Appendix E - Data Transformation

HUD provided Deloitte & Touche with data from the Single Family Data Warehouse for fiscal endorsement years 1975 through 2000 as of June 30, 2000. The following summarizes the process of summarizing the data and preparing the data sets for analysis.

Initial Record Drop Criteria

Our first step in sorting through the data was to take out any files that did not have an original loan amount (orig_mrtg_amt = 0) or a contract rate (int_rt = 0)[1]. The following table summarizes the results of this process.

Table E.1

Fiscal Origination Year	Original Number of Loans in Database	Total Initial Drop	Number Remaining Loans After Initial Drop	Percent of Total Original Loans
1975	185,986	35	185,951	0.019%
1976	222,112	51	222,061	0.023%
1977	256,148	82	256,066	0.032%
1978	294,587	152	294,435	0.052%
1979	389,790	857	388,933	0.220%
1980	337,139	545	336,594	0.162%
1981	216,298	246	216,052	0.114%
1982	149,182	7,616	141,566	5.105%
1983	506,090	119	505,971	0.024%
1984	287,195	19	287,176	0.007%
1985	400,634	11	400,623	0.003%
1986	928,984	30	928,954	0.003%
1987	1,126,833	41	1,126,792	0.004%
1988	615,823	148	615,675	0.024%
1989	634,639	109	634,530	0.017%
1990	715,752	52	715,700	0.007%
1991	643,556	62	643,494	0.010%
1992	637,160	70	637,090	0.011%
1993	991,008	63	990,945	0.006%
1994	1,059,809	71	1,059,738	0.007%
1995	521,218	145	521,073	0.028%
1996	728,022	51	727,971	0.007%
1997	740,127	59	740,068	0.008%
1998	954,987	47	954,940	0.005%
1999	1,116,160	61	1,116,099	0.005%
2000	533,086	67	533,019	0.013%
Total	**15,192,325**	**10,809**	**15,181,516**	**0.071%**

[1] Program filename: /HUD/program/frstdrop.sas and /HUD/program/scnddrop.sas

Identifying Loan Types

We split the database into six different loan types:[2]

1. Fixed rate 30-year (FX30)
2. Fixed rate 15-year (FX15)
3. Adjustable rate (ARM)
4. Streamline refinance 30-year (SRFX30)
5. Streamline refinance 15-year (SRFX15)
6. Adjustable rate streamline refinance (SRARM)

We identified Streamline Refinanced (SR) loans in fiscal origination years 1988 through 2000 according to three criteria:

1. A refinance code (rfnc_cd) of "H", "R", or "S"
2. A streamline flag (pd_strmln_flg) of "R", or
3. A loan-to-value ratio (ratio_loan_to_vl) coded as 30 or 999 (as opposed to our calculated value of LTV).

We used the adjustable rate indicator and the 15-year term indicator in the Data Warehouse to further classify the loans.

Geography

There are some geographic areas covered by the MMIF but for which some of the external economic information was unavailable. These are, specifically: Puerto Rico, the Virgin Islands, and Guam. Since we did not have complete information about these areas, we had to make simplifying assumptions. Given the small size of this subset of the database (see table below), we believe the assumptions to have an immaterial effect on our results.

[2] Program filename: /HUD/program/loantype.sas

We used economic information about Florida as a proxy for information about Puerto Rico. We excluded Virgin Island and Guam records from the regression analysis.

Table E.2

Fiscal Origination Year	Number of Records in Analysis	Puerto Rico	Puerto Rico Percentage	Virgin Islands, Guam	Virgin Islands, Guam Percentage	Number of Records Remaining in Analysis
1975	185,951	1,513	0.814%	436	0.234%	185,515
1976	222,061	1,755	0.790%	177	0.080%	221,884
1977	256,066	2,594	1.013%	214	0.084%	255,852
1978	294,435	3,753	1.275%	168	0.057%	294,267
1979	388,933	2,660	0.684%	56	0.014%	388,877
1980	336,594	2,924	0.869%	26	0.008%	336,568
1981	216,052	1,706	0.790%	2	0.001%	216,050
1982	141,566	1,903	1.344%	71	0.050%	141,495
1983	505,971	1,918	0.379%	115	0.023%	505,856
1984	287,176	3,178	1.107%	111	0.039%	287,065
1985	400,623	3,697	0.923%	42	0.010%	400,581
1986	928,954	5,753	0.619%	31	0.003%	928,923
1987	1,126,792	8,903	0.790%	43	0.004%	1,126,749
1988	615,675	8,422	1.368%	28	0.005%	615,647
1989	634,530	8,559	1.349%	27	0.004%	634,503
1990	715,700	8,810	1.231%	50	0.007%	715,650
1991	643,494	7,779	1.209%	28	0.004%	643,466
1992	637,090	6,782	1.065%	64	0.010%	637,026
1993	990,945	6,231	0.629%	82	0.008%	990,863
1994	1,059,738	7,708	0.727%	64	0.006%	1,059,674
1995	521,073	7,941	1.524%	25	0.005%	521,048
1996	727,971	9,283	1.275%	33	0.005%	727,938
1997	740,068	10,753	1.453%	65	0.009%	740,003
1998	954,940	9,915	1.038%	50	0.005%	954,890
1999	1,116,099	11,303	1.013%	40	0.004%	1,116,059
2000	533,019	5,964	1.119%	19	0.004%	533,000
Total	**15,181,516**	**151,707**	**0.999%**	**2,067**	**0.014%**	**15,179,449**

