

# Supervisory Stress Tests, Model Risk, and Model Disclosure: Lessons from OFHEO\*

W. Scott Frame

University of North Carolina Charlotte,

FRB Atlanta

and FRB Richmond (Charlotte Branch)

Kristopher Gerardi

FRB Atlanta

Paul S. Willen

FRB Boston and NBER

April 3, 2013

**Abstract:** The Federal Reserve has recently embraced stress testing as an important component of its supervisory program for the very largest U.S. banking organizations. While stress tests can provide valuable insights, they are vulnerable to model risk. This paper explores the issue of model risk through a case study of a recent U.S. supervisory experience with a complex and fully disclosed stress test that failed spectacularly: OFHEO's risk-based capital stress test for Fannie Mae and Freddie Mac. Our analysis focuses on a key element of OFHEO's stress test: the performance of 30-year fixed-rate mortgages. After illustrating the poor default and prepayment forecasting performance of the model as implemented, we find that the primary explanation for this model failure was that OFHEO never re-estimated the model and hence left parameters static for almost a decade. We also show that default forecast performance would have been enhanced by including additional variables as the market evolved during the 2000s, like credit scores, documentation levels, vintage effects, and more disaggregated house price indices. Finally, we demonstrate that the house price stress inherent in the OFHEO model was significantly less stressful than the actual recent U.S. experience. Taken together, we believe that our results demonstrate the economic importance of model risk in stress testing and should lend support to efforts to mitigate such risk through continuous model development and independent validation.

**JEL Classifications:** G21, G23, G28

**Keywords:** Bank supervision, stress test, model risk, residential mortgages, government-sponsored enterprise

---

\*The views expressed do not necessarily reflect those of the Federal Reserve Bank of Atlanta, the Federal Reserve Bank of Boston, the Federal Reserve System, or their staffs. The paper has benefited from comments received at seminars at the Federal Reserve Board and Five Bridges LLC. The authors thank Neil Desai and Ellie Terry for outstanding research assistance. Contact information: wframe@uncc.edu, Kristopher.Gerardi@atl.frb.org, Paul.Willen@bos.frb.org

# 1 Introduction

On February 9, 2009, U.S. Treasury Secretary Timothy Geithner unveiled the Supervisory Capital Assessment Program (SCAP) in an effort to restore confidence in the U.S. financial sector in the aftermath of the global financial crisis. The SCAP principally involved a “stress test” of the 19 largest U.S. banking organizations (i.e., those with over \$100 billion in total assets) to determine whether each institution had sufficient capital to survive a protracted recession. The results of this supervisory exercise were made public, and the 10 banking organizations that were judged to have insufficient capital were given six months to raise the required funding from private markets or the U.S. Treasury.<sup>1</sup> The SCAP was widely viewed as credible and as having reduced uncertainty about the financial strength of covered institutions (e.g., Bernanke 2010; Tarullo 2010).<sup>2</sup>

Since the SCAP, the Federal Reserve has embraced the use of stress tests as an important ongoing component of its supervision program for very large, complex banking organizations. In 2010, the Federal Reserve introduced an annual Comprehensive Capital Assessment and Review (CCAR) to evaluate the capital planning processes and capital adequacy (under stress) of the same 19 banking organizations.<sup>3</sup> Shortly thereafter, the Dodd-Frank Act introduced mandatory stress testing for all banking organizations with greater than \$50 billion in total assets, as well as “systemically important non-bank financial institutions.”<sup>4</sup> The introduction of supervisory stress testing requirements may confer substantial benefits, such as enhanced risk measurement and management at covered banking organizations as well as supervisory learning about the institutions and system-wide vulnerabilities.<sup>5</sup>

---

<sup>1</sup>See Board of Governors of the Federal Reserve System (2009) for the SCAP results. Banking organizations unable to raise the required capital in private markets were eligible to receive such funding from the U.S. Treasury through the Capital Assistance Program (part of the Troubled Asset Relief Program, or TARP).

<sup>2</sup>Importantly, the credibility of these disclosures directly emanated from the fact that any identified capital shortfalls could be remedied by public-sector investment. Without such assurances, institutions identified as insufficiently capitalized almost assuredly would have experienced a run.

<sup>3</sup>Capital adequacy in the CCAR is determined based on stress test results produced by the Federal Reserve, although covered institutions are also required to run their own stress tests using both Federal Reserve and internally generated scenarios. Covered banking organizations must be projected to maintain Tier 1 common ratios in excess of five percent throughout the stress test. This requirement is in addition to satisfying the three standard capital adequacy targets post-stress: (1) Tier 1 capital ratio of four percent; (2) Total capital ratio of eight percent; and (3) Tier 1 leverage ratio of (generally) four percent.

<sup>4</sup>Section 113 of the Dodd-Frank Act indicates that systemically important non-bank financial institutions must be designated as such by the Financial Stability Oversight Council. Section 165 of the same law requires banking organizations with \$10-50 billion in total assets to conduct their own stress tests.

<sup>5</sup>Stress testing is not a new risk management practice, as large banks in the U.S. and Europe have actually reported conducting such tests for many years for individual business lines (e.g., Committee on the Global Financial System, 2001) and the International Monetary Fund conducts financial system-wide stress tests in individual countries as part of its Financial Sector Assessment Program. This raises the question of whether, prior to the SCAP, supervisors viewed such tests as somehow flawed or were otherwise not

Unfortunately, the one prior attempt to use supervisory stress testing to measure capital adequacy in the U.S. was, by all accounts, a spectacular failure. We are referring here to the risk-based capital stress test model for Fannie Mae and Freddie Mac, which are two government-sponsored enterprises (GSEs) that are central to the U.S. housing finance market and currently guarantee the performance of \$5.8 trillion of the \$10.0 trillion in home mortgage debt outstanding.<sup>6</sup> By law, the two GSEs operate exclusively in the secondary residential mortgage market by: (1) issuing credit guarantees on mortgage pools (securitization); and (2) engaging in leveraged mortgage investment in mortgage-backed securities (MBS).

Between 1992 and 2008, the Office of Federal Housing Enterprise Oversight (OFHEO) supervised Fannie Mae and Freddie Mac.<sup>7</sup> While the GSEs' minimum leverage requirement was set in statute, OFHEO was required to develop a risk-based capital regulation based on a stress test. The model and risk-based capital rule took OFHEO almost a decade to craft and finalize owing to the struggles of the new agency, the complexity of the stress test, and the politics associated with trying to regulate two very large and growing financial institutions viewed as crucial to the U.S. housing sector. Nevertheless, once in effect, the stress test was hailed as "state of the art" and as a mechanism to ensure that the two GSEs remained financially viable. Indeed, prominent economists concluded that if Fannie Mae and Freddie Mac could meet the OFHEO risk-based capital stress test their risk of insolvency was "effectively zero" (Stiglitz, Orszag, and Orszag 2002).<sup>8</sup>

Fannie Mae and Freddie Mac did maintain capital in excess of regulatory minimums throughout the 24 quarters that the risk-based capital rule was in force, including the stress test as of June 30, 2008. However, as market conditions continued to deteriorate thereafter, both GSEs were placed into federal conservatorship and each entered into preferred stock purchase agreements with the U.S. Treasury.<sup>9</sup> Thus far, Fannie Mae and Freddie Mac have received \$187 billion of direct government support. Hence, it is imperative that we

---

convinced of their utility. See <http://www.imf.org/external/NP/fsap/fsap.aspx> for further details of the IMF program, and Blaschke, Jones, Majnoni, and Peria (2001) for a nice primer on stress-testing.

<sup>6</sup>Data as of year-end 2011 from the Flow of Funds (Table L.218).

<sup>7</sup>OFHEO was created as part of the Federal Housing Enterprises Financial Safety and Soundness Act of 1992. The Housing and Economic Recovery Act of 2008 subsequently created the Federal Housing Finance Agency (FHFA), which consolidated the mission and safety and soundness oversight for Fannie Mae, Freddie Mac, and the Federal Home Loan Bank System. This involved a merger of OFHEO and the Federal Housing Finance Board along with some staff and functions at the U.S. Department of Housing and Urban Development.

<sup>8</sup>This statement is consistent with Hubbard (2004), who characterized the risk of direct economic loss associated with the possible failure of Fannie Mae as "low, both in absolute terms and relative to large commercial banks."

<sup>9</sup>See Frame (2008) for an analysis and discussion of the federal intervention at Fannie Mae and Freddie Mac.

understand how and why the OFHEO risk-based capital stress test failed.

This paper analyzes a key component of the OFHEO risk-based capital stress test: estimates of single-family 30-year fixed rate mortgage performance. This is the most popular mortgage contract in the U.S. and today comprises about 75 percent of Fannie Mae’s and Freddie Mac’s book-of-business. We specifically want to understand how the staleness of the model development and management processes affected model performance over time. Indeed, during the seven years that OFHEO’s risk-based capital stress test was active, the supervisor neither re-estimated (recalibrated) the model nor introduced new variables despite well-documented changes in mortgage underwriting practices during this time.<sup>10</sup>

Since loan-level data for the population of Fannie Mae and Freddie Mac loans is not publicly available, we use a large commercially available dataset that indicates loans acquired by the two GSEs and is representative of their overall book-of-business. Our empirical analysis first involves constructing mortgage default and prepayment forecasts based on our data and the static OFHEO parameters and then comparing these forecasts to realized outcomes over the 2001 to 2009 period. Next, we re-estimate the OFHEO model specification on a quarterly basis using our data to understand whether and how parameter estimates would have evolved, and then compare the associated mortgage default forecasts to those produced by the static OFHEO model and actual outcomes. We then augment the OFHEO mortgage default model with variables understood in the literature to affect mortgage performance (but omitted from OFHEO’s model) and re-estimate quarterly to see if forecasting accuracy is improved. Finally, we explore the role of house prices in OFHEO’s stress test: how the presumed path compared to the recent U.S. experience and its implications for default forecasts.<sup>11</sup>

We find that the OFHEO model did a very poor job of predicting 30-year fixed rate mortgage defaults and prepayments, especially during the recent housing bust. However, a large portion of the forecast error was attributable to the simple fact that the supervisor never re-estimated the model to update parameters. We find that the addition of omitted risk factors, like credit scores and loan documentation, improve the OFHEO model’s in-sample fit, but had only a marginal benefit for predicting future default. Finally, we discuss the role of house prices in the OFHEO risk-based capital stress test and show that: (1) the

---

<sup>10</sup>Why this occurred exactly is unclear, but we note that OFHEO faced significant political and process challenges. Morgenstern and Rosner (2011) offer several examples of significant political meddling in the GSE regulatory process. Moreover, the law required that the risk-based capital stress test be: (1) subject to notice and comment rulemaking; (2) sufficiently specific to permit anyone to apply the test given relevant data; and (3) made public (12 U.S.C. 4611). This may have limited the willingness and/or ability of OFHEO to make changes to the rule.

<sup>11</sup>See Glasserman, Kang, and Kang (2012) and Pritsker (2012) for detailed discussions and analyses concerning stress test scenario selection.

assumed house price path was not actually stressful for the first 10 quarters and was much less stressful than the recent U.S. housing bust overall; and (2) this scenario would have had a material and further negative effect on the supervisors' ability to require increased risk-based capital as U.S. house prices began to rapidly fall during 2007–2008. We believe that our findings represent a cautionary tale about over-reliance on stress testing by large financial institutions and supervisory authorities. Like all statistical representations of economic behavior, the models underlying comprehensive stress tests are generally quite varied, complex, data intensive, and assumption-laden. As a result, stress tests involve substantial “model risk” or the risk of mismeasurement associated with model misspecification, parameter estimation error, data limitations, or operational problems like model coding errors. Some amount of model risk can be mitigated through strong internal controls, as well as independent review and model validation.<sup>12</sup> But significant residual risks can remain owing to modeling choices and limitations that reflect the current (imperfect) state of scientific knowledge about integrated risk measurement and management systems.

