

FHFA STAFF WORKING PAPER SERIES

A Quarter Century of Mortgage Risk

Morris A. Davis
William D. Larson
Stephen D. Oliner
Benjamin R. Smith

October 2019 (revised)
January, 2019 (original)

Working Paper 19-02

FEDERAL HOUSING FINANCE AGENCY
Division of Housing Mission & Goals
Office of Policy Analysis & Research
400 7th Street SW
Washington, DC 20219, USA



Working Papers prepared by staff of the Federal Housing Finance Agency (FHFA) are preliminary products circulated to stimulate discussion and critical comment. The analysis and conclusions are those of the authors alone, and should not be represented or interpreted as conveying an official FHFA position, policy, analysis, opinion, or endorsement. Any errors or omissions are the sole responsibility of the authors. References to FHFA Working Papers (other than acknowledgment) should be cleared with the authors to protect the tentative character of these papers.

Please address correspondence to William Larson (william.larson@fhfa.gov).

A Quarter Century of Mortgage Risk

Morris A. Davis, William D. Larson, Stephen D. Oliner, Benjamin R. Smith
FHFA Staff Working Paper 19-02

October 2019 (revised)

January, 2019 (original)

Abstract

This paper provides a comprehensive account of the evolution of default risk for newly originated home purchase loans over the past quarter century. We bring together several data sources to produce this history, including loan-level data for the entire Enterprise (Fannie Mae and Freddie Mac) book. We use these data to track a large number of loan characteristics and a summary measure of risk, the stressed default rate. Among the many results in the paper, we show that mortgage risk had already risen in the 1990s, planting seeds of the financial crisis well before the actual event. Our results also cast doubt on explanations of the crisis that focus on low-credit-score borrowers. The aggregate series we present in this paper are available for download at <https://www.fhfa.gov/papers/wp1902.aspx>.¹

Keywords: mortgage risk · housing boom · default · foreclosure · house price · leverage

JEL Classification: E32, G21, G28, H22, R31

Morris A. Davis
Department of Finance and Economics
Rutgers Business School
Rutgers University
1 Washington Park Room 627
Newark, NJ 07102
mdavis@business.rutgers.edu

William D. Larson
Federal Housing Finance Agency
Office of Policy Analysis & Research
400 7th Street SW
Washington, DC 20219, USA
william.larson@fhfa.gov

Stephen D. Oliner
American Enterprise Institute for Public
Policy Research
1789 Massachusetts Avenue, NW
Washington, DC 20036
stephen.oliner@aei.org

Benjamin R. Smith
The Wharton School
University of Pennsylvania
3733 Spruce Street
Philadelphia, PA 19104
benjamrs@wharton.upenn.edu

¹These indices are works in progress and all data, tables, figures, and other results in this working paper are subject to change. Earlier versions of this paper were posted in January and March 2019 under the title “Mortgage Risk Since 1990.”

1. Introduction

Since the global financial crisis, there has been an outpouring of research to understand the developments in the U.S. home mortgage market that precipitated the crisis.¹ Nonetheless, as valuable as this research has been, there is still no comprehensive account of the changes in mortgage risk that produced the worst foreclosure wave since the Great Depression.

This paper is an effort to fill that gap, covering essentially the entire market for home purchase loan originations in the United States from 1990 to 2018.² We bring together several data sources to construct this comprehensive historical picture, including the full set of home mortgages guaranteed by the government sponsored enterprises Fannie Mae and Freddie Mac (the Enterprises). To our knowledge, this is the first time the entire Enterprise book has been used in publicly-available research on mortgage risk.³ We supplement the Enterprise dataset with data covering more than 90 percent of the loans in private mortgage-backed securities and data from mortgage servicers for a large fraction of loans in the rest of the market (loans guaranteed by the Federal Housing Administration (FHA) and the Department of Veterans Affairs (VA) and unguaranteed loans held in the portfolio of banks and other lenders).⁴

We track many loan characteristics that, taken together, are highly predictive of loan performance. These characteristics include the borrower's credit score, the debt-payment to

¹ For readers seeking to explore this vast literature, the appendix table in Ferreira and Gyourko (2015) lists more than 30 papers on various aspects of the mortgage market. Foote and Willen (2018) discuss the post-crisis surge of research on the determinants of mortgage default. And Mian and Sufi (2018) cover research on the expansion of mortgage credit supply during the housing boom and its implications for house prices.

² Although this version of the paper focuses on home purchase loans, we are in the process of revising the paper to include refinance loans.

³ The Enterprise data source is the FHFA Mortgage Loan Information System (MLIS). No personally identifiable information is contained in the tables, charts, or the database associated with this paper. Results presented pertaining to these data are aggregates (including model outputs) and rounded where appropriate.

⁴ For FHA and VA loans originated during 2013-2018, we also have Ginnie Mae data provided by the American Enterprise Institute (AEI) Housing Center. Because the Ginnie Mae data maintained at AEI cover virtually all FHA and VA loans, we use these data for 2013-2018 originations in place of the servicer data.

income ratio (DTI), the combined loan-to-value ratio (CLTV) that accounts for any subordinate liens at origination, loan type (fixed or adjustable mortgage rate), loan term, loan purpose, whether the borrower's income is fully documented, and whether the mortgage has a feature that reduces the paydown of loan principal (such as a period of interest-only payments).

Given the multitude of loan characteristics, it is important to develop a publicly available summary measure of default risk. This measure should convey the risk of default in a simple, straightforward manner, and to be of most use to policymakers with prudential oversight responsibilities, it should focus on default under severely stressed conditions. To this end, we calculate what we call the "stressed default rate" for all loans in the dataset based on the observed default experience of similar loans originated nationwide in 2006 and 2007, just before the financial crisis. The stressed default rate for a given loan thus represents its expected default rate if it were hit shortly after origination with a replay of the financial crisis, including the observed national average decline in house prices and ensuing policy responses. Because the drop in house prices varied enormously across localities, the stressed default rate calculated in this way will only be valid for a national portfolio of loans. As an extension, we also calculate a stressed default measure for states and metropolitan areas. This extension embeds the estimated magnitude of a severe local house price shock in each year using the framework in Smith and Weiher (2012) and Smith et al. (2016).

The stressed default rates presented in this paper build on those currently published by the AEI Housing Center and the Urban Institute. AEI unveiled its National Mortgage Risk Index (NMRI) in 2013, while the Urban Institute introduced its Housing Credit Availability Index (HCAI) a year later in 2014. The NMRI uses highly accurate and complete data for government-guaranteed loans but only goes back to late 2012 and does not cover private portfolio loans or

loans in private mortgage-backed securities. The HCAI covers the entire market back to 1998 but misses developments over 1990-1997 that we cover, uses incomplete data for pre-2013 Enterprise loans, pays less attention to imputing missing loan information, and has a much less granular system for assigning stressed default rates.⁵ In short, we substantially improve the scope and quality of stressed default measures that already have a wide audience.

