

Appendix C: Data for Loan Performance Simulations

This appendix describes the methodology we used to produce forecasts of future loan performance. We first describe how loan event data for future time periods were generated to project future loan performance and mortgage-related cash flows. This required creating future event data both for existing books of business and for future loan cohorts not yet originated. Then we summarize how the economic forecasts were applied. The forecasts of the economic factors are discussed in Appendix D. The derivation and application of the dispersion volatility parameters for the national average house price forecast are also explained in detail.

I. Future Loan Event Data

The development of future loan event data was closely integrated with the development of the data used in the statistical estimation of loan performance. As described in Appendix A, the process of building the historical loan event data entailed expanding FHA loan origination records into dynamic quarter-to-quarter event data from loan origination up to, and including, the period of loan termination. The loan event data were augmented with external economic data (house price indices and interest rates) to derive a number of time-varying predictors of conditional prepayment and claim rates.

For loans that did not terminate and still active as of FY 2008 Q2), the process of building the period-by-period event data followed the same procedure as for terminated loans, but used forecasted values of the external economic factors to project future loan termination rates and cash flows.

In addition, we forecasted the loan performance of future FHA books originated through FY 2015. The total endorsement volumes for FY 2008 through FY 2015 are provided by HUD from their internal demand model. These forecasted volumes are allocated among the six loan product types following their distribution in the most recent FY 2008 book. HUD estimates that streamline refinance loans will account for about 5 percent of the future endorsements. Besides the total endorsement volume, HUD also projected detailed compositions by LTV and credit score for 30-year fixed-rate mortgages in future books of business. Exhibits C-1 and C-2 present the composition of future books.

Exhibit C-1

Projected Originations By Mortgage Type (Percentage of Mortgages by Dollar Volume)						
Fiscal Year	Purchase Mortgages			Streamline Refinancing		
	30-Year	15-Year	ARMs	30-Year	15-Year	ARMs
	FRMs	FRMs		SRs	SRs	SRs
2009 to 2015	90.51	1.58	0.89	6.75	0.15	0.12

Exhibit C-2

Projected Portfolio Composition FY 2009 (%)							
Loan-to-Value Ratio	FICO Score Range						
	850-680	679-640	639-600	599-560	559-500	499-300	Missing
X <= 90	3.46	3.39	4.82	3.62	2.41	0.36	0.59
90 < X <= 95	5.85	5.39	5.90	4.19	2.51	0.00	0.42
95 < X <= 96.5	17.75	10.93	10.42	7.02	9.35	0.00	1.62
Projected Portfolio Composition FY 2010 to FY 2015 (%)							
Loan-to-Value Ratio	FICO Score Range						
	850-680	679-640	639-600	599-560	559-500	499-300	Missing
X <= 90	3.83	3.76	5.32	4.00	2.67	0.40	0.65
90 < X <= 95	6.47	5.96	6.52	3.57	1.72	0.00	0.47
95 < X <= 96.5	19.63	12.09	11.52	6.71	2.91	0.00	1.80

We then updated the initial mortgage contract rates of future loan originations according to the corresponding forecasted interest rate environments to reflect conditions at the time of origination. The future mortgage cash flows for individual loan stratifications are then aggregated to derive the total cash flows for the entire MMI Fund. The total cash flows are calculated as weighted average cash flows among individual stratifications, with the weights calibrated to the future demand and compositions forecasted by HUD.

As described in Appendix A, the data used for statistical estimation comprise detailed loan stratifications grouped by age of loan, all possible combinations of the categorical outcomes for the explanatory variables, and additional categories such as mortgage product types. The data for future cohorts are organized in the identical manner.

II. Future Economic Forecasts

FHA received quarterly economic forecast data from Global Insight, Inc. These data served as the independent source of projected future interest rates and house price appreciation rates in our analysis.

For the projection of future changes in housing values, we used the Global Insight forecast of the OFHEO national-level housing price index. Because the national-level housing price series is an average of regional housing price performance, it tends to smooth out deviations in house price trends among individual underlying regional indexes. There is also an additional layer of uncertainty with regard to the dispersion of individual house price appreciation rates around the market average, represented by the national-level house price index (HPI). When using the national-level house price forecast to compute the probability of negative equity, it is important to take into account both sources of uncertainty.