The specific finding that the static nature of the OFHEO risk-based capital model was the principal source of failure demonstrates the importance of allowing stress tests to evolve with innovations to statistical methods, data, and market practice. However, it also naturally raises the question of why this occurred? We believe that OFHEO faced material challenges emanating from statutory model disclosure requirements which, coupled with the political power of Fannie Mae and Freddie Mac, would have made meaningful changes to the risk-based capital rule extremely costly.<sup>13</sup> By publishing model specifications and parameter estimates, Fannie Mae and Freddie Mac were able to take on risks that were not well-captured by the models (e.g., loans to borrowers with weak credit histories or very high loan-to-value ratios). This disclosure, by design, allowed the GSEs to successfully manage to the stress test. However, it may have also distracted Fannie Mae and Freddie Mac from developing more sophisticated internal risk management systems.

The remainder of the paper is structured as follows. Section 2 provides detailed background information about the OFHEO risk-based capital stress test and Section 3 describes the supervisory approach taken to 30-year fixed rate mortgage performance. Section 4 discusses the data and general empirical framework. Section 5 presents results, while Section

---

<sup>12</sup>In fact, such notions are a central part of recently issued supervisory guidance pertaining to model risk management. See Board of Governors of the Federal Reserve System and the Office of the Comptroller of the Currency (2011).

<sup>13</sup>The law required that the OFHEO risk-based capital stress test be: (1) subject to notice and comment rulemaking; (2) sufficiently specific to permit anyone to apply the test given relevant data; and (3) made public (12 U.S.C. 4611). Morgensen and Rosner (2011) offer several examples of significant political meddling in the GSE regulatory process. Taken together, this may have limited the willingness and/or ability of supervisor to make changes to the rule.

6 offers concluding remarks.

## 2 Background: The OFHEO Risk-Based Capital Stress Test

The Federal Housing Enterprise Financial Safety and Soundness Act (the 1992 Act) created a two-part regulatory structure for Fannie Mae and Freddie Mac. Mission regulation was to be conducted by the U.S. Department of Housing and Urban Development (HUD), while safety-and-soundness regulation was to be conducted by a new regulatory agency within HUD called the Office of Federal Housing Enterprise Oversight (OFHEO).

The 1992 Act subjected Fannie Mae and Freddie Mac to minimum and risk-based capital requirements to be enforced by OFHEO. The minimum capital requirement was set at 2.5 percent of on-balance sheet assets plus 0.45 percent for off-balance sheet credit guarantees. By contrast, the risk-based requirement was to be based on a stress test constructed by OFHEO, but subject to certain statutory requirements. Specifically, risk-based capital for Fannie Mae and Freddie Mac was to be sufficient to maintain positive capital throughout a 10-year period of stressful credit and interest rate conditions plus an additional 30 percent for management and operations risk. The law further dictated two important parameters of the risk-based capital stress tests relating to interest rate and credit risks.

In terms of interest rate risk, the 1992 Act specified two stress scenarios for the 10-year U.S. Treasury constant maturity rate (CMT). The first scenario involves the 10-year CMT rate falling by the lesser of 600 basis points below the average yield during the nine months preceding the stress period, or 60 percent of the average yield during the three years preceding the stress period, but in no case to a yield less than 50 percent of the average yield during the preceding nine months. The second path has the 10-year CMT rate rising by the greater of 600 bps above the average yield during the nine months preceding the stress period, or 160 percent of the average yield during the three years preceding the stress period, but in no case to a yield greater than 175 percent of the average yield during the preceding nine months.<sup>14</sup>

In terms of mortgage credit risk, OFHEO was to identify a “benchmark loss experience” based on the worst cumulative credit losses experienced by loans originated during a period

---

<sup>14</sup>OFHEO implemented the proscribed interest rate stress in the following way. For the 10-year stress period, OFHEO assumed that in both interest rate scenarios the 10-year CMT changes in 12 equal monthly increments from the starting point (the average of the daily 10-year CMT yields for the month before the stress period) and stayed at the new level for the remaining nine years of the stress period. OFHEO also established the relevant U.S. Treasury yield curve for the stress period in relation to the prescribed movements in the 10-year CMT. In the down-rate scenario the yield curve was assumed to be upward sloping during the last nine years of the stress period, while in the up-rate scenario the yield curve was flat during the last nine years of the stress period. All other interest rates were set as their average ratio to the comparable CMT for the two years prior to the stress period.

of at least two consecutive years in contiguous states comprising at least five percent of the U.S. population. Loans originated in Arkansas, Louisiana, Mississippi, and Oklahoma in 1983 and 1984 were identified by OFHEO. The mortgage credit risk element of the stress test was to then be “reasonably related” to the benchmark loss experience. As discussed below, this was done through adjustments to mortgage performance models as well as through the assumed path of house prices during the 10-year stress test horizon.

The general approach of the OFHEO stress test to mortgage performance (and hence mortgage credit risk) involved four principal steps. The first was the specification and estimation of statistical models of mortgage default and prepayment for different products. Second, adjustments were made to the statistical models to assure a reasonable relationship to the benchmark loss experience. Third, for the risk-based capital calculation in any particular quarter, contemporaneous mortgage data was run through the fitted and adjusted models to construct ten-year quarterly forecasts of expected default and prepayment probabilities assuming that house prices followed the path of the West South Central Census Region between 1984 and 1993. Finally, ten years of quarterly conditional cash flows were projected by loan group with the fraction of the group’s unpaid principal balance current, prepaid, and defaulted in each period. Defaulted loan balances were assumed to recover at a 61 percent baseline recovery rate, which was adjusted for the actual updated LTV and any prospective mortgage insurance proceeds.

OFHEO promulgated its risk-based capital rule for Fannie Mae and Freddie Mac in three steps. There was a First Notice of Proposed Rulemaking (June 1996) that addressed the methodology for identifying the benchmark loss experience and the use of OFHEO’s Census Division house price indices (HPI) to update original loan-to-value (LTV) ratios for loans held by Fannie Mae and Freddie Mac.<sup>15</sup> A Second Notice of Proposed Rulemaking (April 1999) outlined the remaining specifications of the stress test.<sup>16</sup> The final rule, which included several changes from the proposals, was issued in 2001 and became effective in 2002:Q4.<sup>17</sup> Figure 7 illustrates the overall framework for the risk-based capital stress test.

As shown in Figure 7, Fannie Mae (Panel A) and Freddie Mac (Panel B) each maintained capital in excess of statutory minimums for each of the 24 quarters that both capital requirements were in force. Moreover, the risk-based capital requirement always remained below the minimum capital requirement – suggesting that it was not a binding constraint. For example, while the GSEs’ minimum leverage and risk-based capital levels were always extremely low, Freddie Mac’s estimated risk-based capital requirement remained below 200

---

<sup>15</sup>See Federal Register 61(113), 29592-29621.

<sup>16</sup>See Federal Register 64(70), 18084-18131.

<sup>17</sup>See Federal Register 66(178), 47730-47875. OFHEO also issued a set of technical amendments to the rule in December 2006. See Federal Register 71(240), 75085-75106.

basis points during the entire 2003 to 2007 period. Both capital requirements were suspended with the imposition of the conservatorships at Fannie Mae and Freddie Mac in 2008.

### **3 Background: Single-Family 30-Year Fixed Rate Mortgage Performance in the OFHEO Stress Test**

The principal driver of credit losses in the OFHEO risk-based capital stress test are those associated with 30-year fixed-rate mortgages, and these loans are the focus of our study. OFHEO derived estimates of expected loan performance from statistical models of single-family mortgage default and prepayment, which were treated as competing risks and estimated jointly using a multinomial logit specification. OFHEO defined default as having occurred when a mortgage terminated with a loss. In such cases, default was then recorded as having occurred as of the last mortgage payment. Prepayment was defined as an instance in which the borrower voluntarily pays off the entire outstanding balance of the mortgage. The independent variables in the default and prepayment specifications were: loan age, original loan-to-value ratio, probability of negative equity, burnout, investor, relative spread, yield curve slope, and relative loan size. Each variable was represented categorically – indicating that a loan has a particular characteristic.<sup>18</sup>

Patterns of mortgage default and prepayments have characteristic age profiles, increasing during the first years after origination and then declining. OFHEO accounted for such loan seasoning by including a series of nine indicator variables for mortgage age (AGE) in both the default and prepayment models: six that correspond to each of the first six year’s of a loan’s life and then categories for loans aged seven to nine years, 10-12 years, and older than 12 years. Hence  $AGE = (AGE1, AGE2, AGE3, AGE4, AGE5, AGE6, AGE7-9, AGE10-12, AGE12+)$ .

The original loan-to-value ratio (LTV) is an indicator of the borrower’s financial resources and loans with higher LTVs are more likely to default and less likely to prepay. OFHEO included six original LTV categories in their model:  $LTV = (LTV \leq 60, 60 < LTV \leq 70, 70 < LTV \leq 75, 75 < LTV \leq 80, 80 < LTV \leq 90, 90 < LTV)$ . Virtually all mortgages will have origination LTVs below 100 percent, meaning that they have positive equity and little incentive to default at that time. However, over time, changes in area home

---

<sup>18</sup>Before estimation, OFHEO aggregated the loan-level data into groups of loans having similar characteristics, such as: product type, interest rate, original LTV, age, loan size, Census Division, etc. Hence, the default and prepayment models calculate the proportions of outstanding principal balances of loan groups. This was done to speed up computational time, as computers were significantly slower at that time. We implement the estimation on a random sample of the loan-level data rather than aggregating in the manner done by OFHEO.



prices can affect this equity position (positively or negatively) and hence the borrower’s propensity to default or prepay. To capture this, the OFHEO model also includes a measure of the probability that a borrower is currently in a position of negative equity (PNEQ), which is defined as the cumulative normal density for the ratio of the natural logarithm of the current LTV ratio to the contemporaneous HPI dispersion parameter (historical volatility) for the relevant Census Division.<sup>19</sup> The numerator of the current LTV is the current balance, while the denominator is the current estimated property value (based on the relevant U.S. Census Division HPI series). PNEQ is then assigned to categories,  $PNEQ = (0 < PNEQ \leq 0.05, 0.05 < PNEQ \leq 0.1, 0.10 < PNEQ \leq 0.15, 0.15 < PNEQ \leq 0.20, 0.20 < PNEQ \leq 0.25, 0.25 < PNEQ \leq 0.30, 0.30 < PNEQ \leq 0.35, 0.35 < PNEQ)$ .

Borrowers that have passed-up on previous opportunities to refinance when market rates are significantly below their current coupon rate are generally viewed as being either financially unsophisticated or experiencing financial difficulties. Such borrowers are more likely to default and less likely to prepay, holding other things constant. The indicator variable BURNOUT equals one if the market rate is 200 basis points below the loans coupon rate in any two quarters out of the first eight quarters of a loan’s life. Once detected, the burnout effect is phased-in over the first eight quarters: no effect during the first two quarters of a loan’s life, 25 percent effect during quarters three and four, a 50 percent effect during quarters five and six, and a 75 percent effect during quarters seven and eight.

For a given level of property (negative) equity, it is understood that investors are more likely to default than owner-occupiers. This occurs because the investors do not realize the personal consumption value of the home as shelter. Investors also tend to be more financially sophisticated and less credit constrained on average, and hence more likely to exercise their prepayment option. The variable investor (INVESTOR) indicates mortgages made to investors (including second homes and all 2-4 family properties).

Three additional variables were included in the prepayment model, but omitted from the default model. First, the relative spread between the interest rate on the mortgage and the current rate (RS) is a proxy for the “mortgage premium value”, or value to a borrower of the refinance option.  $RS = (RS \leq -0.20, -0.20 < RS \leq -0.10, -0.10 < RS \leq 0, 0 < RS \leq 0.10, 0.10 < RS \leq 0.20, 0.20 < RS \leq 0.30, 0.30 < RS)$ . Second, the slope of the yield curve (YCS) is measured as the difference between the 10-year CMT and 1-year CMT. The shape of the yield curve reflects expectations of the future levels of interest rates and will thereby affect borrowers’ mortgage prepayment decisions. For purposes of the model the slope is categorized in the following way:  $YCS = (YCS < 1.0, 1.00 \leq YCS < 1.20, 1.20 \leq YCS < 1.50,$

---

<sup>19</sup>The probability of negative equity is included in the model – as opposed to the direct estimate of a borrower’s equity position – in order to account for the measurement error that comes from using an aggregated house price index to estimate the values of individual properties.

