56 percent of FRMs are missing data.

Dataset Coverage and Sample Weighting

The Historical Loan Performance dataset contains nearly complete coverage of single-family mortgages financed by Fannie Mae and Freddie Mac, but the LoanPerformance dataset comprises only a sample of single-family mortgages financed with private-label MBS. To make meaningful comparisons between those two funding channels, we adjusted the LoanPerformance data so that the total dollar volume of originations in each year of the period from 2001 through 2008 is comparable to the universe of mortgages originated in those years and financed with private-label MBS issued from 2001 through 2009. We use estimates of private-label MBS issuance in that period compiled by Inside Mortgage Finance (IMF) Publications, Inc. Because the LoanPerformance dataset includes information on first liens that issuers

⁶ This measure of LTV ratio does not capture any second liens secured by the property that were present at origination, so loans with equal reported LTV ratios may pose significantly different credit risk.

⁷ This universe of loans was targeted as well as the data allow. For example, we could not accurately identify the lien status of many Enterprise-acquired mortgages in the Historical Loan Performance dataset. Because most loans acquired by the Enterprises are first liens, we treat as first liens all mortgages identified as “unknown” in the Historical Loan Performance dataset. In addition, neither dataset identifies mortgages that finance one-unit properties. To target such properties, we include all mortgages associated with one- to four-unit properties with an initial loan balance at or below the conforming limit for one-unit properties that was in effect during the origination year and in the area in which the loan was originated.



characterized as of prime, subprime, and Alternative-A (Alt-A) credit quality that had initial balances below and above the conforming loan limit, as well as information on closed-end second liens, the relevant universe used for scaling the LoanPerformance dataset is the sum of mortgages backing those categories of private-label MBS issues reported by Inside Mortgage Finance Publications, Inc.

The LoanPerformance dataset's total coverage of the universe of loans financed with private-label MBS in each year is estimated as the total dollar volume of mortgages in the dataset that back securities issued during the given year, divided by the estimate of issuance during the year made by Inside Mortgage Finance Publications, Inc. To scale the LoanPerformance data to that of the universe estimated by Inside Mortgage Finance Publications, Inc. and make it more comparable to the universe of Enterprise-acquired loans in the Historical Loan Performance dataset, the balance of each loan in LoanPerformance is multiplied by the scale factor in the year in which the loan was securitized.

Appendix Table A-1 presents the data and calculations used to determine the LoanPerformance coverage ratio and scale factor used for weighting each loan in the LoanPerformance dataset. The weighted LoanPerformance data are aggregated by origination year to produce tables 1 through 3 in the body of this release and appendix tables B-1 through B-3.

Appendix Table A-1. Basis for Weighting Mortgages in the Loan Performance Dataset

Issuance Year	(a) Private-Label Mortgage Securities Issued¹ (\$ in Billions)	(b) Mortgages in Loan Performance Dataset² (\$ in Billions)	(c) = (b)/(a) Coverage Ratio of Loan Performance Dataset (Percent)	(d) = (a)/(b) Scale Factor Used for Weighting (Percent)
2001	\$256.1	\$221.1	86.3%	115.8%
2002	\$372.5	\$325.0	87.2%	114.6%
2003	\$526.9	\$500.6	95.0%	105.3%
2004	\$803.6	\$720.0	89.6%	111.6%
2005	\$1,138.8	\$1,041.0	91.4%	109.4%
2006	\$1,107.6	\$1,046.8	94.5%	105.8%
2007	\$664.9	\$624.4	93.9%	106.5%
2008	\$10.8	\$9.7	89.6%	111.6%

Source: FHFA (using data from Inside Mortgage Finance Publications, Inc. and CoreLogic, Inc.)

¹ Per Inside Mortgage Finance Publications, Inc. Includes securities backed by loans designated as prime, subprime, Alt-A, and second mortgages

² Unpaid principal balance at origination of all mortgages in the LoanPerformance dataset, aggregated by issuance year of the securities with which the loans were financed



Calculation of Ever-90-Day Delinquency Rates

Table 3 and appendix tables B-3 and B-6 show ever 90-day delinquency rates. We compute those rates as the proportion of the original principal balance of a group of mortgages that have ever been 90 days past due, in foreclosure processing, or in real estate owned (REO) status through December 2009.

The information available in the LoanPerformance and Historical Loan Performance datasets do not allow us to compute ever 90-day delinquency in exactly the same way across the two datasets. LoanPerformance reports monthly performance data for every loan in the dataset and identifies each loan as current, prepaid, in foreclosure, REO, or past due in each month using the Mortgage Bankers Association's definition of delinquency.⁸ So, for the LoanPerformance dataset, we identify a loan as 90-plus days delinquent if, for any month through December 2009, its status is 90 days or more delinquent, in foreclosure processing, or REO.