Loan-to-Value Ratio Calculation

In general, the initial loan-to-value ratio, LTV_0, is calculated using the following formula:

$$\frac{orig\_mrtg\_amt - ufmip\_pd\_amt}{\min(prprty\_aprsl\_vl, prc\_excl\_clsng\_amt)}$$

1. **If both prprty_aprsl_vl and prc_excl_clsng_amt are available**, the LTV_0 ratio is estimated based on the above formula.
2. **If one of prprty_aprsl_vl or prc_excl_clsng_amt is not available**, the LTV_0 ratio's denominator takes the value of the available variable.
3. **If both "previous" prprty_aprsl_vl and "previous" prc_excl_clsng_amt are unavailable**, then we use the ratio_loan_to_vl field in the database.
4. **If ratio_loan_to_vl is unavailable**, then the loan record is excluded from the regression analysis for lack of sufficient information.

Streamline Refinanced Loans

Matching to Original Loan

Because Streamline Refinancing doesn't require an appraisal, we needed to estimate LTV_0 for those loans. We did this by attempting to match each SR loan to the refinanced or "previous" loan. We searched all loans prior to each SR loan for a loan where the refinance case number field (rfnc_cs_nbr) matched the case number of the SR loan.

We were able to match roughly 85% of the SR loans to their "previous" loans. The success rate varied by fiscal origination year as shown in the table below.

Table E.3

Origination Year	Streamline Refinancings	Unmatched	Total Streamline Refinancings Remaining	Percent Unmatched
1988	21,547	19,874	1,374	92%
1989	13,497	8,536	4,765	63%
1990	25,255	10,554	14,641	42%
1991	29,085	10,609	18,785	36%
1992	97,393	21,643	80,714	22%
1993	421,398	42,659	388,923	10%
1994	458,543	76,983	388,773	17%
1995	28,127	14,407	13,524	51%
1996	102,854	22,540	81,130	22%
1997	56,167	11,300	45,110	20%
1998	212,659	28,807	184,350	14%
1999	259,244	32,860	226,928	13%
2000	24,254	5,699	18,586	23%
Total	**1,750,023**	**306,471**	**1,467,603**	**18%**

If we could not match an SR loan to an earlier loan record, we dropped the SR from our regression analysis. Note that, if the "previous" loan had already been dropped from the analysis

for lack of sufficient information to calculate LTV_0, then the corresponding SR loan is included in the count of "unmatched" loans.

Estimation of Property Value
Depending on the data available from the "previous" loan, we can estimate the property value of the SR loan based on one of the following scenarios:[3]

1. **If both "previous" prprty_aprsl_vl and "previous" prc_excl_clsng_amt are available**, the SR loan's property value is estimated as the minimum of these two values adjusted by the ratio_loan_to_vl. (Note: If the previous ufmip_pd_amt is unavailable in this scenario, we adjust the estimated property value by an upfront premium factor, based on the upfront premium table shown in Appendix D - The Cash Flow Model).
2. **If exactly one of "previous" prprty_aprsl_vl or "previous" prc_excl_clsng_amt is not available**, the other is assigned as the SR loan's estimated property value.
3. **If both "previous" prprty_aprsl_vl and "previous" prc_excl_clsng_amt are unavailable,** we use the ratio_loan_to_vl field in the database.
4. **If ratio_loan_to_vl is unavailable,** then the SR loan is discarded for lack of sufficient information.

Note that the three scenarios parallel the LTV_0 calculation described in the description of the LTV_0 calculation in the previous section.