1.50 $\leq$ YCS). Finally, the size of a particular loan relative to its state average (RLS) may be related to prepayment behavior insofar as refinancing costs are proportionately higher for lower balance loans.  $RLS = (0 < RLS \leq 0.40, 0.40 < RLS \leq 0.60, 0.60 < RLS \leq 0.75, 0.75 < RLS \leq 1, 1 < RLS \leq 1.25, 1.25 < RLS \leq 1.50, 1.50 < RLS)$ .

While the same set of covariates was included in the empirical specification for both the default and prepayment hazards, certain parameters of the default hazard were constrained to be zero in the estimation routine (i.e., those associated with relative spread, yield curve slope, and relative loan size). The multinomial logit was estimated using a 10 percent random sample of mortgage loans that Fannie Mae and Freddie Mac had securitized or retained between 1979 and 1999 (with origination years from 1979 to 1997) using the CATMOD procedure in SAS.<sup>20</sup>

## 4 Data

OFHEO used proprietary data on residential mortgages held by Fannie Mae and Freddie Mac to estimate their single-family mortgage performance models for the risk-based capital stress test, which are not available to us. As a result, we use commercially available loan-level mortgage data from Lender Processing Services (LPS) for 1993-2009 to re-estimate the OFHEO model specification as well as to conduct three principal empirical exercises (described below).

The LPS data are collected from several of the largest U.S. mortgage servicers and cover a large fraction of active loans.<sup>21</sup> The LPS data include a large number of standard mortgage underwriting fields. Loan-level attributes include borrower characteristics (e.g., origination FICO score, occupancy status, and documentation level), collateral characteristics (e.g., property type, original loan-to-value ratio, and zip code), and loan characteristics (e.g., loan balance, lien holder type, and loan status). The monthly history of each loan appears in the data including their current payment/performance status.<sup>22</sup> One issue with the LPS data is that not all servicers populate all fields, although this was primarily an issue before the mid-2000s and the affected fields were generally not those used in the OFHEO risk-based capital model anyway (investor status excepted). We come back to this issue below.

The LPS field “lien holder type” allows us to identify those loans held or guaranteed by Fannie Mae and Freddie Mac. These comprise our loan sample. To check the repre-

---

<sup>20</sup>OFHEO used the CATMOD procedure in SAS to obtain estimated parameters for all values of the categorical variables. We also use CATMOD in our analysis below.

<sup>21</sup>The LPS loan-level dataset covers approximately 40 million active first lien mortgages and 8 million active second lien mortgages.

<sup>22</sup>See Foote, Gerardi, Goette, and Willen (2009) for a more detailed discussion of the LPS dataset.

representativeness of our sample, we compare the annual sample means for certain key variables (origination loan-to-value ratio, unpaid principal balance at origination, and interest rate at origination) to those provided to us by staff at FHFA for the population of Fannie Mae and Freddie Mac loans held or guaranteed each year between 1995 and 2005. The comparisons are provided in Table 7. There are minor differences between the two datasets in any given year, but the broad patterns are quite consistent and suggest that the LPS data are quite representative.

For each quarter under study (1993:Q1 through 2009:Q4), we pare down the number of loans using the following selection criteria. First, we only include loans that LPS indicates as being held by Fannie Mae or Freddie Mac. By law, these loans must have original balances below the “conforming loan limit” for the year and location that the loan was made.<sup>23</sup> Second, we consider loans only in the 48 contiguous U.S. states, which is consistent with OFHEO’s sample restriction. We further require each loan to: indicate that it finances a single-family residence, maintain a first-lien position, and is a fully amortizing 30-year fixed rate note.

Because of the large number of active loans in any given quarter in the LPS data, we take random samples to speed up estimation. However, because the LPS data coverage relative to the population of outstanding mortgages varies over time, our sampling is not uniform. LPS added mortgage servicers to their database over time, thereby increasing their coverage of the U.S. mortgage market. In order to maintain an approximately constant number of loans in our estimation sample we decrease the proportions of the random samples over time. For loans originated before the end of 1998:Q4, we take a 30 percent random sample of loans meeting our selection criteria. Then, for loans originated during 1999:Q1 through 2004:Q4 and meeting our selection criteria, we use a 21 percent random sample, and for loans thereafter we take a 17 percent random sample. These samples are used to estimate the various models over different time horizons. Also, when comparing forecasts generated by the various mortgage performance models to realized outcomes, we utilize five percent random samples for the outcomes.

Our analysis also requires information about house prices and interest rates. In order to replicate the OFHEO mortgage model, we collect quarterly Census Division house price indices and associated price volatility series from the Federal Housing Finance Agency. In some additional analysis, we also utilize county-level house price series available from CoreLogic. In terms of interest rates, we collect monthly series for 30-year mortgage rates, as well as 1-year and 10-year Treasury rates from the Federal Reserve Board website.

---

<sup>23</sup>See U.S. Federal Housing Finance Agency (2011) for historical data about the conforming loan limits.

## 5 Empirical Analysis

We conduct several exercises aimed at understanding whether and how the OFHEO first-lien 30-year fixed-rate mortgage performance models (default and prepayment) performed in the years leading up to the mortgage bust and subsequent financial crisis. Because the LPS data do not include information about mortgage losses, we define default to occur when a foreclosure is completed and then date the default back to the last observed payment. We define prepayment in the same manner as OFHEO. Table 2 compares the default and prepayment parameter estimates for 30-year fixed rate mortgages published by OFHEO (based on proprietary Fannie Mae and Freddie Mac loan data between 1979 and 1999) to our estimates using the LPS data between 1994 and 2000. For brevity, we display the parameters associated with the LTV and probability of negative equity (PNEQ) variables. The parameter estimates are surprisingly consistent given the fact that the OFHEO and LPS estimation samples have very little overlap (only 6 years). The signs of the parameter estimates are almost identical across all categories, and the magnitudes are very similar.

### 5.1 OFHEO Default Forecasts with Static Parameters and Perfect Foresight

Our first exercise explores how well the OFHEO first-lien mortgage model would have predicted quarterly default propensities assuming the supervisor had perfect foresight about house prices and interest rates in the next quarter. The perfect foresight assumption is made in order to obtain a clear determination of how well the OFHEO model can predict defaults. We construct one-quarter-ahead default probability forecasts using the public OFHEO estimates and compare these to realized default rates in the LPS data.

Figure 7 presents the actual and predicted default rates for each quarter from 2000 to 2009. Actual defaults were very low between 2000 and 2006, hovering around 0.10 percent of total active GSE mortgages. However, at the same time the OFHEO model seems to have been under-predicting defaults, as the average ratio of actual to predicted defaults was approximately 1.5 over this period. Actual defaults then climbed steadily to about 0.53 percent of total GSE mortgages by June 2008. At that time, predicted defaults were only 0.14 percent indicating a forecast error of almost four times!

Figure 7 displays actual and predicted one-quarter-ahead prepayment rates from 2000 to 2009. The model predicts prepayments relatively accurately from 2000 through 2002 and from 2004 through 2007, but badly misses in two sub-periods. It severely under-predicts prepayments during the refinance boom in 2003, and severely over-predicts prepayments

during the financial crisis period in 2008 and 2009.<sup>24</sup>

## 5.2 OFHEO Default Forecasts with Dynamic Parameters and Perfect Foresight

The second exercise extends our analysis by simply updating the parameters of the OFHEO first-lien mortgage model on a quarterly basis (i.e., up to  $k - 1$ ). We do this by re-estimating the OFHEO model using the LPS data based on a seven year rolling window and then once again relating these updated default forecasts to realized defaults assuming perfect foresight about the next quarter's house prices and interest rates. The first estimation window spans 1993:Q1 to 2000:Q1 and is then updated quarterly through 2009:Q4. The idea behind this exercise is to determine whether default forecasts could be improved by simply updating the OFHEO model parameters using data available in real-time.

Figure 7 presents the ratio of the actual default rate to the predicted 1 quarter-ahead default rate for each quarter from 2006 to 2009 for both the static OFHEO model (using the estimates from Figure 7) as well as the forecast based on the rolling regressions. Again, each forecast is predicated on the supervisor having perfect foresight about the next quarters' values of house prices and interest rates. It is quite clear that simply re-estimating the OFHEO model each quarter dramatically reduces the forecast error. While the static OFHEO model under-predicts one quarter-ahead default rates by almost a factor of four in the latter half of 2008, the updated OFHEO model under-predicts by a factor of only 1.5. Another pattern worth noting from Figure 7 is that, unlike the static model which under-predicts default rates throughout the entire post-2006 period, the updated model actually over-predicts defaults in 2006 and the first half of 2007.

To dig a little deeper into the source for this dramatic improvement in predictability, we graph the quarterly time series evolution of the LTV and PNEQ default hazard parameter estimates in Figure 7. Focusing on the LTV parameter estimates, it is clear from the figure that over time, up through the peak of the crisis period (mid-2008), the higher LTV indicators become more positively correlated with default while the lower LTV indicators become more negatively correlated with default. A similar pattern can be seen in the graph of the PNEQ default hazard parameter estimates as well. Thus, LTV and PNEQ became more powerful predictors of default over time in the updated model (although this pattern appears to have reversed itself at the very end of the sample for the LTV variable).<sup>25</sup>

---

<sup>24</sup>For the remainder of the paper we will focus strictly on default rates, but prepayment figures are available upon request from the authors.

<sup>25</sup>We also conducted a similar exercise for the other covariates in the OFHEO model, but did not find as pronounced changes in the parameter estimates for those variables.

### 5.3 OFHEO Default Forecasts with Dynamic Parameters, Perfect Foresight, and Additional Variables

The OFHEO model specification lacks several covariates that have been shown to have predictive power in forecasting mortgage defaults. This is likely a result of initially poor data availability, the lack of updating, and changes in industry practice over time. Our third exercise explores whether adding additional relevant predictors to the OFHEO model (updated every quarter) improves the default forecasts. Specifically, we explore the potential roles of FICO credit scores, loan documentation, loan vintage, and unemployment rates.<sup>26</sup>

First, for credit scores, we include a series of categorical variables in 40-point increments.<sup>27</sup> The specific categories are:  $FICO \leq 620$ ,  $620 < FICO \leq 660$ ,  $660 < FICO \leq 700$ ,  $700 < FICO \leq 740$ ,  $740 < FICO \leq 780$ ,  $780 < FICO \leq 820$ , and  $FICO \geq 820$ . Second, the lack of loan documentation has been previously identified as a risk factor, as well as a contributor to the recent housing bust. Moreover, the GSEs became significant purchasers of low documentation mortgages during the housing boom as such loans became a greater share of the marketplace.<sup>28</sup> Hence we add variables indicating whether the loan was a “no doc” or “low doc” mortgage. We also add year of origination fixed effects in order to capture unobserved changes over time in underwriting standards, as it is well-documented that mortgage underwriting standards decreased dramatically during the housing boom in observable and unobservable ways (e.g., Gerardi, Lehnert, Sherlund, and Willen, 2008; Demyanak and van Hemert 2011). Finally, we also add county-level unemployment rates from the Bureau of Labor Statistics, as job loss is likely to be an important factor in a borrower’s decision to stop making mortgage payments.<sup>29</sup> We add both the level of the unemployment rate as

---

<sup>26</sup>We also tried to expand the set of original LTV indicator variables. First, we redefined  $LTV > 90$  as a series of indicator variables:  $90 < LTV \leq 95$ , and  $LTV > 95$ , to account for the dramatic rise in high-leverage mortgages originated during the boom. Second, we tried including an indicator for loans with loan-to-value ratios exactly equal to 80 percent ( $LTV = 80$ ) to account for the fact that some of these loans had unobserved subordinate liens. However, neither of these changes substantially affected the model forecasts.