Our paper makes numerous contributions to the literature. The first, and most basic, is that we document the characteristics of the entire purchase loan market back to 1990 and provide a summary measure of risk going back almost as far. As a result, researchers and policymakers now have at their disposal more complete and more accurate historical information about the mortgage market.⁶ All the data presented in the paper and many other series are available for free download from the FHFA website at <https://www.fhfa.gov/papers/wp1902.aspx>.⁷

Among the insights provided by the longer historical analysis is that seeds of the financial crisis were planted in the 1990s. We show that across the entire market, the average CLTV on new home purchase loans increased substantially in the first half of the 1990s, and the average DTI trended up from 1993 to 2000. By 2000, the stressed default rate for the purchase loan market was already somewhat above the levels seen in the 1990s, with the largest increase for loans bundled into private-label securities (PLS). This finding calls into question the common

⁵ For more information on the NMRI and the HCAI, see <https://www.aei.org/housing/mortgage-risk-index/> and <https://www.urban.org/policy-centers/housing-finance-policy-center/projects/housing-credit-availability-index>. The HCAI is based on analysis in Li and Goodman (2014).

⁶ From the outset, we wish to make clear that while we focus on mortgage risk, there may be substantial benefits to increased access to credit. It is our hope that the metrics in this paper provide the basis for future benefit-cost analysis of policy decisions.

⁷ The available data include (a) annual series for all of the risk factors at the national level that go into the stressed default rate, (b) the underlying default tables and adjustment factors used to construct stressed default rates, (c) loan counts in the cells of these default tables, and (d) the annual series for the baseline stressed default rate and the extension that incorporates potential shocks to house prices at the state and metro area level.

view that mortgage lending conditions were normal in the early 2000s (CoreLogic, 2017; Goldman Sachs, 2014; Urban Institute, 2018).⁸

Although the stressed default rate cannot identify the influence of credit supply versus credit demand on mortgage risk, we provide separate evidence that credit supply expanded in the early and mid-2000s. We do so by documenting a compression of mortgage rate spreads between the riskiest and least risky loans. To cleanly identify changes in credit supply, we limit the analysis to loans for which lenders and investors fully bear the credit risk – private sector loans with no government guarantee and no private mortgage insurance. To our knowledge, this is the first exploration of spreads across risk tiers.⁹

Our results contribute as well to the ongoing debate about the role of “subprime” borrowers in the housing boom and bust.¹⁰ Mian and Sufi (2009, 2014, 2017a) have focused on an increase in credit availability for subprime borrowers, which they define as borrowers with a credit score less than 660, while several recent papers have argued that the credit expansion was more widespread (Adelino, Schoar, and Severino, 2016, 2017; Albanesi, Di Giorgi, and Nosal, 2017; Conklin et al., 2018; Ferreira and Gyourko, 2015). Our results favor the latter view. We find that the low-credit-score share of home purchase loans fell slightly on net from the late

⁸ To be clear, we are not saying that every risk factor for every market segment experienced increases in mortgage risk before the early 2000s. However, combining the various risk factors, the average risk profile of newly-originated purchase-money mortgages had already risen by 2000.

⁹ For earlier discussions of mortgage spreads, see Demyanyk and van Hemert (2011), Justiniano, Primiceri, and Tambalotti (2017), and Mian and Sufi (2017a). Anenberg et al. (2018) complement this evidence by showing that the credit supply frontier – defined as the largest loan available for a given set of loan and borrower characteristics – shifted up during the early and mid-2000s.

¹⁰ Throughout the paper, we define low-credit-score borrowers as those with a credit score of less than 660 at origination. Such borrowers are often referred to as “subprime” in the academic literature. In other contexts, there is no agreed-upon definition of the term. Accordingly, we refrain from using the subprime label throughout the paper except in reference to others’ work.

1990s to the mid-2000s and that the stressed default rate for these loans rose in line with that of loans to higher-score borrowers.

Another question of interest is the role played by loans with risky product features in the buildup of risk during the boom. We define such loans to be those having one or more of the following features: low or no documentation of borrower income, less than full amortization of loan principal, and a term greater than 30 years.¹¹ We find that these risky product features accounted for slightly more than half of the overall rise in the stressed default rate, with more standard forms of borrower leverage such as high payment burdens and small down payments accounting for the rest. Even though risky product features are largely absent from the mortgage market today, standard forms of leverage boosted risk a lot during the boom.

The final part of the paper brings house price risk explicitly into the analysis. We use a method based on Smith and Weiher (2012) and Smith et al. (2016) to construct severe potential house price shocks. The shocks vary over time and geography based on FHFA house price indices for states and metropolitan areas. We adjust each loan's reported CLTV at origination to reflect the lower house value if the shock were realized. With the large house price increases starting in the late 1990s, these "shock CLTVs" move further and further above the reported CLTVs. The average shock CLTV peaks in 2006 at nearly 140 percent, illustrating the enormous house price risk borne by lenders and mortgage insurers at that time.

2. Data

We rely on several sources of loan-level data over the period 1990-2018. For Enterprise loans, we use the internal, non-public data maintained by FHFA, the Mortgage Loan Information System (MLIS). The MLIS dataset covers all mortgage loans acquired by the Enterprises,

¹¹ This definition adopts some of the key product features that make a loan ineligible for Qualified Mortgage status under the Dodd-Frank Act and regulations adopted by the Consumer Financial Protection Bureau.

providing detailed information about loan characteristics and performance. As noted in the introduction, we believe our paper is the first to use the full book of Enterprise loans over such a long history.¹² The internal FHFA data have important advantages over the public-use loan-level data made available by the Enterprises, which are limited to fixed-rate mortgages with full documentation and full amortization of loan principal that were not part of designated affordable housing programs.¹³ By omitting adjustable-rate mortgages, loans with risky product features (such as low-doc loans), and many loans made under affordable housing programs, the publicly available datasets understate the risk in Enterprise-guaranteed loan portfolio, especially during the housing boom.¹⁴

For loans securitized in the private market without an Enterprise guarantee, we use the CoreLogic non-agency residential mortgage-backed securities dataset. The CoreLogic dataset is the most comprehensive source for loans included in private-label securities. It covers all segments of the PLS market, with the loans in the dataset accounting for more than 90 percent of the entire market. Many previous studies have used this dataset, including Adelino, Frame, and Gerardi (2017), Demyanyk and Van Hemert (2011), Houghwout, Peach, and Tracy (2008), Keys et al. (2010), and Mayer, Pence and Sherlund (2009). Our analysis builds on these studies by constructing much longer historical time series.¹⁵

For loans held in portfolio by private lenders, we use the Loan-Level Market Analytics (LLMA) dataset from CoreLogic and the McDash dataset from Black Knight, Inc. Both datasets

¹² Other research has studied loan performance using internal Fannie Mae or Freddie Mac data over shorter periods. For example, Fout et al. (2018) use internal Fannie Mae loan data over 2002-2013, while Firestone, Van Order, and Zorn (2007) use internal Freddie Mac data over 1993-1997.

¹³ The publicly available data are Fannie Mae's Single-Family Loan Performance Dataset and Freddie Mac's Single-Family Loan-Level Dataset, which are posted at <http://www.fanniemae.com/portal/funding-the-market/data/loan-performance-data.html> and http://www.freddiemac.com/research/datasets/sf_loanlevel_dataset.html, respectively.

¹⁴ The publicly available datasets exclude 45 percent of all Enterprise home purchase loans originated during 2000-2007, with especially poor coverage of high LTV loans.