Once we have estimated the property value based on the available information from the "previous" loan, it is then adjusted by a house price appreciation factor. These factors were derived from the house price index (HPI) published by OFHEO by MSA, by state and by census division.

Payment to Income Fix Subroutine

Analyzing the payment to income ratio in the database (ratio_tmp_tei), we have found that a number of records contain a value of zero in this field. Therefore, we replaced the zero values with a reasonable estimate for the ratio, loan by loan.[4] **For each loan type and each fiscal year**, we followed three simple steps to fix the records containing a zero value in this field:

1. Find all the loans where the ratio_tmp_tei field contains a non-zero value.
2. Calculate a weighted average of ratio_tmp_tei using the non-zero ratios determined in item1 with weights based on the corresponding orig_mrtg_amt.
3. Replace the zero values for ratio_tmp_tei with this weighted average ratio.

The table below shows the calculated average payment-to-income ratio by year and by loan type.

[3] Program Filename: /HUD/program/sr_aprsl.sas
[4] Program Filename: /HUD/program/r_tmptei.sas

Table E.4

Fiscal Origination Year	Average Payment-to-Income Ratio (%)					
	Fixed Rate, 30-year Loans	Fixed Rate, 15-year Loans	Adjustable Rate Loans	Streamline Fixed Rate, 30-year Loans	Streamline Fixed Rate, 15-year Loans	Streamline Adjustable Rate Loans
1975	20.1511	17.2079	N/A	N/A	N/A	N/A
1976	20.3891	17.2068	N/A	N/A	N/A	N/A
1977	20.2267	16.9051	N/A	N/A	N/A	N/A
1978	21.5974	17.0854	N/A	N/A	N/A	N/A
1979	22.2422	17.2330	N/A	N/A	N/A	N/A
1980	23.3820	18.5238	N/A	N/A	N/A	N/A
1981	24.4836	19.3648	N/A	N/A	N/A	N/A
1982	24.7050	20.6668	N/A	N/A	N/A	N/A
1983	23.4249	22.9526	N/A	N/A	N/A	N/A
1984	24.1687	22.8545	N/A	N/A	N/A	N/A
1985	23.2641	22.8489	22.8935	N/A	N/A	N/A
1986	21.4678	20.4746	21.8601	N/A	N/A	N/A
1987	21.3340	19.7993	21.5152	N/A	N/A	N/A
1988	23.3420	22.4618	23.0469	22.1870	21.6939	22.2386
1989	25.3144	23.4088	25.4874	23.1435	19.6855	25.7244
1990	23.7710	21.7246	23.2304	24.9409	21.2093	
1991	22.9586	20.9526	23.8647	25.4199	22.4616	23.2330
1992	22.7206	20.0839	23.4342	23.7293	22.2075	22.3007
1993	22.4510	19.5223	23.6768	23.9142	21.6048	23.5660
1994	22.8193	19.3276	24.1883	21.4928	20.6469	21.5072
1995	23.9851	20.1807	24.8910	23.9238	21.8267	23.8350
1996	24.0224	20.5012	24.9596	24.3507	21.2986	24.1403
1997	24.3540	21.0764	24.9655	25.5657	22.1886	25.4720
1998	24.2689	21.1922	25.0506	29.0337	22.3627	27.4633
1999	25.0280	21.9188	26.1987	25.1613	21.3394	27.1458
2000	26.8307	23.7212	27.2908	27.7689	23.5373	26.8987

Reasonable Range of LTV_0

We further attempted to remove erroneous records from the data set for regression analysis by checking the calculated LTV_0. We excluded any loan where LTV_0 was less than or equal to 10%, and any loan where LTV_0 was greater than or equal to 140%. The results of this step are summarized for fixed rate, 30-year loans in the table below.

Table E.5

Origination Year	Number of Loans, All Loan Types	LTV 10% or Less	LTV 140% or Greater	Remaining Loans	Percent Excluded
1975	185,515	26,512	431	158,572	15%
1976	221,884	28,291	662	192,931	13%
1977	255,852	24,169	950	230,733	10%
1978	294,267	41,259	1,623	251,385	15%
1979	388,877	67,176	1,855	319,846	18%
1980	336,568	36,780	2,135	297,653	12%
1981	216,050	47,196	1,603	167,251	23%
1982	141,495	20,769	725	120,001	15%
1983	505,856	88,591	1,043	416,222	18%
1984	287,065	8,021	599	278,445	3%
1985	400,581	4,434	7,358	388,789	3%
1986	928,923	5,076	4,101	919,746	1%
1987	1,126,749	2,418	2,812	1,121,519	0%
1988	595,773	307	2,600	592,866	0%
1989	625,967	1,560	1,607	622,800	1%
1990	705,096	196	2,454	702,446	0%
1991	632,857	5,881	1,546	625,430	1%
1992	615,383	3,907	5,418	606,058	2%
1993	948,204	42	8,044	940,118	1%
1994	982,691	31	8,046	974,614	1%
1995	506,641	20	4,194	502,427	1%
1996	705,398	12	6,625	698,761	1%
1997	728,703	10	7,418	721,275	1%
1998	926,083	34	9,940	916,109	1%
1999	1,083,199	10	11,229	1,071,960	1%
2000	527,301	2	3,369	523,930	1%
Total	**14,872,978**	**412,704**	**98,387**	**14,361,887**	**3%**