<sup>27</sup>The first FICO score was made available to the three major U.S. credit bureau agencies in 1991. However FICO scores were not introduced into mainstream mortgage models until the mid-1990s (for more details we direct the reader to <http://www.fico.com/en/Company/Pages/history.aspx>). Thus, FICO scores were unavailable to OFHEO when it estimated the risk-based capital model. Low documentation mortgages did not become popular until the 1990s, and thus were likely not prevalent in the sample of mortgages used by OFHEO to estimate the model. See Foote, Gerardi, and Willen 2011 for more details on the history of low documentation loans.

<sup>28</sup>At the peak of the housing boom in 2006, almost 40 percent of newly originated subprime mortgages had less than full documentation of income and assets (Gerardi, Lehnert, Sherlund, and Willen 2008). Low documentation mortgages were even more common in the Alt-A segment of the market, reaching a peak of 78 percent of originations in early 2007 (Sengupta 2010).

<sup>29</sup>Previous empirical default studies have not found a strong correlation between the incidence of mortgage default and unemployment rates at the state or county level. However, Gyourko and Tracy (2013) show that a weak correlation between aggregate unemployment rates and default could be consistent with a strong correlation between household-level unemployment shocks and default due to a large attenuation bias that

well as the cumulative change in the unemployment rate since the quarter of origination. The first variable likely captures persistent differences in unemployment across geographies, while the second variable captures differences in the evolution of unemployment rates across geographies during the life of the mortgage.<sup>30</sup>

The OFHEO residential mortgage default model included a variable, PNEQ, intended to capture the probability that a given mortgage is in a negative equity position (i.e., “underwater”) based on updated property values and amortization. Property values were updated using the OFHEO/FHFA house price index (and index dispersion measure) for the Census Region in which the property was located. While a reasonable attempt to capture the effect of changes over time in home equity positions, the use of regional house price indices may significantly reduce the usefulness of this variable, as the correlation between changes in individual property values and changes in such an aggregated index are likely weak. We attempt to at least partially address this issue by reconstructing the PNEQ variable using a more disaggregated house price index at the county-level from CoreLogic.

Figure 7 compares the ratio of actual-to-predicted defaults under the baseline OFHEO model to revised models that include each of the new covariates: Corelogic county-level house price index for computing PNEQ, FICO credit scores, documentation level, origination-year fixed effects, and the unemployment rate variables. We add each covariate sequentially to determine which one has the most substantial impact on the forecasting ability of the model. We start by substituting the county-level house price indices for the regional indices in the computation of the PNEQ variable. This change improves the forecasts during the crisis period of 2007 and 2008, as the model under-predicts defaults by significantly less. We then add FICO scores and documentation levels to the model, but find no significant difference in the aggregate, out-of-sample forecasts of the model. Next, we add origination year fixed effects, which has the effect of improving the default forecasts during the crisis period in 2008. However, adding these effects causes the model to over-predict defaults by significantly more during the period right before the crisis in 2006 and 2007. Finally we add the unemployment variables, which has virtually no effect until 2008, when it then causes the model to further over-predict defaults.

It is somewhat surprising that the addition of credit scores and documentation status to the model does not improve its ability to forecast aggregate default rates. It is possible though that the addition of these variables may improve the ability of the model to predict defaults at the individual loan level. In Figure 7 we consider this possibility by displaying

---

occurs by using aggregate unemployment rates to proxy for individual unemployment shocks. Since we do not have individual data, we are unfortunately forced to use aggregate rates.

<sup>30</sup>We also experimented with shorter term changes in unemployment rates, such as the change in unemployment over the previous 4 quarters, but found no significant differences in the forecasting results.

the C-Statistic for each variation of the model that we considered in Figure 7. The C-statistic is an in-sample goodness-of-fit measure that can take values between 0.5 and 1, where 0.5 corresponds to the case where the model is no better than chance at predicting which mortgages default and which do not and 1 corresponds to the case where the model is perfect in distinguishing defaults from non-defaults. The first observation worth noting in Figure 7 is that the baseline OFHEO model in which parameter estimates are updated each quarter is characterized by a relatively high C-statistic that ranges between 0.79 and 0.82 over the entire sample period. The second noteworthy result is that the addition of FICO scores provides the largest improvement in terms of in-sample predictability at the loan-level. Thus, while the addition of credit scores to the model does not seem to improve the model's ability to forecast aggregate defaults, it does improve the ability of the model to distinguish between defaults and non-defaults at a more disaggregated level.

## 5.4 House Price Stress

Our final exercise explores the role of house price stress in the OFHEO model. Figure 7 compares the 5-year expected house price path used by OFHEO in the risk-based capital stress test with the 5-year realized path of U.S. house prices, measured using FHFA's national house price index since the beginning of the housing bust (fourth quarter of 2006). Recall that the OFHEO house price stress is the realized path of house prices for the West South Central Census Region between 1984 and 1993. To construct the figure we took the quarterly growth rate of house prices in the West South Central Census Region between 1984 and 1988, and applied them to the level of house prices that prevailed at the beginning of the housing bust at the end of 2006. It is striking that, after the first ten quarters of the OFHEO stress test, home prices were assumed to increase by two percent from their starting point. By contrast, the recent US experience would have had them down nine percent.

Recall that all of our previous exercises thus far were predicated on evaluating the OFHEO model assuming that the supervisor had perfect foresight about future house prices and interest rates. This final exercise replaces realized house price movements with those mandated for the OFHEO risk-based capital stress test. Figure 7 presents the predicted cumulative three-year expected default rates each quarter between 2000:Q1 and 2007:Q4 from the static OFHEO model (without updating parameters) based on the house price path assumed by OFHEO versus that observed during the recent crisis. It appears that the OFHEO house price path results in three-year expected cumulative default rates that are, on average, around 20-25 basis points lower, which reflects the fact that the house price stress scenario used by OFHEO is significantly less stressful than the actual experience of house prices through the housing bust and financial crisis.



## 6 Conclusion

Stress tests have become an increasingly important part of financial institution risk management programs, and an important tool used by supervisory authorities to evaluate the financial health of large banking organizations and financial systems. While stress testing exercises can provide valuable insights, they are vulnerable to model risk.

This paper studied a recent U.S. supervisory experience with a complex and fully disclosed stress test that failed spectacularly: OFHEO’s risk-based capital stress test for Fannie Mae and Freddie Mac. Our analysis focused on a key element of OFHEO’s stress test: the model used to predict default and prepayment of 30-year fixed-rate mortgages. We first demonstrated the poor out-of-sample forecasting performance of OFHEO’s default and prepayment models, especially during the recent housing bust. The principal cause of this failure appears to have simply been that the supervisor never re-estimated the model and hence left parameters static for almost a decade. We show that this was problematic because certain parameters, like those associated with borrower leverage, were unstable likely due to changes in market practice.

Another important reason for the OFHEO model’s failure was the exclusion of certain variables which became increasingly common in residential mortgage modeling over the past decade, such as credit scores, indicators for level of documentation, and more disaggregated house price indices. Interestingly, we find that such factors significantly improved model fit, but resulted in only modest improvements in mortgage default forecasts.

We also reviewed the role of house prices in the OFHEO risk-based capital stress test and showed that the assumed house price path was not actually stressful for the first 10 quarters of the stress test horizon and was much less stressful than the recent U.S. housing bust overall. As a result, the assumed house price path would have had a material negative effect on the supervisor’s ability to require increased capital as US house prices began to rapidly fall in 2007 and 2008.

The poor performance of OFHEO’s 30-year fixed-rate mortgage default and prepayment model used for setting Fannie Mae’s and Freddie Mac’s risk-based capital represents a concrete example of model risk. Our analysis illustrates that simple updating and some modest research and development would have provided a much clearer picture of the emerging distress in the mortgage portfolios held or guaranteed by the two GSEs. While this may seem like a straightforward case of supervisory failure, OFHEO did face some legal and political constraints. By law, OFHEO’s risk-based capital stress test was to be “sufficiently specific to permit anyone to apply the test given relevant data.” This provision likely contributed to OFHEO’s decision to publish all of the model specifications and parameter estimates. How-

ever, once in the public domain, any model changes would have been very costly. These costs would have been direct (required notice-and-comment, interagency clearance, and republication) and indirect (political fallout from proposing any changes that would disadvantage Fannie Mae or Freddie Mac).

Overall, we draw three broad lessons from the U.S. experience with the OFHEO risk-based capital stress test for Fannie Mae and Freddie Mac. First, while stress tests can provide valuable insights, they are subject to significant model risk. Second, real efforts should be taken to mitigate model risk through continuous development and model validation. Finally, in the case of supervisory stress tests, the full disclosure of models and parameters can result in financial institutions attempting to game the models and mute incentives for investment in proprietary risk management systems.

## 7 References

- Bernanke, Ben, 2010. "The Supervisory Capital Assessment Program One Year Later" Speech at the Federal Reserve Bank of Chicago 46th Annual Conference on Bank Structure and Competition, Chicago (May 6).
- Blaschke, Winfrid, Matthew T. Jones, Giovanni Majnoni, and Soledad Maria Peria, 2001. "Stress Testing of Financial Systems: An Overview of Issues, Methodologies, and FSAP Experiences." IMF Working Paper 01-88.
- Board of Governors of the Federal Reserve System, 2009. "The Supervisory Capital Assessment Program: Overview of Results" Available online at: <http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20090507a1.pdf>.
- Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency, 2011. "Supervisory Guidance on Model Risk Management." OCC Bulletin 2011-12.
- Demyanyk, Yuliya, and Otto Van Hemert. "Understanding the Subprime Mortgage Crisis." *Review of Financial Studies* 24(6), 1848-1880.
- Foote, Christopher L., Kristopher S. Gerardi, Lorenz Goette, and Paul S. Willen, 2009. "Reducing Foreclosures: No Easy Answers," *NBER Macroeconomics Annual*, 24, 80-138.
- Foote, Christopher L., Kristopher S. Gerardi, and Paul S. Willen, 2012. "Why Did so Many People Make so Many Ex-post Bad Decisions? The Causes of the Foreclosure Crisis" in *Rethinking the Financial Crisis*, edited by Blinder, A., Lo, A., and R. Solow, Russell Sage Foundation: New York.
- Frame, W. Scott, 2008. "The 2008 Federal Intervention to Stabilize Fannie Mae and Freddie Mac," *Journal of Applied Finance*, 18, 124-133.
- Gerardi, Kristopher S., Andreas A. Lehnert, Shane S. Sherlund, and Paul S. Willen, 2008. "Making Sense of the Subprime Crisis," *Brookings Papers on Economic Activity*, Fall, 69-159.
- Glasserman, Paul, Chulmin Kang, and Wanmo Kang, 2012. "Stress Scenario Selection by Empirical Likelihood," Working Paper, Columbia University.
- Goldstein, Itay and Haresh Sapra, 2012, "Should Banks Stress Test Results be Disclosed? An Analysis of the Costs and Benefits." Working Paper, University of Chicago Booth School of Business.
- Gyourko, J. and J. Tracy, 2013. "Unemployment and Unobserved Credit Risk in the FHA Single Family Mortgage Insurance Fund," NBER Working Paper 18880.
- Hubbard, R. Glenn, 2004. "The Relative Risk of Freddie and Fannie," *Fannie Mae Papers*, Vol. 3, Issue 3.

Morgenstern, Gretchen and Joshua Rosner, 2011. *Reckless Endangerment: How Outsized Ambition, Greed, and Corruption Led to Economic Armageddon*.

Pritsker, Matthew, 2012. “Enhanced Stress Testing and Financial Stability,” Federal Reserve Bank of Boston Working Paper.

Schuermann, Til, 2012. “Stress Testing Banks,” Wharton Financial Institutions Working Paper 12-08.

Sengupta, Rajdeep, 2010. “Alt-A: The Forgotten Segment of the Mortgage Market,” *Federal Reserve Bank of St. Louis Review*, 92(1), 55-71.

Stiglitz, J., Orszag, J., and P. Orszag, 2002. “Implications of the New Fannie Mae and Freddie Mac Risk-based Capital Standard,” *Fannie Mae Papers* Vol. 1, Issue 2.

Tarullo, Daniel K., 2010. “Lessons from the Crisis Stress Tests,” Speech at the Symposium on Building the Financial System of the 21st Century, Armonk, New York (March 18).