¹⁵ Our dataset for PLS loans currently ends in 2017; the next version of the paper will add the 2018 data.

compile information provided by large loan servicers, though the set of servicers differs across the two datasets. Studies that have used one of the servicer datasets include Bubb and Kaufman (2014), Courchane, Kiefer, and Zorn (2015), DeFusco, Johnson, and Mondragon (2017), Foote et al. (2010), and Keys, Seru, and Vig (2012), among others. In contrast to these studies, we merge the two datasets to provide broader coverage than with either one alone.

The identification of portfolio loans in the servicer datasets is challenging because one has to separate these loans from other conventional loans that were purchased by the Enterprises or securitized in the PLS market. Both datasets include investor codes to identify whether a loan was acquired by the Enterprises, but the information in this field is incomplete in both LLMA and McDash. We use the procedure described in Appendix A to remove loans that are reported to have been acquired by the Enterprises or are likely to be Enterprise loans. In addition, both datasets have limited information on whether a loan was securitized in the PLS market. We remove suspected PLS loans through loan-level matches to the PLS dataset. To be confident that we are only retaining portfolio loans, we purge from the dataset not only perfect one-to-one matches but other loans that are likely to be PLS loans. By drawing a wide net around potential Enterprise and PLS loans, some true portfolio loans inevitably will be removed as well.¹⁶

For FHA and VA loans, we use the two servicer datasets through 2012 and then switch over for 2013-2018 to the National Mortgage Risk Index (NMRI) dataset maintained by the AEI Housing Center. The NMRI dataset is a near-census of government guaranteed loans that

¹⁶ These identification challenges, along with McDash's limited market coverage in the 1990s, hold down the size of our portfolio dataset. We estimate that our coverage of portfolio loans peaks at roughly 60 percent in 2006 and is considerably lower in the 1990s. To correct for potential biases in the data, we weight the portfolio loans to be nationally representative, as described below.

provides detailed information about risk characteristics at origination. It is more comprehensive than the servicer datasets but is only available starting in late 2012.¹⁷

For all of these market segments, we use data on first-lien, 1-4 unit home purchase loans that pass a variety of data quality filters. These filters remove, for example, loans that lack information on such variables as origination year, loan term, and product type (ARM vs. fixed-rate).¹⁸ In all, the filters remove relatively small shares of the source data: about 1-2 percent of Enterprise and PLS loans, 5 percent of FHA/VA loans, and 19 percent of the private-sector loans that are candidates for inclusion in our portfolio loan dataset.¹⁹ After applying the filters and removing duplicate loans, the dataset includes roughly 93 million home purchase loans originated from 1990 through 2018.

Despite using the best available mortgage data, information on key risk factors is missing for many loans. This is especially true for the early years of our analysis, when the reporting of credit scores and DTIs is very spotty. Although we could simply drop observations with missing data, we would lose the non-missing data fields that contain useful information on the likelihood of default. Furthermore, this information is not missing solely by chance, as the likelihood of a missing risk factor is correlated with observable borrower and loan characteristics. Hence, dropping observations with missing information would bias the sample and resulting statistics.

¹⁷ The NMRI dataset covers not only FHA and VA loans, but also loans guaranteed by the Rural Housing Service (RHS), which are not included in this analysis, and loans securitized by the Enterprises. The coverage of Enterprise loans is essentially the same as in FHFA's internal dataset, but the available history is much shorter. The source data for FHA, VA, and RHS loans are posted by Ginnie Mae at https://www.ginniemae.gov/data_and_reports/disclosure_data/pages/disclosure_history.aspx.

¹⁸ See Appendix A for full detail on these data quality filters and all other aspects of our data preparation.

¹⁹ These private-sector loans are only candidates for the portfolio dataset because we still have to remove PLS loans and suspected Enterprise loans that remain after our initial screening. In addition, because we merge data from the two servicer datasets to arrive at the candidate portfolio loans and the set of FHA/VA loans, there are many duplicates that must be removed. The percentages cited here represent the shares of loans in both groups that are removed by the data quality filters over and above the removal of duplicate loans.

To be able to use loans with incomplete information in our analysis, we impute missing values for five risk factors: credit score, DTI, loan documentation status, amortization status, and occupancy status. Table 1 shows the share of each of these risk factors that must be imputed in each market segment. As can be seen, we impute more than half of the DTIs for portfolio, PLS, and FHA/VA loans. In addition, roughly 40 percent of portfolio loans require imputations for credit score and documentation status. For the other cells in Table 1, imputations are less frequent.²⁰

Figure 1 displays the share of imputations by year for the entire set of purchase loans. As shown, nearly all credit scores and DTIs are imputed for 1990-1992. The imputation share for DTIs falls dramatically in 1993 and the share for credit scores drops in 1993, 1994, and 1996. Improved reporting for Enterprise loans largely accounts for these changes. For the other risk factors shown in the figure, the imputed share is relatively low in all years.

The near-complete reliance on imputed DTIs and credit scores in the early 1990s raises concerns about the accuracy of those results. To ensure that our findings are anchored to real data, we do not present results for DTIs and credit scores until the first year in which these risk factors are well reported by the Enterprises, the dominant institutions in the mortgage market. This rule implies that our DTI and credit score results begin in 1993 and 1994, respectively. By 1993, more than 90 percent of Enterprise purchase loans had a reported DTI, up sharply from

²⁰ Table 1 indicates that we do not impute occupancy status for FHA/VA loans or documentation status for Enterprise loans. This requires a bit of explanation. Some FHA/VA loans in the servicer dataset and the Ginnie Mae dataset maintained at AEI have missing data on occupancy status. We assume that these loans are actually for primary owner-occupied properties given that second-home loans and investor loans make up a very small share of the loans with reported occupancy status. A similar issue arises in the FHFA dataset, which shows many Enterprise home purchase loans as having an unknown documentation status. One of the Enterprises has missing documentation status for nearly all loans except those that are affirmatively no/low doc, and both Enterprises have some loans with missing documentations status even in the post-crisis period when all Enterprise loans are required to have full documentation. Accordingly, we assume all Enterprise loans with missing documentation status are full-doc.

less than 10 percent in prior years. In 1994, one of the two Enterprises began to systematically report credit scores, with the other following suit in 1996. Because the credit profiles for the two Enterprises are similar, the data available as of 1994 provide a very good estimate for the full Enterprise book. Given the central role of credit scores as a risk factor, we begin reporting the stressed default rate in 1994. Anchoring to well-reported Enterprise data supports the accuracy of our results for the purchase loan market as a whole, even though many non-Enterprise loans require imputations in 1994 and beyond.

As detailed in Appendix A, the imputations for all variables are done under the general framework of multiple imputation (Rubin, 1976).²¹ Within this framework, we employ predictive mean matching (PMM) as developed by Little (1988). The PMM procedure uses regression-based predictions to find the nearest “donors” of a given variable for loans with missing values. We randomly select one of the eight closest donors to provide the missing value. PMM is well suited for imputing variables that are not normally distributed, such as credit scores and DTIs, and it avoids the convergence problems that can affect other methods, such as logit models, for predicting categorical variables.