Relative House Price

HUD provided us with median house prices (MHP) through 1997 for some MSAs, and for all states. We estimated MHPs for 1998-2000 based on changes in HPI.

We calculated the relative house price (*RHP*) for a given loan to be consistent with our calculation of LTV_0. For each loan,

$$RHP = \frac{orig\_mrtg\_amt - ufmip\_pd\_amt}{LTV_0} \cdot \frac{1}{MHP}.$$

This guarantees that the "price" used in the RHP calculation for each loan was the same as the property value used to calculate the loan-to-value ratio. We used the MHP by MSA where it was available; otherwise we used MHP by state.

RHP and LTV Categories

Table E.6

LTV Range		Percentage of Loans in Range	Cumulative Percentage	
10%	15%	0.0073%	0.0073%	
15%	20%	0.0073%	0.0145%	
20%	25%	0.0170%	0.0315%	
25%	30%	0.0334%	0.0649%	
30%	35%	0.1134%	0.1783%	
35%	40%	0.1082%	0.2865%	
40%	45%	0.1564%	0.4429%	
45%	50%	0.2306%	0.6735%	
50%	55%	0.3328%	1.0063%	
55%	60%	0.4482%	1.4545%	
60%	65%	0.6417%	2.0962%	Low
65%	70%	0.9378%	3.0340%	
70%	75%	1.8147%	4.8487%	
75%	76%	0.4200%	5.2686%	
76%	77%	0.4671%	5.7358%	
77%	78%	0.5385%	6.2743%	
78%	79%	0.5429%	6.8172%	
79%	80%	0.7442%	7.5614%	
80%	81%	0.8004%	8.3618%	
81%	82%	0.8633%	9.2251%	
82%	83%	0.8622%	10.0872%	
83%	84%	0.9248%	11.0120%	
84%	85%	1.9703%	12.9823%	Investors
85%	86%	1.0161%	13.9984%	
86%	87%	1.2151%	15.2135%	
87%	88%	1.3237%	16.5372%	
88%	89%	1.4539%	17.9910%	
89%	90%	2.3021%	20.2932%	
90%	91%	1.8492%	22.1423%	
91%	92%	2.3805%	24.5228%	Mid
92%	93%	2.9257%	27.4485%	
93%	94%	3.7430%	31.1915%	
94%	95%	5.9437%	37.1352%	
95%	96%	11.0768%	48.2120%	
96%	97%	17.6611%	65.8731%	
97%	98%	15.9449%	81.8180%	
98%	99%	6.4814%	88.2994%	
99%	100%	8.3829%	96.6823%	
100%	101%	0.7842%	97.4665%	
101%	102%	0.4056%	97.8720%	
102%	103%	0.2430%	98.1151%	High
103%	104%	0.1485%	98.2636%	
104%	105%	0.0640%	98.3276%	
105%	110%	0.3405%	98.6681%	
110%	115%	0.1998%	98.8679%	
115%	120%	0.1558%	99.0237%	
120%	125%	0.1300%	99.1537%	
125%	130%	0.8463%	100.0000%	