U.S. Congressional Oversight Panel, 2009. “Stress Testing and Shoring up Bank Capital,” Washington D.C., U.S. Government Printing Office.

**Table 1: Comparison of OFHEO and LPS Datasets**

<b>Panel A: Fannie Mae</b>						
Year	Avg. LTV Ratio (%)		Avg. UPB (\$)		Avg. Interest Rate (%)	
	OFHEO	LPS	OFHEO	LPS	OFHEO	LPS
1995	80.1	79.5	101,518	101,393	8.1	8.6
1996	79.1	77.3	105,059	107,358	8.0	8.1
1997	78.1	78.5	111,398	115,546	7.8	8.0
1998	76.2	78.0	122,646	129,966	7.1	7.1
1999	77.6	76.8	123,600	128,224	7.4	7.2
2000	78.9	77.9	128,041	137,490	8.2	8.1
2001	76.2	74.9	145,435	148,313	7.1	7.1
2002	74.3	74.2	153,982	155,927	6.7	6.7
2003	72.2	72.4	162,743	160,537	5.9	5.9
2004	74.4	70.8	162,513	161,472	6.0	6.0
2005	73.8	72.4	175,886	164,631	6.0	6.1

<b>Panel B: Freddie Mac</b>						
Year	Avg. LTV Ratio (%)		Avg. UPB (\$)		Avg. Interest Rate (%)	
	OFHEO	LPS	OFHEO	LPS	OFHEO	LPS
1995	78.8	75.8	103,682	111,936	8.1	7.8
1996	78.2	71.6	106,414	115,373	8.0	7.9
1997	77.6	74.9	112,231	123,512	7.8	7.8
1998	75.5	73.8	122,976	123,961	7.1	7.0
1999	77.2	76.2	123,772	127,018	7.4	7.2
2000	78.4	71.1	128,781	126,022	8.2	8.1
2001	76.1	72.1	145,741	145,714	7.0	7.1
2002	74.5	72.5	153,380	158,182	6.6	6.8
2003	72.5	68.5	159,715	181,222	5.8	5.7
2004	74.3	72.6	164,079	177,177	5.9	5.9
2005	72.7	72.1	178,889	178,042	5.9	5.8

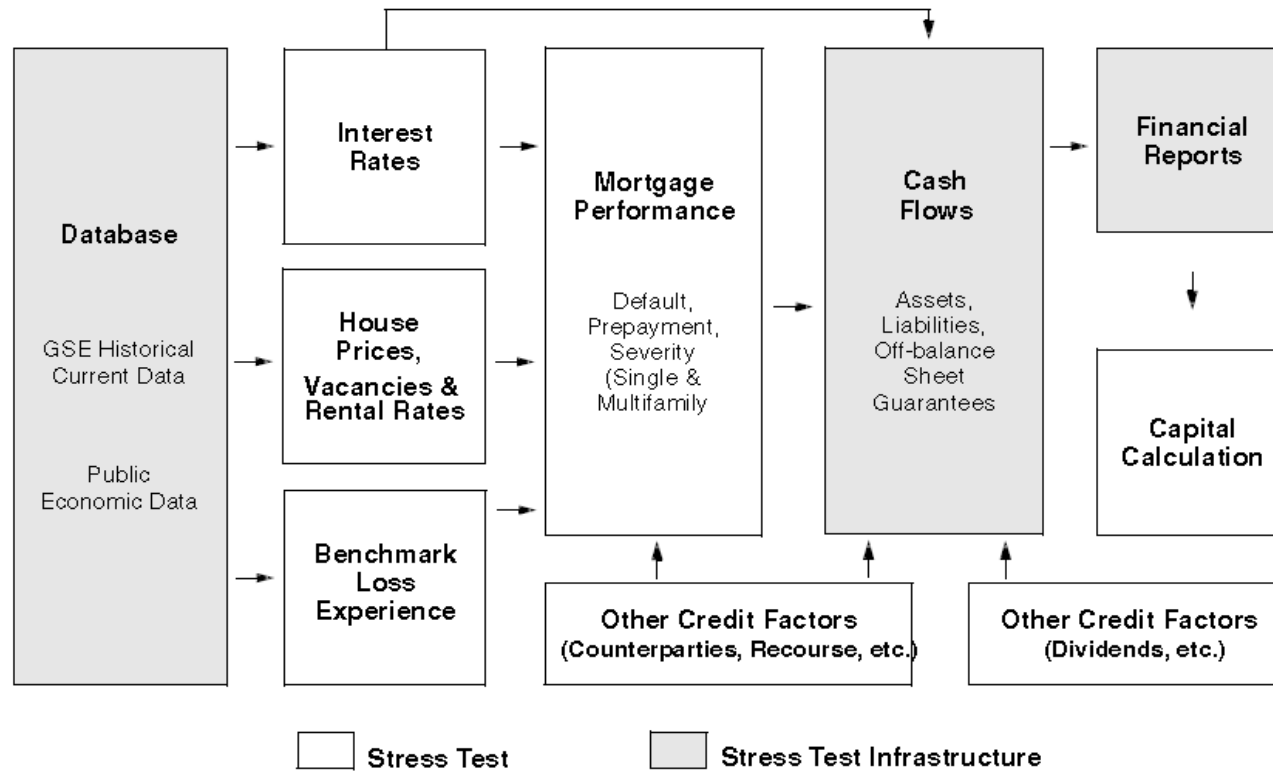
Notes: This table presents annual comparisons between the OFHEO and LPS datasets for three key mortgage contract terms: loan-to-value ratio (LTV), unpaid principal balance (UPB), and interest rate. Sample average values for new originations are provided separately for Fannie Mae (Panel A) and Freddie Mac (Panel B) for each year 1995 through 2005. The OFHEO data is based on the population of single-family mortgages purchased or guaranteed by each GSE. LPS data reflects loans identified in the data as being held or guaranteed by each GSE.

**Table 2: Comparison of Estimates from Default and Prepayment Hazard Models for 30-Year Fixed Rate Mortgages as Specified in the Risk-based Capital Stress Test**

Variable	Default Hazard Estimates		Prepayment Hazard Estimates	
	OFHEO (1979 – 1999)	LPS (1994 – 2000)	OFHEO (1979 – 1999)	LPS (1994 – 2000)
<i>Loan-to-Value (LTV)</i>				
$LTV \leq 60$	-1.150	-0.983	0.048	-0.035
$60 < LTV \leq 70$	-0.104	-0.057	-0.031	-0.095
$70 < LTV \leq 75$	0.597	0.473	-0.099	-0.077
$75 < LTV \leq 80$	0.224	0.109	-0.041	-0.050
$80 < LTV \leq 90$	0.200	0.177	-0.005	0.058
$90 < LTV$	0.233	0.280	0.128	0.199
<i>Probability of Negative Equity (PNEQ)</i>				
$0 < PNEQ \leq 0.05$	-1.603	-1.260	0.591	0.694
$0.05 < PNEQ \leq 0.1$	-0.524	-0.371	0.370	0.456
$0.1 < PNEQ \leq 0.15$	-0.181	-0.053	0.229	0.322
$0.15 < PNEQ \leq 0.2$	0.080	0.093	-0.020	0.034
$0.2 < PNEQ \leq 0.25$	0.255	0.240	-0.160	-0.212
$0.25 < PNEQ \leq 0.3$	0.515	0.459	-0.246	-0.386
$0.3 < PNEQ \leq 0.35$	0.652	0.426	-0.294	-0.357
$0.35 < PNEQ$	0.806	0.466	-0.464	-0.551

Notes: This table compares parameter estimates associated with discrete measures of the loan-to-value ratio and probability of negative equity produced by OFHEO-defined default and prepayment hazard models for 30-year fixed rate mortgages. OFHEO estimates reflect those provided in the risk-based capital stress test, which were produced using a large sample of loans from Fannie Mae and Freddie Mac between 1979 and 1999. LPS estimates are produced using a sample of loans identified as owned or guaranteed by Fannie Mae or Freddie Mac between 1994 and 2000.

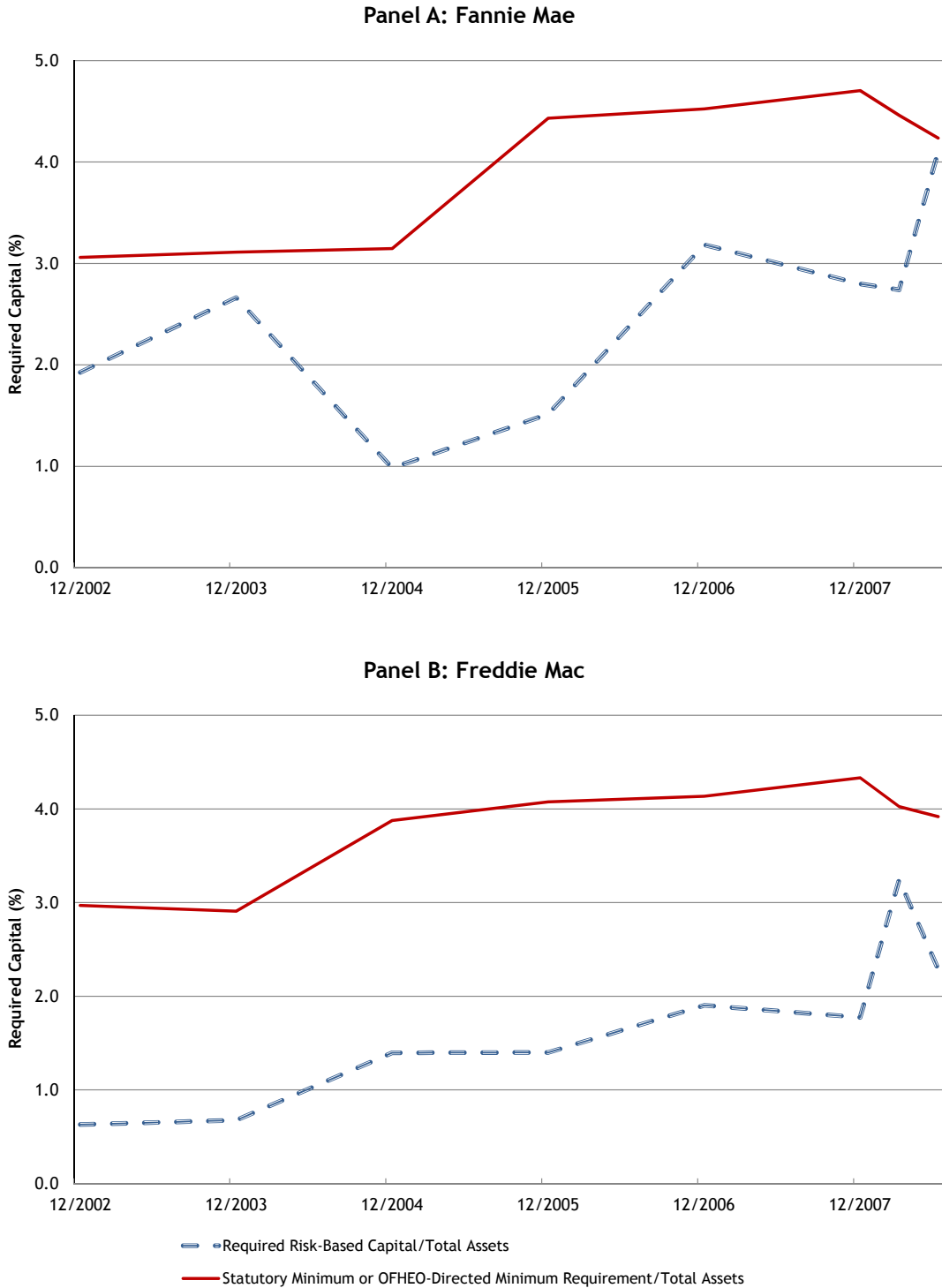
Figure 1: OFHEO Risk-Based Capital Stress Test Framework



Notes: This figure illustrates the process used by OFHEO to calculate required risk-based capital for Fannie Mae and Freddie Mac. Database represents all of the GSEs historical data pertaining to conforming mortgages acquired (held or guaranteed), as well as mortgage insurance, investment securities, liabilities, and derivatives. It also includes public economic data, such as interest rates and house price indices. Interest Rates and House Prices represent the specific series and 40 quarter paths assumed by OFHEO under the stress test. Benchmark Loss Experience represents the adjustments made by OFHEO to the mortgage performance models to equate them to the benchmark loss experience. Mortgage Performance represents the set of default and prepayment models developed by OFHEO. Cash Flows combines estimates generated by the mortgage performance models with those from other assets, liabilities, and off-balance sheet exposures. Some of these cash flows are discounted by Other Credit Factors, which account for counterparty credit risk. The resulting net flows are aggregated into quarterly Financial Reports from which Capital Calculations are derived.