The PMM models are estimated separately for each market segment. To allow the estimated parameters to vary over time, we use the shortest estimation window that we judge to have sufficient data for reliable estimation. For Enterprise loans, we estimate separate PMM regressions every month starting in 1995, with some pooling of months for earlier years; for

²¹ Multiple imputation preserves both the first and second moments of the underlying data, which is critical to our exercise; the relation between risk factors and default is often highly convex, making preservation of the second moment in the imputed data of paramount concern.

other market segments, we use a mix of windows ranging from monthly to grouped years. In addition, all the models use a rich set of loan characteristics as explanatory variables.²²

Because the coverage of portfolio and FHA/VA loans in the servicer datasets is incomplete, we weight the included loans to be representative of the full national market. The first step is to construct national loan totals for both market segments using HMDA data, which we adjust to account for loans not subject to HMDA reporting.²³ These adjustments provide estimated annual loan counts for the entire conventional mortgage market and the FHA/VA market. We back out the total annual count of portfolio loans as the conventional market count minus the Enterprise loan count (from the FHFA dataset) and the PLS loan count (from the PLS dataset). With HMDA-based control totals for portfolio and FHA/VA loans in hand, we then divide the portfolio loans and the FHA/VA loans into a set of buckets for origination year, loan amount, and state. The final step is to construct separate bucket-level weights for portfolio loans and FHA/VA loans equal to the adjusted HMDA count in each bucket divided by the count from our servicer-based dataset.

Despite this weighting, the servicer data still may not be fully representative of the national market if the loans in a given origination year, loan amount, and state bucket have risk characteristics that differ from the universe of loans in that bucket. This issue can only be addressed by benchmarking to a representative external database. The National Mortgage Database (NMDB), an ongoing project jointly managed by FHFA and the Consumer Financial Protection Bureau, has the potential to provide such an external benchmark. The NMDB, still under development, is drawn from a 5 percent random sample of mortgage files at a national

²² We exclude some risk factors, such as prepayment penalties and the incidence of fraud, that are frequently missing from our data or not measured at all. Insofar as these are correlated with other explanatory variables, the predictive power from these omitted variables will be subsumed in the coefficients for the observed factors.

²³ The details of these adjustments and other aspects of our weighting procedure can be found in Appendix A.

credit repository dating back to 1999 (see Avery et al., 2017, for details). As described in Appendix A, we make some use of the NMDB and other information to benchmark credit scores and DTIs calculated from the servicer data.²⁴

3. Stressed default rates

We calculate a stressed default rate for every loan in the dataset based on the observed default experience of similar loans originated nationwide in 2006 and 2007. We define a loan to have defaulted if it was ever 180 days delinquent or was terminated with the borrower forfeiting the property. Stressed default rates are calculated by measuring lifetime loan performance for fine disaggregations of the 2006-2007 cohort and then applying the cell-level default rates to all loans originated between 1994 and 2018. The cells are highly granular and span a number of loan-level risk factors. This approach can be thought of as equivalent to estimating a standard logit default model with nearly unlimited interaction terms among the modelled risk factors. It is less flexible than standard machine-learning algorithms such as CART models, which let the data determine the structure of the cells into which the loans are allocated. Nonetheless, as we show in Appendix B, the results from our approach are very similar to those from machine-learning models.

It is important to be clear about the interpretation of our default measure. The stressed default rate for a given loan represents its expected performance had it been hit shortly after origination with a replay of the financial crisis (including the observed policy response) *and* experienced the national average decline in house prices. The second part of the interpretation holds because the observed performance of 2006-2007 originations nationwide is used to

²⁴ We also use the NMDB to benchmark credit scores for PLS loans starting in 2008. After the housing bust, PLS pools included securitizations of defaulted PLS loans that were originated in earlier years. The CoreLogic PLS data likely contains such loans, which could impart some error in measuring the characteristics of newly originated PLS loans. The NMDB provides a benchmark based on new originations.

estimate the stressed default rates. Given that the drop in house prices during and after the crisis varied enormously across localities, the stressed default rate calculated in this way will only be valid for a national portfolio of loans. In section 5, we introduce an extension of the stressed default rate that incorporates location-specific house price shocks and thus is useful for states and metropolitan areas.

The stressed default rate is the analogue in the mortgage market to crash tests for motor vehicles or wind ratings for doors and windows in hurricane zones. In all of these cases, the goal is to assess performance under severe stress. This measure of mortgage risk should be of primary interest to policymakers because the stability of the mortgage market depends on its ability to withstand extreme events.

To capture the wide variation in default rates, we calculate eight separate default tables for loans originated in 2006-2007. Four tables apply to different types of fixed-rate mortgages, with a parallel set of tables for ARMs. As shown in Table 2, the four default tables cover each combination of documentation status (full doc or low/no doc) and amortization of principal (full amortization or less than full). The loans that we classify as not fully amortizing have an interest-only period, a period of negative amortization, or a balloon payment at the end of the loan. Each of the eight tables contains 320 cells to account for all combinations of eight credit score buckets, eight CLTV buckets, and five DTI buckets.²⁵ Given the eight tables and 320 cells per table, we allocate the 2006-07 home purchase originations across 2,560 cells in total.²⁶

²⁵ The credit score buckets are 579 or less, 580-619, 620-639, 640-659, 660-689, 690-719, 720-769, and 770+; the CLTV buckets (in percent) are 60 or less, 61-70, 71-75, 76-80, 81-85, 86-90, 91-95, and 96+; and the DTI buckets (in percent) are 33 or less, 34-38, 39-43, 44-50, and 51+.

²⁶ For the cells with 100 or more loans, we use the actual default rate for the loans in that cell. For cells with less than 100 loans, we use the average estimated default rate from a logit model built using all loans in the default tables; details of the logit model are available on request. Cells with no loans use the average estimated default rate from the logit model for a loan with the median value for each risk factor that defines the bucket. While a cutoff of only 100 loans leaves room for the actual default rate in the cell to be affected by idiosyncratic factors, our focus is

We produce separate default tables for Enterprise, PLS, FHA, and VA loans. For portfolio loans, we are unable to produce reliable tables because we have incomplete information on loan performance. Accordingly, we use the default tables for Enterprise loans to assess the riskiness of portfolio loans. The maintained assumption is that portfolio loans originated in 2006-07 with the same risk characteristics would have the same default experience.²⁷

As an illustration of the default tables, Table 3 displays two slices from the default table for Enterprise 30-year fixed-rate home purchase loans with full documentation and full amortization. The top panel shows the slice of the table for credit scores of 720-769 and the bottom panel shows the slice for credit scores of 660-689. To help visualize the pattern of default rates in the table, the cells with default rates of 8 percent or less have green shading, those with default rates of 8.01 to 16 percent have orange shading, and those with higher default rates have red shading.²⁸

Both panels show that higher DTIs and CLTVs increase stressed default rates, though the CLTV effect is considerably stronger. Moving from the lowest CLTV bucket to the highest in either panel increases the stressed default rate by a factor between 6 to 12, while moving from the lowest DTI bucket to the highest raises stressed defaults by a factor of only 1.5 to 2.5. The effect of credit scores can be seen by comparing the same cells across the two panels. This

on aggregate metrics, not cell-specific results. We are therefore able to rely on the central limit theorem to wash out these idiosyncratic errors. A final point that applies to the entire default table is that cells with few observations have little effect on the aggregate stressed default rate due to the extremely low share of loans in those cells.