Table E.7

	RHP Range		Percentage of Loans in Range	Cumulative Percentage	RHP Range		Percentage of Loans in Range	Cumulative Percentage	
	0%	10%	0.1025%	0.1025%	100%	101%	1.0669%	67.5465%	
	10%	20%	0.1551%	0.2576%	101%	102%	1.0352%	68.5817%	
	20%	30%	0.8741%	1.1317%	102%	103%	1.0411%	69.6228%	
	30%	40%	2.8404%	3.9721%	103%	104%	0.9912%	70.6140%	
	40%	50%	5.6306%	9.6026%	104%	105%	1.0063%	71.6203%	
	50%	60%	9.1307%	18.7333%	105%	106%	0.9849%	72.6052%	
	60%	61%	1.0200%	19.7533%	106%	107%	0.9283%	73.5335%	
	61%	62%	1.0593%	20.8126%	107%	108%	0.9272%	74.4607%	
	62%	63%	1.0651%	21.8777%	108%	109%	0.8970%	75.3577%	
Low	63%	64%	1.1066%	22.9842%	109%	110%	0.8743%	76.2321%	
	64%	65%	1.1210%	24.1052%	110%	111%	0.8682%	77.1003%	
	65%	66%	1.1236%	25.2288%	111%	112%	0.8346%	77.9348%	
	66%	67%	1.1509%	26.3797%	112%	113%	0.8157%	78.7506%	
	67%	68%	1.1407%	27.5204%	113%	114%	0.7791%	79.5297%	
	68%	69%	1.2027%	28.7231%	114%	115%	0.7596%	80.2893%	
	69%	70%	1.2057%	29.9288%	115%	116%	0.7476%	81.0369%	
	70%	71%	1.1832%	31.1120%	116%	117%	0.7066%	81.7435%	
	71%	72%	1.2180%	32.3300%	117%	118%	0.7130%	82.4565%	
	72%	73%	1.2105%	33.5405%	118%	119%	0.6783%	83.1348%	
	73%	74%	1.2546%	34.7951%	119%	120%	0.3276%	83.4625%	High
	74%	75%	1.2458%	36.0409%	120%	130%	5.5114%	88.9738%	
	75%	76%	1.2488%	37.2898%	130%	140%	3.7097%	92.6835%	
	76%	77%	1.2512%	38.5409%	140%	150%	2.4547%	95.1382%	
	77%	78%	1.2662%	39.8071%	150%	160%	1.6266%	96.7648%	
	78%	79%	1.2949%	41.1020%	160%	170%	1.0510%	97.8157%	
	79%	80%	1.2884%	42.3905%	170%	180%	0.6863%	98.5021%	
	80%	81%	1.2551%	43.6456%	180%	190%	0.4572%	98.9593%	
	81%	82%	1.2555%	44.9010%	190%	200%	0.2995%	99.2588%	
	82%	83%	1.2888%	46.1898%	200%	210%	0.2044%	99.4632%	
	83%	84%	1.2632%	47.4530%	210%	220%	0.1395%	99.6027%	
	84%	85%	1.2603%	48.7133%	220%	230%	0.0999%	99.7026%	
	85%	86%	1.2612%	49.9746%	230%	240%	0.0737%	99.7763%	
Mid	86%	87%	1.2716%	51.2461%	240%	250%	0.0536%	99.8299%	
	87%	88%	1.2352%	52.4814%	250%	260%	0.0386%	99.8685%	
	88%	89%	1.2370%	53.7184%	260%	270%	0.0280%	99.8965%	
	89%	90%	1.2188%	54.9372%	270%	280%	0.0200%	99.9164%	
	90%	91%	1.2124%	56.1496%	280%	290%	0.0159%	99.9323%	
	91%	92%	1.2077%	57.3573%	290%	300%	0.0000%	99.9323%	
	92%	93%	1.2059%	58.5632%	300%	300%	0.0677%	100.0000%	
	93%	94%	1.1627%	59.7259%			100.0000%		
	94%	95%	1.1472%	60.8731%					
	95%	96%	1.1559%	62.0290%					
	96%	97%	1.1379%	63.1668%					
	97%	98%	1.1087%	64.2756%					
	98%	99%	1.1073%	65.3828%					
	99%	100%	1.0967%	66.4796%					

The two previous tables illustrate the distribution of loans (across all loan types) by LTV ratio and by RHP ratio, respectively. (The calculation of each of these ratios for individual loans was described above.) Our definition of the LTV and RHP ranges was based on examination of these tables.

We further subdivided the LTV categories into increments for purposes of accuracy. In particular, the calculation of the probability of negative equity for a "cell" of loans requires a finer definition of the LTV range. The table below shows the definitions of the LTV increments, as well as the value for each increment that we used as a proxy for each value within the range in calculating the probability of negative equity.

Table E.8

LTV Category	Proxy Value	Incremental Range	
Low	77.5%	0%	80%
	81.5%	80%	83%
Investor	84%	0%	85%
	86%	85%	87%
	90%	87%	140%
Mid	88.5%	87%	90%
	91%	90%	92%
	93%	92%	94%
	95%	94%	96%
High	97%	96%	98%
	99%	98%	100%
	105%	100%	140%

Age

Throughout this document, we will refer to the age of a pool of loans in terms of time t or policy year. In each case, we are defining the age of the pool of loans in terms of the number of years since the inception of the fiscal origination year (or endorsement year, if applicable). Therefore, policy year 1 for fiscal origination year 1985 is the time period between the inception of the period, October 1, 1984, and the date one year later, October 1, 1985. Fiscal origination year 1998 will reach age 4 ($t = 4$) on October 1, 2001.