To address this issue, we developed a methodology to estimate the historical dispersion of the OFHEO regional (Census division) and metropolitan HPIs relative to the national HPI series. This analysis is described in greater detail in the remaining parts of this appendix. To summarize, estimates of additional dispersion among local housing markets were combined with the dispersion among individual house appreciation rates within the same local housing market to derive the total dispersion among individual house appreciation rates from the national average appreciation rate. The additional dispersion among local housing markets is added only after the beginning of the forecast period, i.e., as the computation of the probability of negative equity switches from using the actual MSA-level indexes to using the forecasted national-level HPI. The additional dispersion increases with time in a similar manner to the dispersion of individual property appreciation rates based on the MSA-level index.

Recall that the source of house price appreciation rates for historical loans was the HPIs obtained from OFHEO. The rule for assigning metropolitan area indexes was to use the Metropolitan Statistical Area Division (MSAD) index if one exists for the loan's Federal Information Processing Standards (FIPS) state-county code, otherwise we used the Core Based Statistical Area (CBSA) index if one is available. If no MSAD or CBSA index is available, we applied one of the nine Census-division HPIs.

As described in Appendix A, the indexes are used in conjunction with estimates of house price diffusion parameters to compute probabilities of negative equity at each loan age for individual borrowers. The dispersion estimate reflects the deviations among individual house price appreciation rates around the national average appreciation rate.

III. Dispersion of Local House Price Indices

As also described in Appendix A, the distributions of individual house values relative to the value at loan origination were computed using estimates of house price drift and diffusion volatility estimated by OFHEO.

To forecast the future changes in housing values, we utilized Global Insight's national-level HPI forecast. The Global Insight's national average house price data and forecast do not provide estimates of the diffusion volatility between a single property and the national index. Although OFHEO publishes a national-level HPI, it is based on a weighted average of indexes for the nine Census divisions, and no separate diffusion parameters are produced by Global Insight, Inc. at the national level. Thus, direct estimates of house price dispersion around a national index do not exist. Since we used Census-division-level and metropolitan-level indexes for calculating historical probability of negative equity, we adopted an approach that would build upon the OFHEO local house price volatility estimates and modify them to be consistent with the forecasting period when a national-level index is used.

We estimated the variance of the growth rates of housing values implied by the division indexes around the growth rates of the national-level index. The following discussion uses the case of MSA indexes as an example, but the same approach is also applied in the case of the Census-division indexes.

The growth rate for property i between time periods t and s relative to its MSA index is given by:

$$\ln(G_{i,t,s}) = \ln(H_{i,t}) - \ln(H_{i,s}) = \ln(H_{MSA,t}) - \ln(H_{MSA,s}) + \varepsilon_{i,t,s} \quad (1)$$

where $H_{i,t}$ is the value of the house i at time t , $H_{MSA,t}$ is the HPI of the surrounding MSA at time t , and $\varepsilon_{i,t,s}$ is the deviation of the growth rates of the house i and its surrounding MSA index from time s to t . Similarly, the growth rate implied by the MSA index relative to the national average forecast can be decomposed as follows:

$$\ln(G_{MSA,t,s}) = \ln(H_{MSA,t}) - \ln(H_{MSA,s}) = \ln(H_{N,t}) - \ln(H_{N,s}) + \varepsilon_{MSA,t,s} \quad (2)$$

where $H_{N,t}$ is the national HPI at time t , and $\varepsilon_{MSA,t,s}$ is the deviation of the growth rates of the MSA index and the national index from time s to t . Intuitively, one can think of the growth rate of a particular MSA as the national average growth rate plus a deviation term ($\varepsilon_{MSA,t,s}$). Similarly, the growth rate of a particular house is equal to the MSA-level average growth rate plus another deviation term ($\varepsilon_{i,t,s}$).

Plugging equation (2) into equation (1), we find that the individual house price growth rate equals the national average HPI growth rate and the sum of the dispersions of individual property around MSA levels, and the specific MSA levels around the national average growth rate:

$$\ln(G_{i,t,s}) = \ln(H_{N,t}) - \ln(H_{N,s}) + \varepsilon_{i,t,s} + \varepsilon_{MSA,t,s} \tag{3}$$