²⁷ This assumption may be testable with the NMDDB data, which tracks loan performance. Exploring the feasibility of such a test is on our agenda for future work. For FHA loans, we confirmed that the FHA default tables produce essentially the same stressed default rates as running the FHA loans through the Enterprise default tables, which indicates that the FHA and Enterprise tables are similar for the high CLTV loans that dominate the FHA book. In contrast, the default tables for VA loans generate lower stressed default rates for VA loans than does running these loans through either the FHA or Enterprise tables. This finding is consistent with results in Goodman, Seidman, and Zhu (2014).

²⁸ The default tables do not control for the influence of house prices, as they pool loans originated across the U.S. In Section 5, however, we amend the baseline default tables to incorporate loan-level effects of house price movements. Those tables display the same general effects of CLTVs, credit scores, and DTIs on stressed defaults as in Table 3.

comparison indicates that lower scores are associated with substantially higher default rates. Although Table 3 does not show the lowest and highest credit score buckets, the difference in default rates across those buckets is comparable to that for CLTVs. Thus, the default tables embed standard results from empirical studies of mortgage default – that credit scores and CLTVs at origination are highly predictive of loan performance, with DTIs contributing some additional information (see, for example, Mahoney and Zorn, 1997; Haughwout, Peach, and Tracy, 2008; Dunskey, Kelly, and Lam, 2013; and Fout et al., 2018).

A few studies have asserted that DTIs not only have less predictive power than credit scores and CLTVs, but add little to default models (Avery et al., 1996; Foote et al., 2010). This conclusion is not supported by our results. The early research cited by Avery et al. (1996) was conducted in an era when underwriting practices generally kept DTIs within strict limits (Quercia and Stegman, 1992; Herzog and Earley, 1970). The limited variation may explain why these studies found that DTIs had so little explanatory power. More recently, Foote et al. (2010) studied the influence of DTIs on defaults for mortgages originated during 2005-2008, finding much smaller effects than those shown in Table 3. Their analysis, however, only tracked loan performance through the end of 2008, thus omitting about half of the total jump in the unemployment rate during the financial crisis. Thus, their results miss a good part of the income shock during the crisis, which mutes the estimated effect on households with heavy payment burdens. In contrast, our analysis – and that in DeFusco, Johnson, and Mondragon (2017) – tracks loan performance over a much longer period, which allows the full effect of payment burdens to emerge. A key takeaway from our research is that rising DTIs were a significant factor behind the wave of mortgage defaults during the housing bust.

All of the default tables pertain to 30-year mortgages used to purchase owner-occupied properties. We then apply adjustment factors to these tables for loans with 15- or 20-year terms, loans with 40-year terms, loans originated to investors, and loans used to purchase second homes. To calculate these adjustment factors, we estimate separate logit default models for Enterprise, PLS, and FHA/VA loans originated in 2006-2007. Each logit includes dummy variables for the four characteristics listed above as well as numerous other variables. Then, for each of the four characteristics, we run loans with that characteristic (say, investor loans) through the model as reported and a second time with the characteristic set to the baseline value (primary owner-occupied in this case). The average default rate with the characteristic as reported divided by the average default rate with the characteristic “turned off” represents the adjustment factor. The adjustment factors reduce the stressed default rates for 15- and 20-year loans and increase the rates for 40-year loans, investor loans, and second-home loans.²⁹

4. Main results

This section reports the main results from our analysis of the data on home purchase loans described in Section 2.

Market shares

Figure 2 displays the annual shares of home purchase loans by market segment over 1990-2018. The Enterprises generally accounted for about 40 to 50 percent of the total number of purchase loans over this period, with a spike above 50 percent in 2007. FHA/VA loans

²⁹ We apply the adjustment factors from the Enterprise logit results to portfolio loans, consistent with our use of the Enterprise-based stressed default tables for these loans. The actual adjustment factors used in the analysis include a refinement beyond this basic description. After calculating the logit-based estimate of each loan’s default rate, we divide the loans with the characteristic in question into two groups – those with estimated default rates above the median and those below the median. We then calculate the adjustment factor for each group separately. This refinement allows the adjustment factors to have a different effect on defaults for relatively low risk and relatively high risk loans.

represented roughly another 20 percent of the market in the 1990s, then shrank during the boom years in the 2000s, as they gave up market share to PLS loans, before rebounding in importance after the financial crisis. Altogether, government guaranteed loans have accounted for roughly 80 percent of the purchase loan market since 2009, a historically high share. The increased share has coincided with a significant reduction in loans held in portfolio and the near-disappearance of the PLS market.

Stressed default rate

Figure 3 shows the results for the stressed default rate – our summary measure of risk – from 1994 to 2018. The top left panel plots the series for the mortgage market as a whole, while the other panels disaggregate the full market into major segments. All panels use the same scale so that differences in levels across segments are readily apparent. Also, to facilitate comparisons of the rate at which risk changed across the market segments, we use a ratio scale in each panel.

The top left panel shows that mortgage risk for the market held in a narrow band from 1994 to 1998. It then fluctuated within a somewhat higher range between 1999 and 2003, followed by a steep ascent over 2004-2006. From 2007 through 2013, mortgage risk fell sharply as credit standards tightened in the wake of the housing bust. In recent years, the risk measure has edged higher, but as of 2018 it remained near the bottom of the range observed since 1994.

The low level of the stressed default rate in recent years owes importantly to the relatively high credit scores for most borrowers. If the average credit score in 2018 (about 730) were instead at the average level before the crisis (about 695 for the period 1994-2007), we estimate that the stressed default rate in 2018 would have been similar to its average over 2000-2003 – that is, during the early stage of the housing boom. We take a closer look at DTIs, credit scores, and other risk factors after completing this discussion of stressed default rates.

The remaining panels in Figure 3 display the stressed default series for loans guaranteed by the Enterprises and FHA/VA and for two splits of unguaranteed private-sector loans. The Enterprise stressed default series has the same broad contour before the crisis as the total market series, though it increased more noticeably from 1998 to 2003 than did the market-wide aggregate. The stressed default rate for FHA/VA loans changed little through 2007, which contrasts with the explosion of risk in the rest of the market. In terms of the levels for the stressed default rate, Enterprise loans consistently have had lower risk than the market as a whole, while FHA/VA loans have been more risky except in the peak years of the housing boom, when risk surged for PLS and portfolio loans.

The lower left panel shows that risk in the PLS market skyrocketed from 1998 through 2006, while the sharp rise for portfolio loans began later, in 2004. With the collapse of the PLS market in 2007, risk dropped sharply for the modest volume of new originations, and in recent years, both PLS and portfolio loans have had risk levels well below those for any pre-crisis year back to 1994.³⁰

When we perform a different decomposition of private-sector loans, interesting results emerge. This alternative decomposition separates private-sector loans that fit inside the Enterprises' conforming loan limits from jumbo loans that are larger than the limit amount and hence are ineligible for purchase by the Enterprises. Each group includes a mix of portfolio and PLS loans. Because jumbo loans have the largest loan amounts in the mortgage market, they tend to be taken out by affluent borrowers. From 1994 through 2000, the stressed default rates for jumbo and private conforming loans diverged, as the jumbo rate declined on net, while the

³⁰ The PLS series ends in 2017 because, as noted above, we have not yet incorporated the 2018 data. For 2018, results for the total market include Enterprise, FHA/VA, and portfolio loans; the exclusion of PLS loans is inconsequential because of their negligible market share.

rate for private conforming loans moved up. Importantly, this implies that risk was already on the rise by 2000 for both Enterprise loans and the private conforming loans with which they compete.