Unemployment

We used a time series of historical countrywide unemployment rates. We did not incorporate a lag into the variable, as we did in the previous Actuarial Review. The unemployment rate associated with a given observation "cell" for loans originating in fiscal year 1980, for example, at policy year 5 is the countrywide unemployment rate for fiscal year 1984.

Time-adjusted Loan-to-Value Ratio (LTV_t)

We calculated LTV_t by individual loan. The time variable, t, represents the age of the fiscal origination year, where t = 1 represents the end of the fiscal year itself, t = 2 is the date one year later, and so on. Therefore, LTV_t is evaluated for a given loan as of October 1 of the fiscal year, plus t years, minus 1 (or as of 10/31/[FY + t – 1]).

$$LTV_t = LTV_0 \cdot \frac{SAF_t}{HPAF_t}, \text{ where}$$

$HPAF_t = \frac{HPI_t}{HPI_0}$, an adjustment for change in house prices between the time of the origination of the loan and the age t, and SAF_t is the scheduled amortization factor, or the percentage of the original loan amount estimated as still outstanding at age t.

Time-adjusted Payment-to-Income Ratio ($PAY.INC_t$)

$$PAY.INC_t = PAY.INC_0 \cdot \frac{contractrate_t}{contractrate_0} \cdot \frac{personalincome_0}{personalincome_t}$$

We obtained personal income per capita by MSA through 1998, and by state through the first quarter of 2000, from the Bureau of Economic Analysis (BEA) website. The BEA data was supplemented with population data from the Census Bureau in order to estimate per capita personal income by MSA for the most recent years.

The adjustment for change in personal income levels were made loan by loan. We made the adjustment for changes in the contract rate for groups of loans. The contract rate changes between time t and time 0 only on adjustable rate loans. The adjusted rate is estimated for a group of loans based on the historical changes in the index for adjustable rate loans, the 1-year, constant maturity T-bill rate. We also assumed that, on average, MMIF loans originated on April 15, which accounts for the seasonality in MMIF originations.

Refinance Incentive Ratio and Related Values

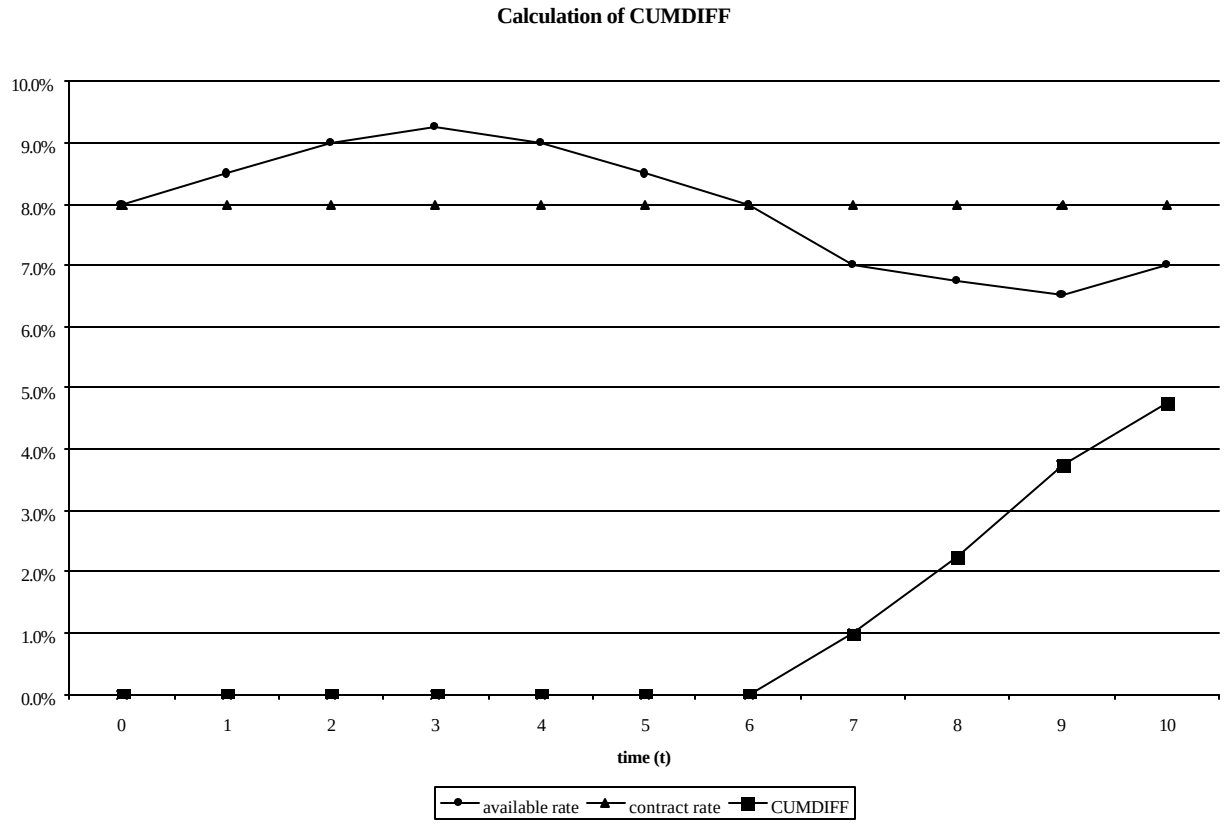
The refinance incentive ratio at a given time t, R_t, is defined as the ratio of the contract rate on a given loan to the available refinance rate at time t. If R_t is greater than one, the contract rate is higher than currently available rates at time t, and refinancing is an attractive prospect. A refinance incentive ratio less than one would imply little or no incentive to refinance at time t.