Source: Office of Federal Housing Enterprise Oversight

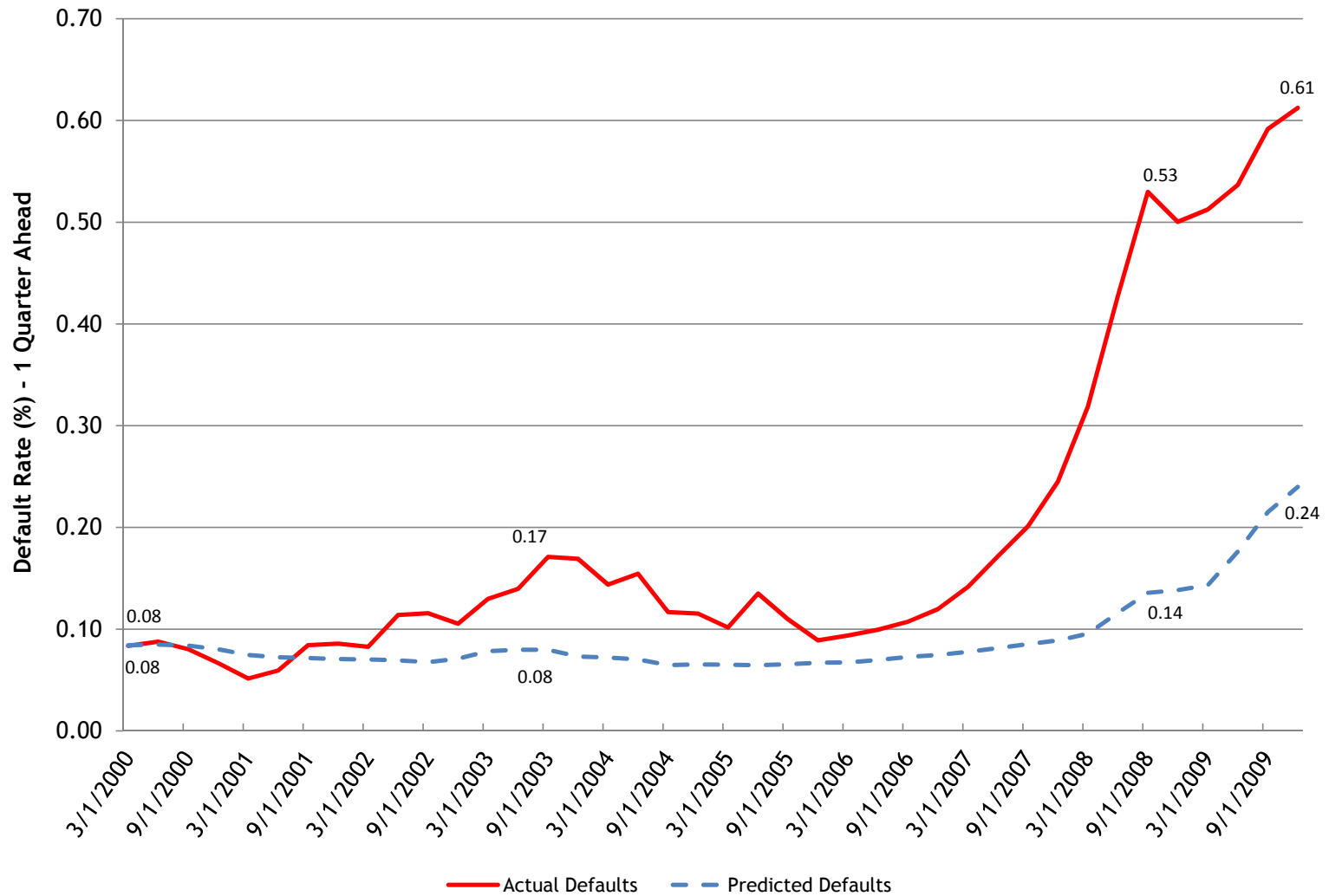
Figure 2: GSE Required Minimum and Risk-Based Capital Requirements



Notes: This figure illustrates the required minimum and risk-based capital requirements (as a percent of total assets) for Fannie Mae (Panel A) and Freddie Mac (Panel B). Minimum capital is simply computed as 2.50 percent of total assets plus 0.45 percent of off-balance sheet guarantees. Minimum capital also includes surcharges. Freddie Mac faced a 30 percent surcharge between 2004:Q1 and 2007:Q4, which was reduced to 20 percent thereafter. Fannie Mae faced a 30 percent surcharge between 2005:Q3 and 2007:Q4, which was reduced to 20 percent thereafter. Risk-based capital is computed by the stress test as represents the larger of the results produced under the two interest rate scenarios (“up rate” and “down rate”).

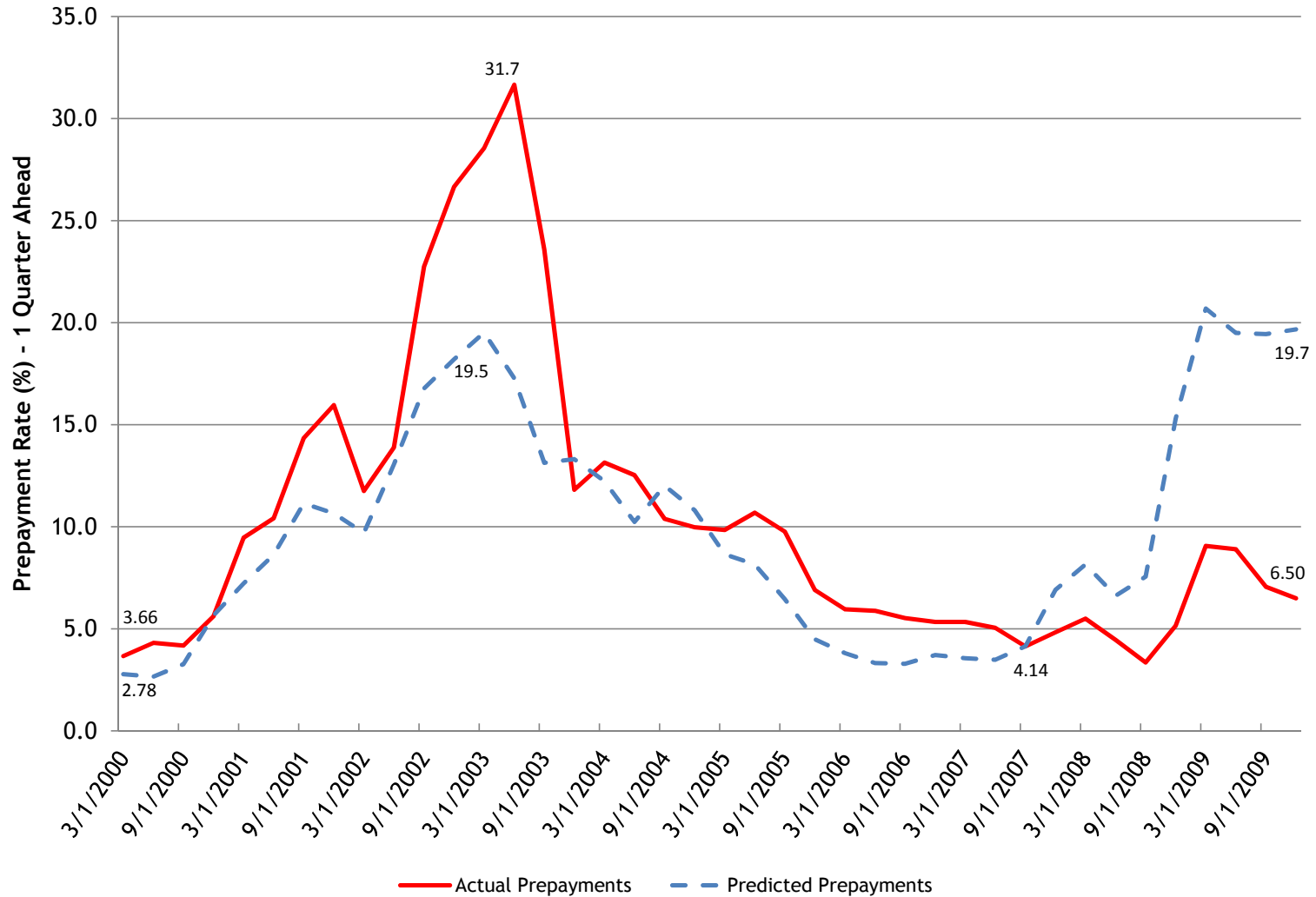


Figure 3: Actual vs. Predicted Defaults Using OFHEO's Static Model for 30-Year FRMs



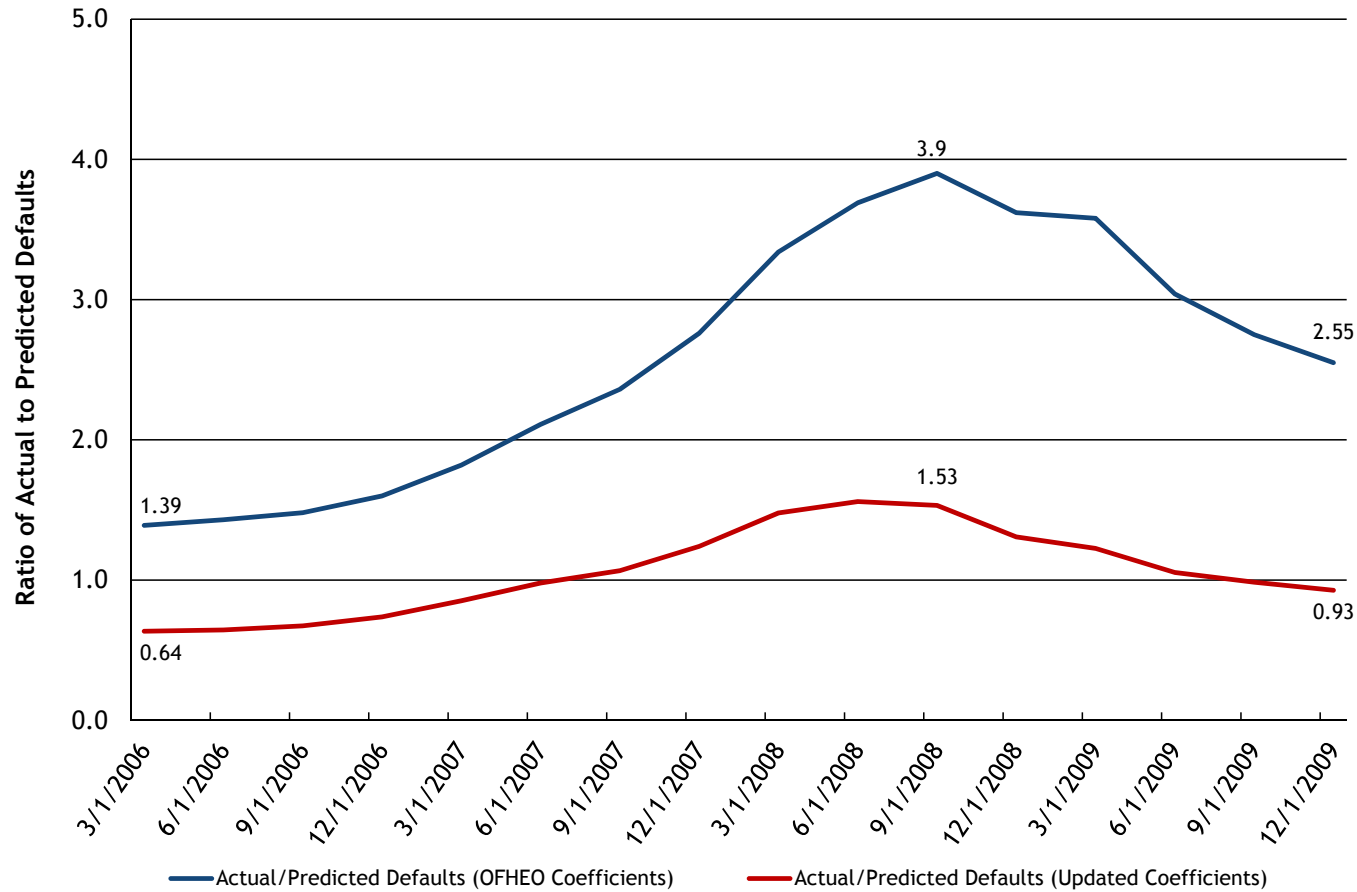
Notes: This figure presents the actual and predicted quarter-end default rate on 30-year fixed rate mortgage loans between 2000:Q1 and 2009:Q3. Actual default rates are based on mortgages identified as being held by Fannie Mae or Freddie Mac in the LPS data. Predicted default rates represent the prior quarters one-quarter-ahead forecasts using the same LPS data projected through the parameterized default model published as part of the OFHEO risk-based capital stress test.