Notice that the variance of the first component of dispersion error given by $\varepsilon_{i,t,s}$ can be computed directly from the “ a ” and “ b ” parameters estimated by OFHEO using the three-stage weighted-repeat-sales methodology:

$$E\left(\sigma^2(\ln G_{i,t,s} - \ln G_{MSA,t,s})\right) = E\left(\sigma^2(\varepsilon_{i,t,s})\right) = a \cdot (t - s) + b \cdot (t - s)^2 \tag{4}$$

where $E(\bullet)$ is the expectation operator and $(t - s)$ is the number of quarters since the loan origination time s . We estimated the variance of the second component error $\varepsilon_{MSA,t,s}$ for the dispersion of the MSA index growth rate around the national index growth rate forecast by a linear regression in the following form:

$$\sigma^2(\ln G_{MSA,t,s} - \ln G_{N,t,s}) = \sigma^2(\varepsilon_{MSA,t,s}) = c \cdot (t - s) + e \tag{5}$$

where e is the residual of the regression. Because equation (4) was estimated by OFHEO as a residual term when estimating the MSA HPI using all houses within that location, the individual property diffusion term must be orthogonal to the diffusion term between the MSA and the national HPIs. That is, the noise term $\varepsilon_{i,t,s}$ is independent of $\varepsilon_{MSA,t,s}$, or:

$$\rho(\varepsilon_{i,t,s}, \varepsilon_{MSA,t,s}) = 0 \tag{6}$$

This implies the following model for the variance of individual house price cumulative appreciation rates around the national average forecast:

$$E(\sigma^2(\ln G_{i,t,s})) = a \cdot (t-s) + b \cdot (t-s)^2 + c \cdot (t-s) \tag{7}$$

The parameter “c” required for projecting the additional dispersion of the MSA index around the national average forecast was estimated as follows. For each quarter *t* we computed the cross-sectional (across MSAs) dispersion variance (MSA versus national) for each possible value of $(t-s) > 0$, which corresponds to time since loan origination, *i.e.*, mortgage age. We then computed the average dispersion variance according to each age of loan:

$$\sigma^2(\bar{\epsilon}_{t-s}) = \frac{1}{N} \sum_{\text{MSA}} [\ln(G_{MSA,t,s} / G_{N,t,s})]^2 \quad t = s+1, s+2, \dots, T \tag{8}$$

where *N* is the number of MSAs used in the estimation sample, and *T* is the most recent quarter that HPIs are available (first quarter of 2008 in this Review). This gives us a cross-section/time-series sample of average MSA index dispersion variance around the national average forecast that we assume is a linear function of $(t-s)$:

$$\sigma^2(\bar{\epsilon}_{t-s}) = c \cdot (t-s) + u_{t-s} \quad t = s+1, s+2, \dots, T; \quad s = 0, 1, \dots; s < t \tag{9}$$

where *u_{t-s}* is the regression residual. We estimated the unknown parameter “c” using a weighted least square regression with the number of average variance observations at each value of *t-s* as weights. The estimated quarterly standard deviations (\sqrt{c}) at age one, *i.e.*, $(t-s) = 1$, were 2.77 percent for MSA indexes and 2.72 percent for Census division indexes.

One of the following two formulas was applied depending on whether the time period was historical or future:

$$\begin{cases} E(\ln G_{i,t,s}) = \ln G_{MSA,t,s} \\ \sigma^2(\ln G_{i,t,s}) = a \cdot (t-s) + b \cdot (t-s)^2 \end{cases} \quad \text{if } s < t \leq T \tag{10a}$$

$$\begin{cases} E(\ln G_{i,t,s}) = \ln G_{MSA,T,t} + \ln G_{N,t,T} \\ \sigma^2(\ln G_{i,t,s}) = a \cdot (t-s) + b \cdot (t-s)^2 + c \cdot (t-T) \end{cases} \quad \text{if } s < T \leq t \quad (10b)$$

Equations (10a) were applied to historical sample time periods when either an MSAD or CBSA index was used to update expected housing values, and equations (10b) were applied during future simulated time periods when the national average forecast was used to update expected housing values.

For future loan originations only a single formula set is required:

$$\begin{cases} E(\ln G_{i,t,s}) = \ln G_{N,t,s} \\ \sigma^2(\ln G_{i,t,s}) = a \cdot (t-s) + b \cdot (t-s)^2 + c \cdot (t-s) \end{cases} \quad \text{if } T \leq s < t \quad (11)$$

Equations (11) were applied to future loan originations when only the national average forecast was used to update the expected housing values.