The variable used to indicate the level of the propensity to refinance is the exponentially weighted, moving average refinance incentive ratio at age t, or R'_t. $R'_t = z \cdot \overline{R}_t + (1-z) \cdot R'_{t-1}$,

where $\overline{R}_t$ = the arithmetic mean of prior refinance incentive ratios up to time t, and z = the weight assigned to prior refinance incentive ratios. For this Review, we selected $z = 0.75$.

The variable $CUMDIFF_t$ and the age of the loan pool determine the degree to which the pool has burned out. $CUMDIFF_t$ is defined as the cumulative positive difference between the loan interest rate and the historically available refinance interest rate. The graph below illustrates this definition for the case of a loan with a fixed rate of 8 percent.

Chart E.9



As long as the available (refinance) rates are higher that the contract rate, there is no incentive to refinance and $CUMDIFF_t$ is zero. As the rates drop below the contract rate, however, there is incentive to refinance. As the positive differences accumulate, there will be very few borrowers left who will prepay and the pool "burns out".

In this Review, we calculated R_t, R'_t, and $CUMDIFF_t$, at the "cell" level of detail. That is, we calculated R_t as the ratio of the average contract rate for a group of loans at a given age to the market rate available at the same point in time. R'_t was calculated based on the cell-level R_t. Similarly, we calculated $CUMDIFF_t$ based on the average contract rate for the group relative to the available market rate. It is our belief that there is very little difference between the values calculated at the cell-level and those calculated at the loan level of detail and weighted by amortized loan values.

House Price Appreciation

There are two house price appreciation variables used in the claims and prepayment rate models, an annual rate and a cumulative rate. Both are based on the historical house price index published by OFHEO.

We calculate the cumulative rate of house price appreciation by individual loan, and weight it based on the amortized values of loans surviving to age *t*. The cumulative rate for an individual loan is the ratio of the index value for the MSA (or state or census division) where the property is located at time *t* (plus three months) to the index value at the time the loan began amortizing (plus three months). We built a lag of three months into the index.

The annual rate of house price appreciation was based on the ratio of the average cumulative rate at time *t* to the cumulative rate at the previous age. This estimate of annual house price appreciation is slightly less clean than the calculation of the cumulative rate in that the mix of surviving loans by MSA may be slightly different between the two points in time. We do not consider that this "impurity" had a material effect on the results of our analysis.

The Probability of Negative Equity

We calculated probabilities of negative equity based on historical house price volatilities by MSA, by state, and by rural census division, published by OFHEO. The threshold for negative equity is an LTV ratio of 100%. Therefore, the calculated probabilities represent the probability that a loan with a given initial LTV will achieve a time-adjusted LTV of 100% or greater by time *t*.