The Historical Loan Performance dataset contains monthly information on the number of days each delinquent mortgage is past due and whether loans are in foreclosure processing. However, mortgages are identified as being in REO only at the end of each quarter. We identify a loan in Historical Loan Performance as having been ever 90-days delinquent if it is 90 or more days past due or in foreclosure in any month through December 2009, or if it is identified as being REO in any quarter through the fourth quarter of 2009.

The ever 90-day delinquency rates shown in Table 3 and appendix tables B-3 and B-6 should be interpreted with caution. Relative comparisons of those rates are likely to be most revealing within origination years for two reasons. First, because ever 90-day delinquency rates can only increase as time passes, rates for identical groups of mortgages that finance properties that experienced identical economic conditions after origination but were originated in different years will be different. Second, low interest rates and rapid house price appreciation allowed many borrowers who took out loans early in the 2001-2008 period to refinance their loans, reducing the ever 90-day delinquency rates of mortgages in those origination years relative to those loans taken out later in the period. In addition, the ever 90-day delinquency rates of groups of loans originated in any year may not correspond to their ultimate default rates and will differ from the rates of loss on the loans.

⁸ The Mortgage Bankers Association considers a loan delinquent if the payment is not received by the day before the next payment is due. Thus, a loan would be 30 days delinquent if, for example, the March 1 installment had not been paid as of March 31. Similarly, a loan would be 60- or 90-days delinquent if, for example, the February 1 or January 1 installment was unpaid as of March 31.



Treatment of FICO Credit Scores

Loans in the Historical Loan Performance dataset may have multiple FICO credit scores. First, a loan may have multiple borrowers. Second, each borrower may have multiple credit scores for a given date collected from different sources, such as each of the major credit reporting agencies. Third, each Enterprise monitors over time the credit scores of borrowers whose loans it has acquired. To select a credit score for the analysis reported here, we first considered only those scores closest to the origination date. Of those, we chose the lowest credit score associated with the primary borrower. In choosing the lowest credit score, we were conservative. In contrast, the LoanPerformance dataset contains just one credit score for each loan, which LoanPerformance describes as the “FICO score at origination.” In both datasets, out-of-range scores (those below 300 or above 850) are treated as missing.

Determining Origination Years

The Historical Loan Performance dataset contains the origination date of each loan, and that date is used to determine the origination year of Enterprise-acquired mortgages. The LoanPerformance dataset contains two variables that can be used to identify the origination year—the origination date and the first payment date. For most loans, these data are reasonably consistent and yield the same origination years. However, for a significant subset of loans, the years recorded for origination and first payment are different, sometimes by as much as eight years.

This data release uses origination date rather than first payment date to determine origination year. The effects of that decision relative to determining origination year using first payment data are presented in Appendix Table A-2. They include reducing the total number of loans in our study by about 125,000 and the total unpaid principal balance of those loans by approximately \$29 billion, or about 1.5 percent. In addition, the numbers of loans and their unpaid principal balances are higher from 2001 through 2004 and lower from 2005 through 2008. The overall ever 90-day delinquency rate across all years remains unchanged, but the pattern of delinquencies is shifted across years.



Appendix Table A-2: Comparison of Annual and Total Loan Counts, Unpaid Principal Balances, and Ever-90-Day Delinquency Rates for Loans Financed with Private-Label Securities When Origination Year is Determined by Origination Date or by First Payment Date

Loan Counts									
	2001	2002	2003	2004	2005	2006	2007	2008	All Years
Origination Date	517,358	837,221	1,490,041	2,449,189	2,939,734	2,132,378	485,778	592	10,852,291
First Payment	496,547	774,529	1,371,152	2,287,643	2,994,084	2,280,808	769,249	3,270	10,977,282
Difference	(20,811)	(62,692)	(118,889)	(161,546)	54,350	148,430	283,471	2,678	124,991

Unpaid Principle Balance (\$ in Billions)									
	2001	2002	2003	2004	2005	2006	2007	2008	All Years
Origination Date	\$66.6	\$121.5	\$234.4	\$432.7	\$553.8	\$430.3	\$104.3	\$0.3	\$1,943.9
First Payment	\$61.6	\$112.1	\$212.4	\$400.1	\$563.1	\$458.1	\$163.6	\$1.8	\$1,972.6
Difference	-\$5.0	-\$9.4	-\$22.1	-\$32.6	\$9.3	\$27.8	\$59.3	\$1.5	\$28.7

Ever-90-day Delinquency Rates									
	2001	2002	2003	2004	2005	2006	2007	2008	All Years
Origination Date	20.4%	15.1%	11.8%	15.1%	28.6%	45.1%	42.2%	14.5%	26.8%
First Payment	21.2%	15.7%	11.7%	14.2%	25.8%	43.4%	44.2%	13.5%	26.8%
Difference	0.9%	0.7%	-0.1%	-0.9%	-2.8%	-1.7%	1.9%	-1.0%	0.0%

Source: FHFA (using data from CoreLogic, Inc.)