This result is relevant for evaluating the common view that the early 2000s represent a normal period in the mortgage market, a benchmark by which to assess whether lending standards in other periods are loose or tight (CoreLogic, 2017; Goldman Sachs, 2014; Urban Institute, 2018). The shading in Figure 3 covers the period 2000-2003 to facilitate the comparison to earlier years. For the market as a whole, the average stressed default rate over 2000-2003 was about 2 percentage points above the average over 1994-1999. Determining what constitutes “normal” lending conditions is inherently subjective. Some observers will view the late 1990s as normal, others will focus on the early 1990s, and others still will prefer an earlier period. The new evidence presented here calls into question the view that the early 2000s represented a period of normal lending conditions.

Risk factors

Figures 4-8 present the time series for key risk factors using the same four-panel format as in Figure 3.

CLTVs

The upper left panel of Figure 4 shows that the average CLTV for purchase loans rose in two phases: a large net increase in the first half of the 1990s and a second leg up from 2004 to 2007. The average CLTV changed little on balance from the mid-1990s through 2004 and again over the decade after the financial crisis. Its level in 2018, about 86 percent, remains close to peak for the series, indicating that high borrower leverage in the form of small down payments continues to be the norm. The longest similar measure in the literature, compiled by Ferreira and

Gyourko (2015), starts in 1997 and thus does not document the large run-up in the average CLTV in the first half of the 1990s, highlighting the value of our longer history.

A possible concern is that the CLTVs in our dataset may not fully capture the second liens taken out in connection with home purchases (so-called “piggyback” loans). We examined this question by comparing the prevalence of such loans in our dataset to the HMDA-based estimates in Bhutta and Keys (2018) and LaCour-Little, Calhoun, and Yu (2011), who show that the use of piggybacks jumped between 2003 and 2006 and then plummeted over the next three years. Although our dataset does not explicitly identify piggyback loans, we deem a piggyback to exist whenever the CLTV at origination exceeds the LTV by at least 5 percentage points.³¹ For 2003 and later years, the occurrence of piggybacks in our data closely tracks the HMDA-based estimates. Thus, the average CLTV shown in the upper left panel of the figure appears to be accurate from 2003 forward. Before 2003, however, the reporting of piggybacks in our dataset is incomplete. Our rough estimate is that the true CLTV for 2002 and earlier years could be understated by $\frac{1}{2}$ to $\frac{3}{4}$ percentage point. Correcting for this understatement would not materially change the contour shown in the upper left panel.³²

Notably, the sharp rise in CLTVs for Enterprise loans that started in 1991 was not evident in the private mortgage market until a few years later. The average CLTV for portfolio loans dropped from 1990 to 1992 and did not exceed the 1990 level until 1994. For PLS loans, the

³¹ This assumption is based on the observation that piggybacks tend to occur at 5 percentage point intervals; histograms have major spikes at 10, 15, and 20 percentage points, with a small bump at 5 percentage points. Accordingly, we set the minimum threshold at a 5 percentage point difference between the LTV and CLTV to infer the existence of a piggyback loan.

³² In addition to piggyback loans, second liens can be taken out after the purchase date by borrowing against accumulated home equity. We do not track these subsequent second liens. Using data for one of the Enterprises, Leventis (2014) estimated that approximately 5 percent of the first liens originated in 2006-2007 had a subsequent second lien. Thus, the performance of the 2006-2007 loans used to calculate our default tables would reflect the CLTV at origination and, to a modest extent, unobserved subsequent claims against home equity.

average CLTV was little changed from 1990 to 1992 and then rose gradually over the next three years.

DTIs

As shown in the upper left panel of Figure 5, the average market-wide DTI rose sharply from 1993 to 2007.³³ The average DTI retraced about half of this increase during and shortly after the financial crisis, before increasing again from 2012 to 2018. As of 2018, the average market-wide DTI stood only slightly below its peak in 2007, indicating that high debt payment burdens were again common.

The average DTI for FHA/VA loans never came down much during the financial crisis, and in recent years it has moved up, reaching an historic high in 2018. The average DTI on Enterprise loans fell sharply from its 2007 peak through 2012 and has rebounded since then. The same is true for portfolio loans, though the post-crisis decline and the subsequent rise are muted compared to Enterprise loans. For PLS loans, the data through 2017 show that the average DTI had remained near the bottom of the range seen since the early 1990s.

The sharper rise in the average DTI since 2012 for Enterprise loans than for portfolio reflects the impact of the Qualified Mortgage (QM) rule issued by the Consumer Financial Protection Bureau in 2013.³⁴ The Dodd-Frank Act and the QM rule grant private lenders a safe harbor against borrower lawsuits for loans that have DTIs of 43 percent or less (and other risk-reducing features). As shown by DeFusco, Johnson, and Mondragon (2017) and confirmed by our data, lenders have responded by limiting their portfolio holdings of loans with DTIs above 43

³³ Although Foote, Loewenstein, and Willen (2018) do not directly measure DTIs, they present evidence consistent with this increase in the average DTI. Specifically, using loan-level data, they regress the purchase loan amount at origination on borrower income and a set of control variables. They find that the relationship between the loan amount and income became weaker over the 1990s and into the 2000s, indicating that income constraints on loan size had loosened.

³⁴ For a summary of the QM rule, see https://files.consumerfinance.gov/f/201411_cfpb_atr-qm_small-entity-compliance-guide.pdf.

percent. However, Dodd-Frank permanently exempted the FHA and VA (along with other smaller guarantee programs) from the 43 percent limit, and the QM rule granted the Enterprises an exemption until January 2021.³⁵ Accordingly, the risk associated with high DTI loans has migrated to government guarantee programs.