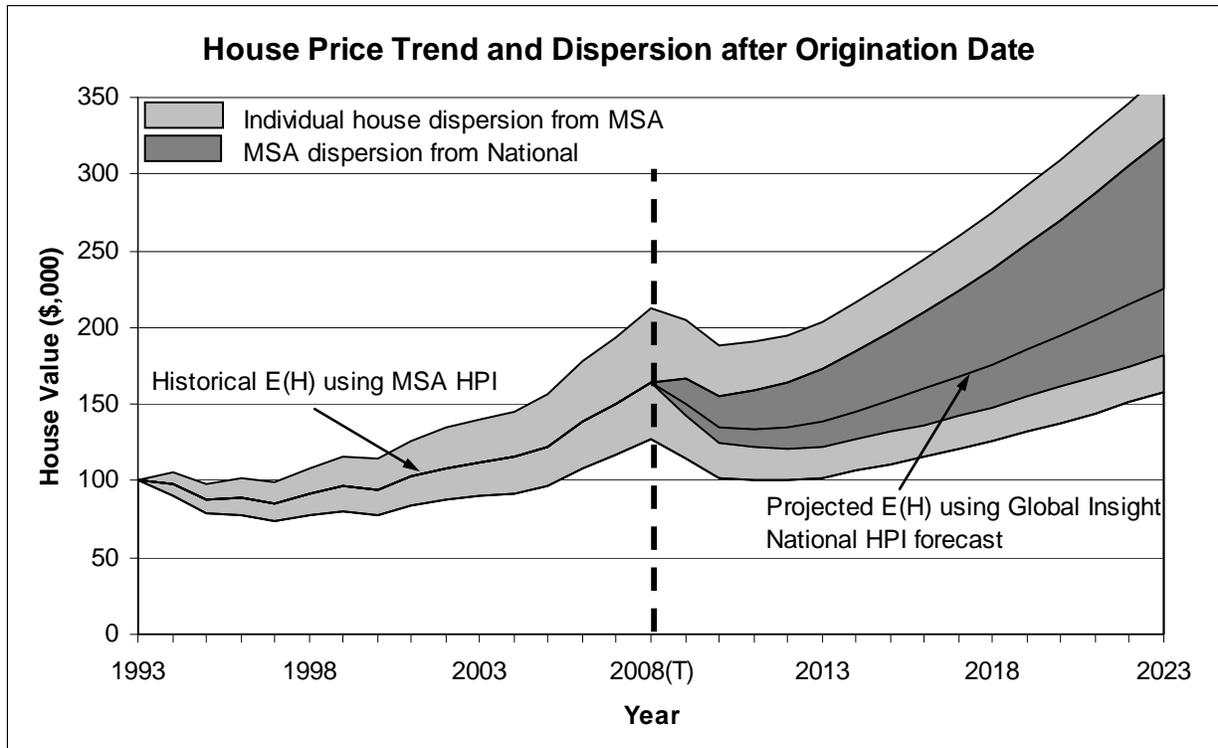
The additional term associated with the dispersion of the growth rate of an MSA HPI around the growth rate of the national HPI increases the overall dispersion volatility and results in higher probabilities of negative equity. This is counterbalanced by the reduced relative frequency of low expected HPI values when using a national average house price forecast instead of the more volatile metropolitan area HPIs.

IV. A Numerical Example

Exhibit C-3 presents an example to depict the expected value and the dispersion of the growth of the value of a house $[E(H)]$ over time. Let's assume that the sales price of an underlying house of a mortgage loan originated in a particular MSA at the first quarter of FY 1993 (s) was \$100,000. We further assume that the last historical HPIs available are as of the first quarter of FY 2008 (T).

The expected value and the dispersion standard deviation of this house at each exposure year (t) can be computed following equations (10b). The center line in Exhibit C-3 demonstrates the expected value of the house at each exposure year. This expected value was updated by using the OFHEO HPI for the particular MSA up to T . After T , the future expected value of the house is then updated according to the national house price growth rate forecasted by Global Insight, Inc. since a forecast of the HPI growth rate for the particular MSA is not provided.

Exhibit C-3



The house price growth rate of an MSA could be higher or lower than the national average for the next 15 years. The dark shaded area represents the possible future MSA house price index if the realized MSA growth rate is within one standard deviation above or below the national average growth rates. Note that, prior to T , such MSA-level dispersion did not exist in historical distributions since the expected house price was updated by using the MSA-specific index.

The light shaded area represents the additional dispersion of one standard deviation due to the individual house-level volatility around the MSA average. The size of this source of dispersion starts immediately after the house was sold in the first quarter of 1993 and continues to grow over the age of the mortgage loan. The OFHEO dispersion volatility parameters were used in computing this source of dispersion.

As a result, the outer boundary of the shaded area in Exhibit C-3 provides a visual demonstration of the dynamics of the house price distribution over the life of the mortgage. The probability of negative equity variable is directly computed as the probability that the house price may fall below the UPB of the mortgage at any point in time.