The calculation of the probability of negative equity is by far the most labor-intensive calculation in terms of the required computer processing time. In order to save processing time, at what we felt was little or no cost in accuracy, we summarized the loans in our regression data sets by MSA. (Loans belonging to no MSA [i.e., rural properties] were grouped by census division, while non-rural properties that could not be assigned to an MSA were grouped by state.) We calculated a probability of negative equity for each MSA (or state or census division) at each point in time *t*, for each LTV increment proxy value. We could then weight the calculated probabilities for each "cell" based on the amortized value of surviving loans by MSA (or state or census division).

For clarification of the weighting calculation, see the example below:

Table E.10

		A Calculated Probability of Negative Equity			B Amortized Loan Value ($000)			A x B Extension		
LTV Increment Proxy:		**88.5**	**91.0**	**97.0**	**88.5**	**91.0**	**97.0**	**88.5**	**91.0**	**97.0**
MSA	**0040**	0.005	0.014	0.023	38	94	29	0.190	1.316	0.667
	0150	0.015	0.024	0.033	78	87	37	1.170	2.088	1.221
	4150	0.002	0.011	0.020	45	58	59	0.090	0.638	1.180
	7800	0.008	0.017	0.026	102	21	91	0.816	0.357	2.366
State	**AZ**	0.009	0.018	0.027	69	69	178	0.621	1.242	4.806
	NY	0.004	0.013	0.022	58	187	43	0.232	2.431	0.946
	PA	0.005	0.014	0.023	154	57	196	0.770	0.798	4.508
Rural Census	**New England**	0.008	0.017	0.026	68	98	87	0.544	1.666	2.262
Division	**Pacific**	0.018	0.027	0.036	38	138	29	0.684	3.726	1.044
	sum:				**650**	**809**	**749**	**5.117**	**14.262**	**19.000**
Weighted Average Probability of Negative Equity:								**0.79%**	**1.76%**	**2.54%**

For example, 0.79% = 5.117/650

External Data and Sources Used in Building Regression Data Sets

External Data	Source	Website
Countrywide unemployment rates - seasonally adjusted monthly civilian unemployment rate - by State - by Metropolitan Statistical Area (MSA)	U.S. Department of Labor, Bureau of Labor Statistics	http://stats.bls.gov/
One-Year U.S. Treasury Constant Maturity Rate	H.15 Release – Federal Reserve Board of Governors	http://www.stls.frb.org/fred/index.html
One-Year Treasury Constant Maturity Series (Weekly ARM Index) since 1975	HSH Associates	http://www.hsh.com
30-Year U.S. Treasury Constant Maturity Rate	H.15 Release – Federal Reserve Board of Governors	http://www.stls.frb.org/fred/index.html
30-Year Conventional Mortgage Rate Average Contract Rate on Commitments for Fixed Rate First Mortgages	Federal Home Loan Mortgage Corporation	http://www.stls.frb.org/fred/index.html
National Monthly Average Mortgage Rates for 30-Year Fixed, 15-Year Fixed, and Adjustable Rate Loans since 1983	HSH Associates	http://www.hsh.com/
Monthly Average Commitment Rates on 30-Year Fixed Rate Mortgages since 1971	Freddie Mac	http://www.freddiemac.com/pmms/pmms30.htm
Monthly Average Commitment Rates on 15-Year Fixed Rate Mortgages since 1991	Freddie Mac	http://www.freddiemac.com/pmms/pmms15.htm
Monthly Average Commitment Rates on 1-Year Adjustable Rate Mortgages since 1984	Freddie Mac	http://www.freddiemac.com/pmms/pmmsarm.htm
30-Year FHA Mortgage Rate, Secondary Market	U.S. Department of Housing and Urban Development	http://www.stls.frb.org/fred/index.html

External Data and Sources Used in Building Regression Data Sets, continued …

External Data	Source	Website
House Price Indices (as of 2000 2nd quarter) - by State, including District of Columbia - by MSA - by Census Division - by Rural Area	Office of Federal Housing Enterprise Oversight (OFHEO)	http://www.ofheo.gov/
House Price Volatility Parameters - by State, including District of Columbia - by MSA - by Census Division - by Rural Area	Office of Federal Housing Enterprise Oversight (OFHEO)	http://www.ofheo.gov/
Per Capita Personal Income - by MSA (1969-1998 annual) - by State (1958-1999 annual) - by State (1969-2000 quarter)	U.S. Department of Commerce, Bureau of Economic Analysis	http://www.bea.doc.gov
State Population Projections (1995-2000)	U.S. Department of Commerce, U.S. Bureau of the Census	http://www.census.gov/
Median House Price - by MSA - by State	1975 through 1997 from PriceWaterhouseCoopers File, 1998 through 2Q2000 estiimated based on OFHEO HPI series	