Credit scores

Figure 6 presents the average credit score for the total purchase loan market and the same market segments as in previous figures. As shown in the top left panel, from 1994 to 2000 the average credit score for the market as a whole remained in a narrow band. From 2000 until 2003, the average score increased from a bit below 690 to about 700 and remained at that level through 2006. It then posted a large and sustained rise through 2012 and has edged down in recent years. Although there are some differences among the market segments, the basic storyline – especially the jump in the average score with the financial crisis – is similar across the various panels in Figure 6.³⁶

There are two important takeaways from this figure. First, the substantial increase in risk for home purchase loans before the financial crisis was not due to a shift toward lower-score borrowers. We provide more detail below on the low-credit-score part of the purchase loan

³⁵ Technically, the rule allows lenders to obtain QM status for a loan that is eligible for purchase by the Enterprises even if the lender retains the loan on its balance sheet. Our data show, however, that lenders have restricted their holdings of both conforming and jumbo loans with DTIs above 43 percent. On July 25, 2019, the CFPB announced that it plans to allow the Enterprise exemption to expire in January 2021 or after a short extension to ensure an orderly market transition. Until that announcement, it had been unclear whether the exemption would be renewed or allowed to expire. For details, see https://files.consumerfinance.gov/f/documents/cfpb_anpr_qualified-mortgage-definition-truth-in-lending-act-reg-z.pdf

³⁶ Because we compare credit scores over many years, an important question is whether a given score indicates the same level of credit risk under successive FICO scoring models. At the time of a model changeover, FICO ensures that borrowers assessed to have the same risk under both the old and the new models will have no change in their FICO score. Some borrowers, though, will be assessed as more risky under the new model than the old, and their scores will fall. Others will be assessed as less risky, and their scores will rise. These “swaps” could have a small impact on the distribution of scores. Overall, however, changes in scoring models do not appear to seriously compromise the comparability of credit scores over time. We thank Joanne Gaskin, Vice President, Scores and Analytics, at FICO for this information, which was obtained through email communication dated June 27-July 1, 2019.

market, but it is already clear that higher CLTVs and higher DTIs contributed to the buildup of risk while credit scores played no parallel role. Second, credit scores remain unusually high despite the downdrift in recent years for government guaranteed loans. The higher average scores likely represent a combination of tighter credit supply and weaker demand from high-risk borrowers. Regardless of the relative importance of supply versus demand factors, high credit scores have caused stressed default rates to remain low despite the high level of borrower leverage in today's market.

Other risk factors

Our analysis tracks a number of other significant factors that affect loan performance. Figure 7 shows the share of home purchase loans with less than full documentation of the borrower's income or assets. Borrowers who are self-employed or who have highly variable income often rely on such loans, though they also have been used to fraudulently overstate income. In 1990, nearly 15 percent of all home purchase loans had less than full documentation, a share that fell to an average of 11 percent for the rest of the 1990s as lenders cut back these offerings after experiencing worse-than-expected loan performance (Sichelman, 1990; Pacelle, 1991). During the first half of the 2000s, low/no doc loans grew in popularity, accounting for about one-third of all home purchase loans in 2005-2006. Many borrowers used low/no doc loans to overstate their true income and preserve buying power in the face of rising house prices (see, for example, Mian and Sufi, 2017b). Since the housing bust, low/no doc loans have nearly disappeared from the market, as the ability-to-repay rules in Dodd-Frank require the documentation of borrower income for virtually all closed-end home mortgage loans.³⁷

³⁷ See section 3.13 of https://files.consumerfinance.gov/f/201411_cfpb_atr-qm_small-entity-compliance-guide.pdf for the very limited exceptions to this requirement.

The story is much the same for loans with less than full amortization, which includes interest-only, negative amortization (“neg am”), and balloon loans. As shown in Figure 8, from 1990 through 2002, about 5 percent of all home purchase loans, on average, had less than full amortization. This share jumped over the next three years and exceeded 30 percent in 2005 for the purchase loan market as whole. The peak share varied widely by market segment, ranging from a bit more than 10 percent for the Enterprises to nearly 70 percent for jumbo loans. Amromin et al. (2018) and Garmaise (2018) show that the borrowers taking out loans with less than full amortization were primarily higher income households with prime credit scores buying homes in states with rapidly rising house prices. Like low/no doc loans, these loans were used to maintain buying power in an environment in which household income was rising much more slowly than house prices. This strategy worked as long as house prices continued rising but was associated with high default rates once prices came down. After the housing bust, these loans became extremely uncommon, largely because loans must fully amortize to qualify for QM status.

Figure 9 shows the prevalence of three additional risk factors for home purchase loans, without the segment detail in previous figures. All three factors contributed to the build-up of mortgage risk during the boom. The ARM share of purchase loans (upper panel) jumped from 2002 to 2005, peaking at almost 50 percent in 2005 before plummeting. The investor share (middle panel) also peaked in 2005, and the share of loans with terms of 15 or 20 years (bottom panel) trended lower from 1998 to 2005, which slowed the accumulation of borrower equity.³⁸ The shift from fixed-rate mortgages to ARMs, from owner-occupied loans to investor loans, and

³⁸ The swing in the first half of the 1990s is also of interest. The 15- and 20-year share rose from 1990 to 1992 as the sharp drop in mortgage rates reduced monthly payments, encouraging borrowers to take out loans that paid off quickly. However, the rise in the 15- and 20-year share had fully unwound by 1994 with the jump in mortgage rates in that year.

from 15- or 20-year loans to 30-year loans each propelled the stressed default rate to higher levels.

Role of risky product features during the boom

How much of the build-up of risk during the housing boom can be traced to standard forms of borrower leverage (such as rising DTIs) versus the growing use of risky products (such as interest-only loans)? We address this question by comparing the stressed default rate shown above, which incorporates the full set of risk factors, to an alternative stressed default rate that omits a specified set of risky product features.

The risk factors we classify as risky product features are guided by the QM rule. To meet the QM definition, a loan must fully amortize, must have a term of 30 years or less, and must have full documentation of income. Thus, we define risky product features as those characteristics that make a loan ineligible for QM status – less than full amortization, less than full documentation, and a term greater than 30 years. To calculate our alternative stressed default rate that excludes risky product features, we run low/no doc loans and loans with less than full amortization through the default tables as loans without those features and we remove the adjustment factor for loans with a 40-year term.

Figure 10 displays the results of this exercise. The solid line shows the standard stressed default rate that embeds all risk factors, while the dashed line shows the alternative series that excludes risky product features. From 1998 through the peak in 2006, the alternative series increased by 7.0 percentage points, which represents 45 percent of the 15.7 percentage point rise

for the standard series. This decomposition indicates that both “plain-vanilla” leverage and risky product features played an important role in the run-up of mortgage risk during the boom.³⁹

Mortgage rate spreads

Our results document a sharp rise in the riskiness of home purchase loans from the late 1990s to the mid-2000s. This finding, on its own, does not identify whether the increase in risk owed to an expansion of credit supply, greater credit demand, or both. However, the literature provides ample evidence that an easing of credit supply was at least part of the story.⁴⁰

One of the key pieces of evidence supporting an expansion of credit supply is that mortgage rate spreads came down during the boom period (Demyanyk and van Hemert, 2011; Justiniano, Primiceri, and Tambalotti, 2017; Mian and Sufi, 2017a). This decline implies that the supply of credit increased by more than enough to accommodate any outward shift in mortgage demand. Our data add to this body of evidence. In particular, we show that rate spreads for the mortgages with the highest stressed default rates compressed sharply during the boom years, which indicates that lenders and mortgage investors became significantly more willing to extend credit to high-risk borrowers. To our knowledge, this is the first exploration of spreads across risk tiers.

³⁹ Li and Goodman (2014) perform a similar decomposition and find a larger role for risky product features. This difference likely relates to their estimate of a very small rise in stressed defaults for Enterprise loans, with portfolio and PLS loans accounting for nearly the entire increase in the market-wide stressed default rate. Since risky product features were more common in portfolio and PLS loans than in Enterprise loans, these features end up dominating plain-vanilla leverage in the Li-Goodman decomposition. However, our results imply that they substantially understated the rise in risk for Enterprise loans, which calls their decomposition into question.