Figure 4: Actual vs. Predicted Prepayments Using OFHEO's Static Model for 30-Year FRMs



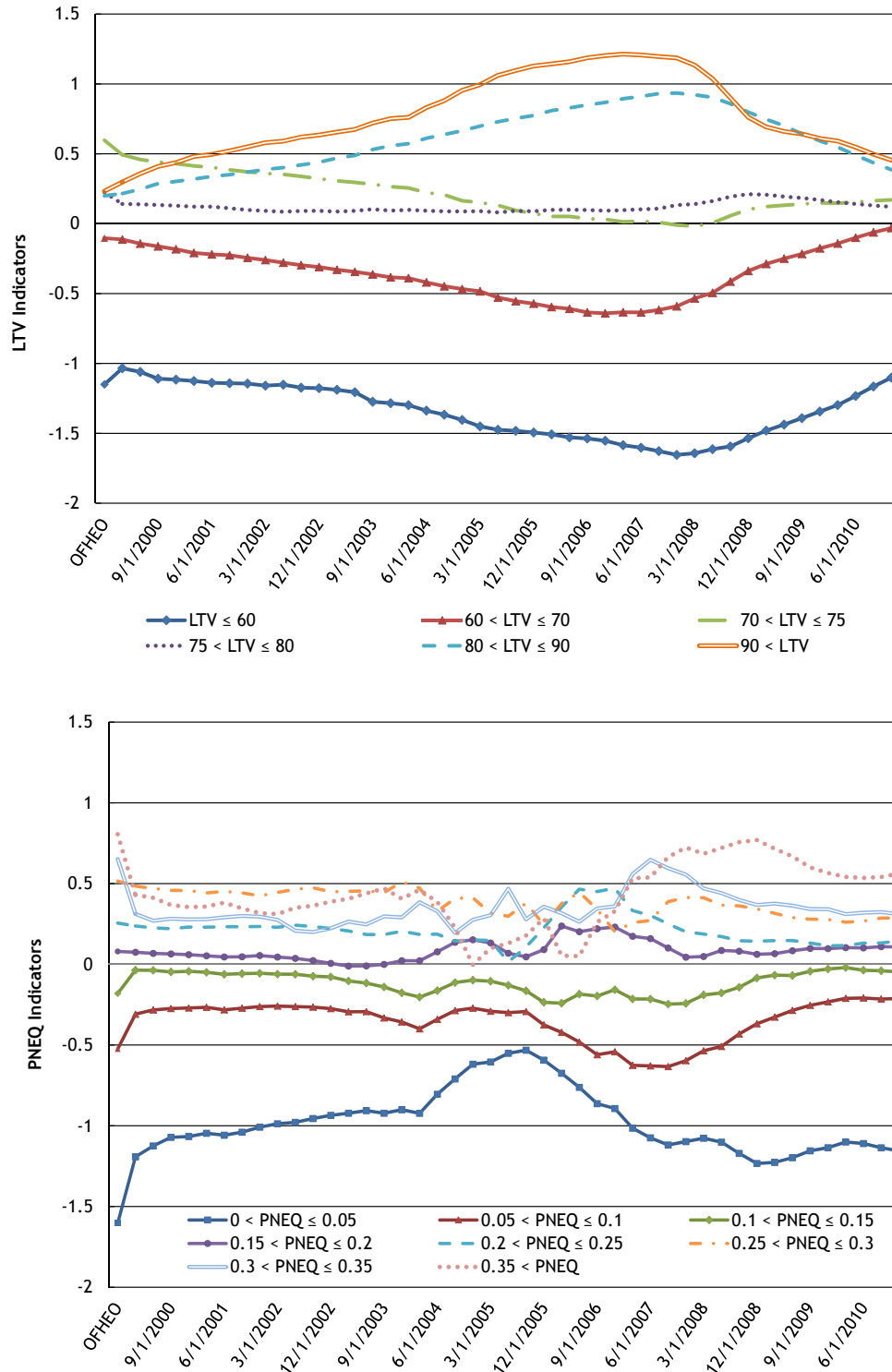
Notes: This figure presents the actual and predicted quarter-end prepayment rate on 30-year fixed rate mortgage loans between 2000:Q1 and 2009:Q3. Actual prepayment rates are based on mortgages identified as being held by Fannie Mae or Freddie Mac in the LPS data. Predicted prepayment rates represent the prior quarters one-quarter-ahead forecasts using the same LPS data projected through the parameterized prepayment model published as part of the OFHEO risk-based capital stress test.

Figure 5: Ratio of Actual to Predicted Defaults Using One-Quarter-Ahead Default Forecasts from OFHEO’s Static and Updated Models



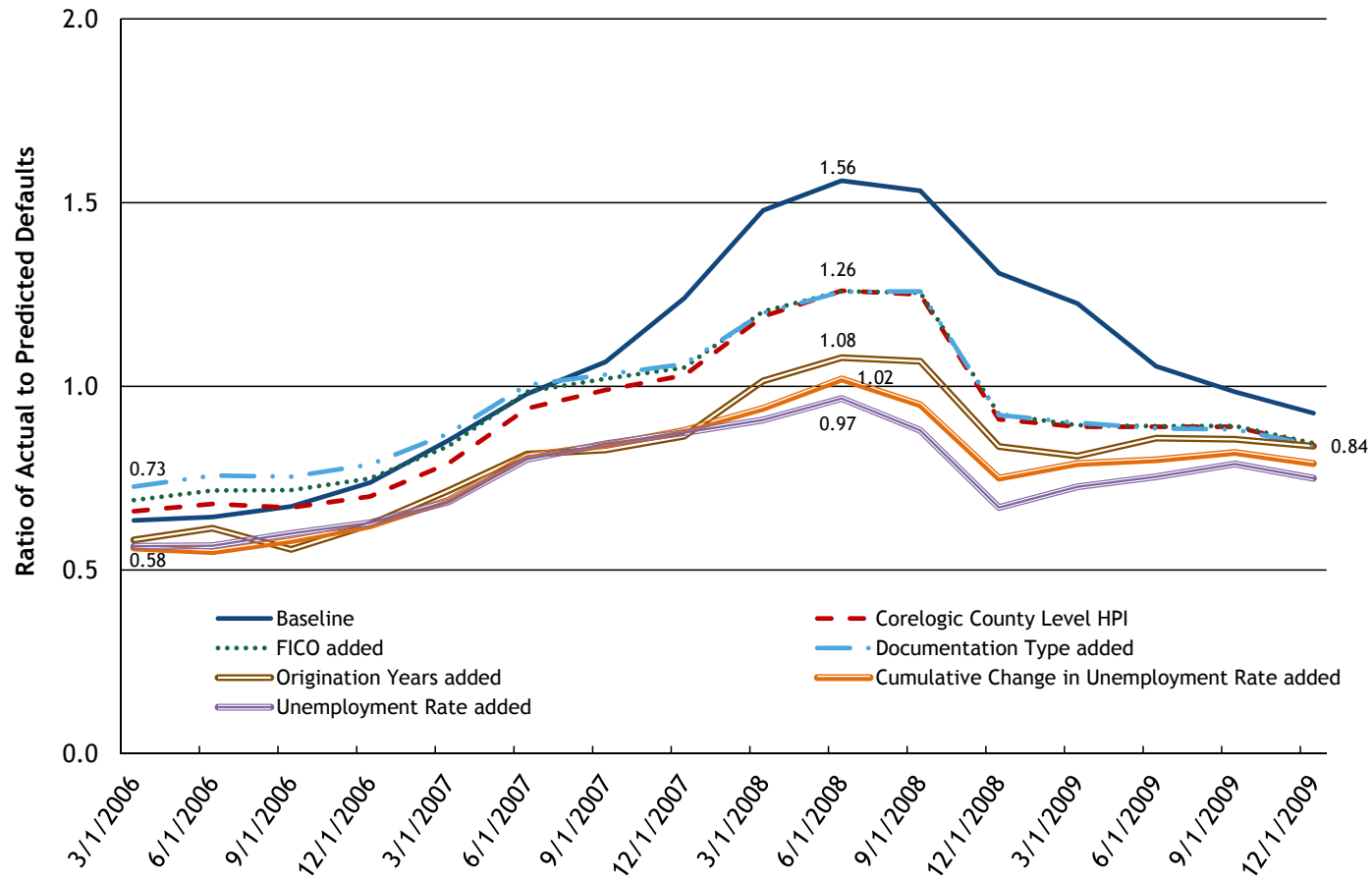
Notes: This figure presents the ratio of actual to predicted defaults using one-quarter-ahead default forecasts from 2006:Q1 through 2009:Q4 based on the published OFHEO parameters (“static model”) and the same model re-estimated quarterly using a seven-year rolling window (“updated model”). The first window begins in 1999:Q1 and runs through 2005:Q4 Both models estimated using 30-year fixed rate mortgages identified as being held by Fannie Mae or Freddie Mac in the LPS data.

Figure 6: Evolution of Default Hazard Parameter Estimates



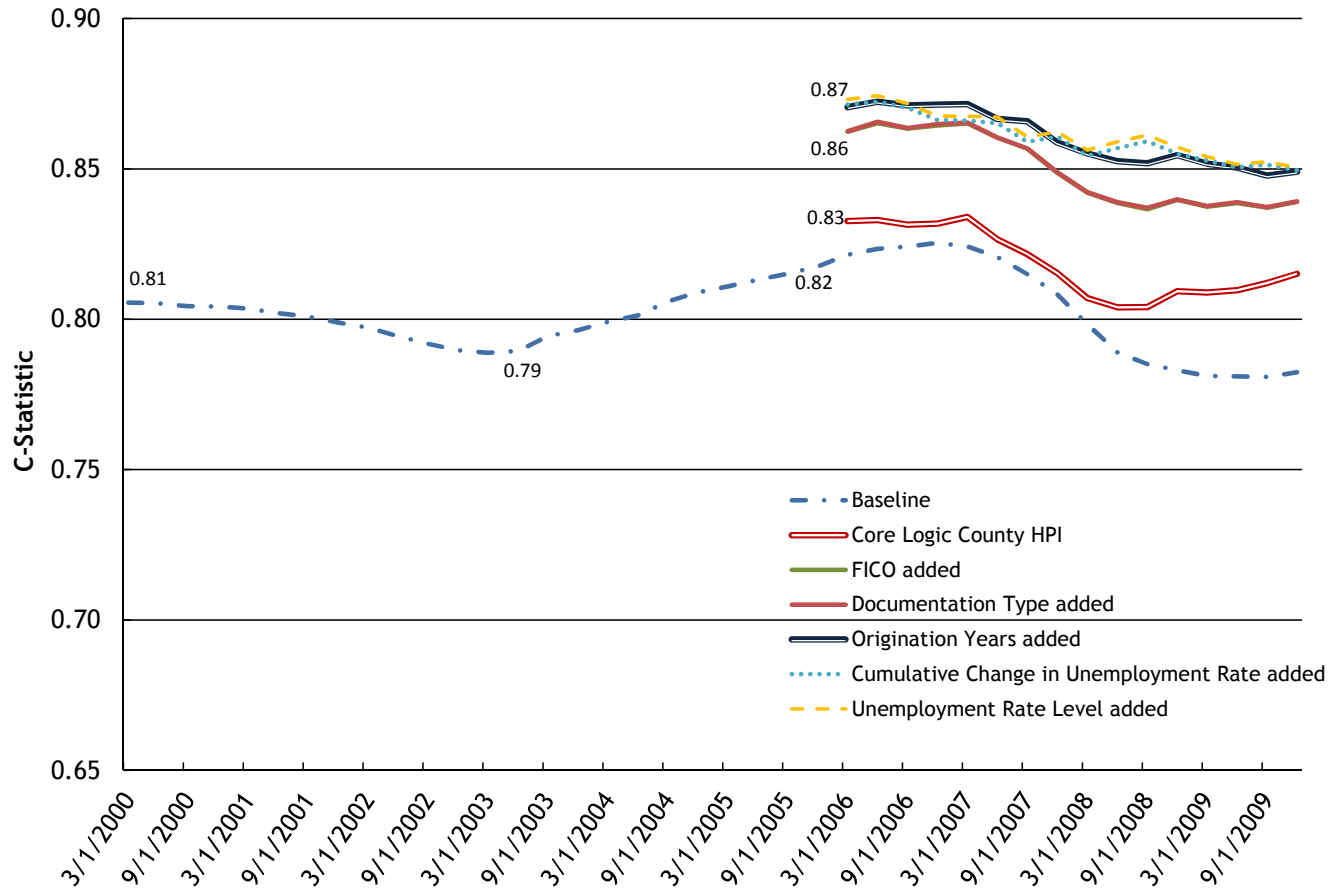
Notes: This figure displays the evolution of coefficient estimates associated with the LTV and PNEQ variables in the default hazard for the updated OFHEO model. All LTV and PNEQ variables are expressed as indicator variables in the model. The CATMOD procedure was used to obtain estimates for all values of the categories.

Figure 7: Ratio of Actual to Predicted Defaults Using One-Quarter-Ahead Default Forecasts from the Updated OFHEO



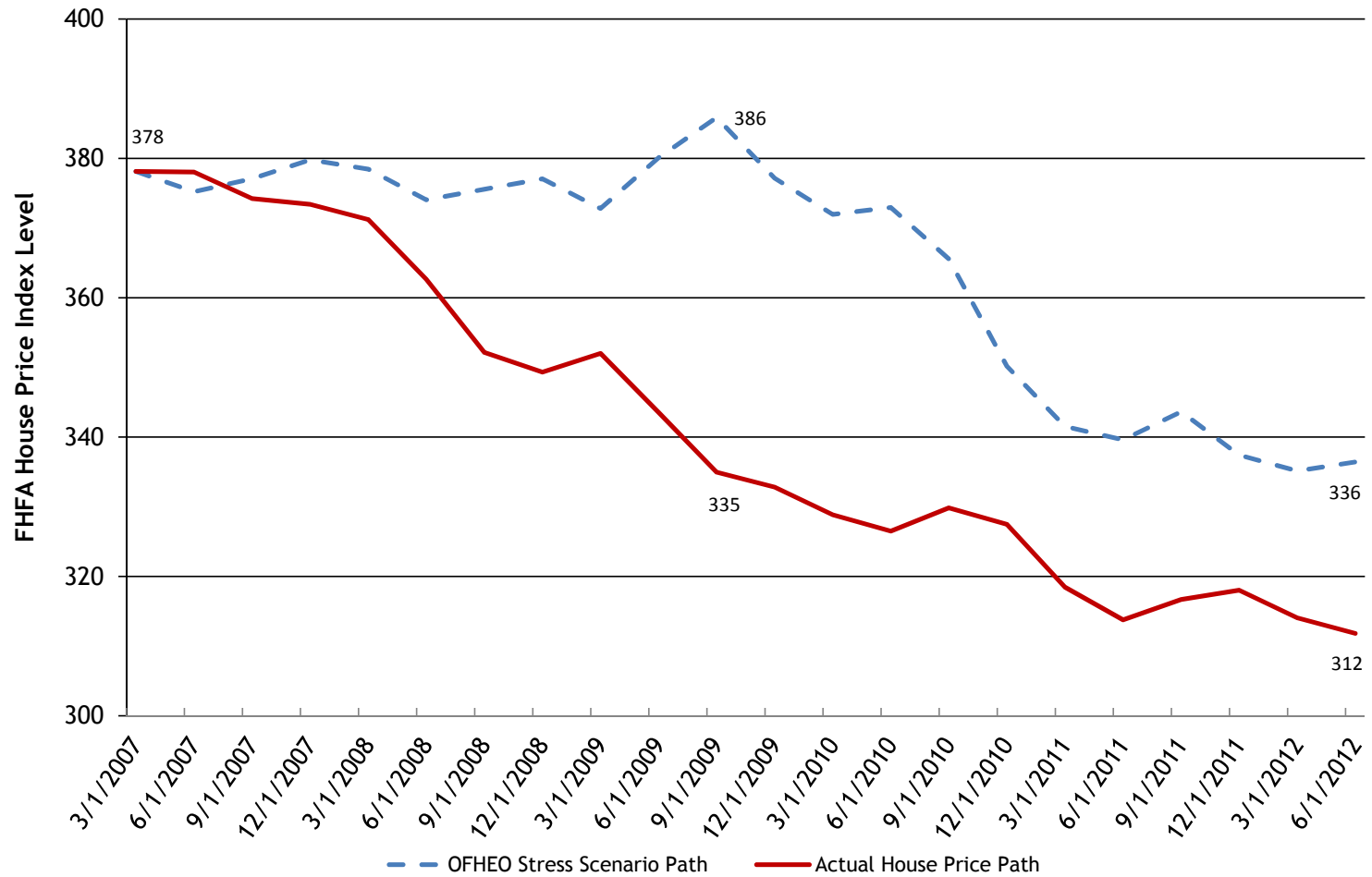
Notes: This figure presents the ratio of actual to predicted defaults using one-quarter-ahead default forecasts from 2006:Q1 through 2009:Q4 based on the OFHEO model re-estimated quarterly using a seven-year rolling window and introducing new covariates. Baseline represents the forecast error ratio for the OFHEO updated model. The model is then cumulatively supplemented by CoreLogic County-Level HPI, FICO, Documentation Type, and Origination Years. All models are estimated using 30-year fixed rate mortgages identified as being held by Fannie Mae or Freddie Mac in the LPS data.

Figure 8: In-Sample Fit (Default): Area under the Receiver Operating Characteristic Curve (C-Statistic)



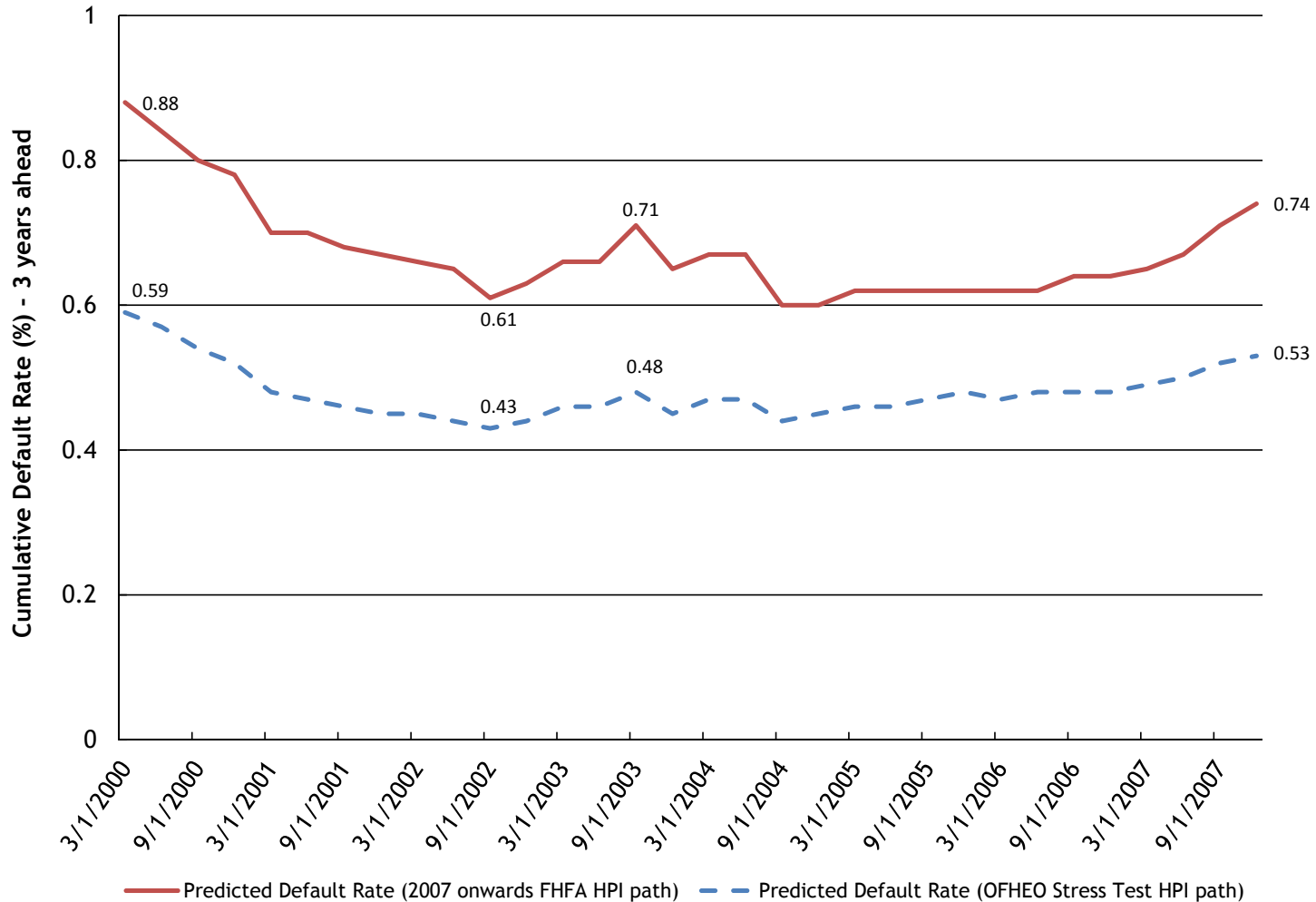
Notes: This figure presents quarterly C-Statistics, or the area under the receiver operating curve, based on the OFHEO updated model re-estimated quarterly using a seven-year rolling window and introducing new covariates. Baseline represents the C-Statistic for the OFHEO updated model 2000:Q1 through 2009:Q4. The model is then cumulatively supplemented by CoreLogic County-Level HPI, FICO, Documentation Type, and Origination Years. Quarterly C-Statistics for these additional models are for 2006:Q1 through 2009:Q4. All models are estimated using 30-year fixed rate mortgages identified as being held by Fannie Mae or Freddie Mac in the LPS data.

Figure 9: Comparison of OFHEOs House Price Stress Scenario to the Actual Path of U.S. House Prices During the Recent Housing Bust



Notes: This figure compares the first five years of OFHEOs house price stress scenario to the actual path of U.S. house prices during the recent housing bust 2007:Q1 through 2012:Q1. OFHEOs house price stress was the path of house prices in the West South Central Census Division. The recent path of U.S. house prices is measured by the OFHEO National House Price Index.

Figure 10: The Effect of House Prices on Quarterly Predicted Cumulative 3-Year Default Rates Using OFHEOs Static Model: OFHEOs Stress Test versus the Recent Housing Bust



Notes: This figure examines the effect of house prices on quarterly predicted cumulative 3-year default rates using OFHEOs static model for 2000:Q2 2007:Q4 Estimates of quarterly 3-year cumulative defaults are based on the path of house prices from the West South Central Census Division from 1984 through 1986 as well as the path of U.S. house prices from 2007 through 2009.