⁴⁰ See Mian and Sufi (2017a) for a discussion of the evidence. Anenberg et al. (2018) provide a nice summary measure of the expansion of mortgage credit supply during the boom by estimating a supply frontier, which represents the maximum loan size that could be obtained for a given set of observable borrower and loan characteristics; they show that the frontier moved substantially higher from 2001 to 2005-2006. Despite this evidence, Albanesi et al. (2017) question whether there was any increase in credit supply based on their finding that mortgage debt growth was strongly correlated with income growth over this period. This correlation, however, would be consistent with a broad-based increase in credit supply throughout the income distribution.

Figure 11 presents the results. We limit our attention to loans for which lenders and investors fully bear the credit risk – PLS and portfolio loans without any private mortgage insurance – to get a clean read on changes in supply. Each panel of the figure shows the average annual spreads from 1998 to 2007 for loans with stressed default rates in progressively higher risk buckets measured relative to the mortgage rate for loans in the lowest stressed default rate bucket (0-10 percent).

For PLS loans (left panel), the level of spreads corresponds one-for-one with the ordering of the risk buckets, confirming that the market was pricing for risk. Spreads for the highest two risk buckets fell dramatically between 1999 and 2005, before turning up in 2006-2007, when investors began to have concerns about the performance of PLS loans. Spreads for the third highest risk bucket also fell noticeably. For portfolio loans, spreads also compressed relative to the lowest risk bucket, with essentially all of the narrowing having occurred by 2001.⁴¹ Overall, these results provide additional evidence of an expansion of credit supply for riskier borrowers during the housing boom.

Borrowers with low credit scores

An ongoing debate in the literature concerns the role of “subprime” borrowers in the housing boom and bust.⁴² Mian and Sufi (2009, 2014, 2017a) have emphasized the importance of mortgage lending to marginal borrowers, who traditionally would have been denied credit. However, a number of recent papers have questioned the focus on this sector, arguing that (a) the growth of mortgage debt was spread across the credit score distribution, (b) loan defaults during

⁴¹ It is not clear why the decline was so abrupt or why the spreads for the two highest risk buckets were much lower over 2001-2007 than in the PLS market. These features of the portfolio loan spreads warrant further analysis.

⁴² As stated previously, we define low-credit-score borrowers as those with a credit score of less than 660 at origination. Such borrowers are often referred to as “subprime” in the academic literature. In other contexts, there is no agreed-upon definition of the term. Accordingly, we refrain from using the subprime label throughout the paper except in reference to others’ work.

the crisis were more skewed toward middle- and higher-income borrowers than in previous periods, and (c) house prices increased the most in areas with relatively weak lending to borrowers with low credit scores (Adelino, Schoar, and Severino, 2016 and 2017; Albanesi, Di Giorgi, and Nosal, 2017; Conklin et al., 2018; Ferreira and Gyourko, 2015).

We make two contributions to this debate. First, using our data for the entire mortgage market, we track the low-credit-score share of home purchase loans. And, second, we examine whether low-credit-score loans became markedly riskier than other loans during the boom. Our definition of low-credit-score loans as those with scores below 660 follows Adelino, Schoar, and Severino (2016, 2017), Conklin et al. (2018), and Mian and Sufi (2009).

Figure 12 displays the annual low-credit-score share of purchase loans over 1994-2018. For the total market, the shares in each year from 2003 to 2007 were below those in every prior year shown. The pattern varied somewhat across the market segments, but FHA/VA loans are the only segment with a reasonably clear upward trend in the share before the financial crisis. These results are inconsistent with an emphasis on low-score borrowers having driven the growth of home purchase lending during the boom.

Figure 13 compares the stressed default rate and selected risk characteristics for loans below the 660 credit score threshold versus loans with higher scores. As shown, the stressed default rate for low-score loans moved closely with that for higher-score loans before the crisis, indicating there was no systematic increase in the riskiness of low-score loans relative to other loans. The risk characteristics shown in the figure explain why. Although the average CLTV and average DTI increased a bit more for low-score loans over 1994-2006, the low/no doc share rose less.

All in all, these results point to a broad-based rise in mortgage risk before the financial crisis. Risky loans became more plentiful and on more favorable terms to both low- and high-credit-score borrowers.⁴³

5. Incorporating House Price Shocks

Measuring Shock CLTVs

An implicit assumption in the national stressed default rate is that each loan is subjected to the average house price decline in the U.S. during and after the Great Recession. The purpose of this assumption was to intentionally constrain the resulting risk indicator to reflect only borrower and mortgage characteristics. But for some applications, it may be desirable to add additional loan-level information. For instance, local risk indicators require the use of location-specific house price risk given the wide variation in house price movements across the country. In this section, we implement local stress measures in a direct and straightforward manner by incorporating hypothetical local house price shocks in the construction of risk indicators.⁴⁴

We define a “shock combined loan-to-value” ratio at origination (SCLTV) as the ratio of the unpaid principal balance at origination (UPB) divided by the “shock house value,” which is set equal to the house value at origination (V) multiplied by one plus a projected 3-year house price shock.

$$SCLTV = \frac{UPB}{V \times (1 + \Delta V^s)}$$

⁴³ An alternative definition of subprime loans in the literature uses 620 as the credit score threshold. We think this alternative definition is of less interest than our primary one because the below-620 share of the purchase loan market is small – only about 15 percent on average over 1994-2006. That said, if we replicate Figure 13 with 620 as the threshold credit score, the results are unfavorable to the subprime-centric view of the housing boom, as the stressed default rate rises considerably less over 1994-2006 for below-620 loans than for loans with higher scores. Clearly, the rise in risk was not concentrated among borrowers with the lowest credit scores.

⁴⁴ Other location-specific stresses could potentially be implemented, including negative income shocks (via DTIs) and possible credit score deterioration.

The SCLTV represents the expected combined loan-to-value ratio of a loan three years after the onset of severely stressed conditions. Because the assumed shock involves a drop in house prices, ΔV^S is negative, causing the SCLTV to be higher than the stated CLTV at origination.

Our price shock variable is constructed following Smith and Weiher (2012) and Smith et al. (2016), who describe a simple method for constructing a house price path associated with severe economic stress. The concept behind this method is that house prices tend to fall below a long-term trend during a period of severe stress. The trend serves as a reduced-form proxy for economic fundamentals that are more traditionally modeled, such as the house price-to-income ratio or the house price-to-rent ratio. Different locations face different degrees of variation around this trend, so both the house price level in relation to the trend and the history of negative price movements are taken into account when setting the shock value.

We begin with a real house price series constructed using a nominal series divided by the consumer price index for all urban consumers (CPI). The construction of the SCLTV then proceeds in 5 steps:

1. Fit a trend line to the series for real house prices using data for 1975 through 2018. If the trend slope is negative, set it to 0.⁴⁵
2. Calculate the “maximum differential” as the largest percent difference between the trend and the below-trend values of the series, or 5 percent, whichever is larger.
3. Calculate the “lower bound” in each year as the trend minus the maximum differential defined in the previous step. If the lower bound calculated in this way is less than 5 percent below the house price index for a given year, set the lower bound for that year to be 5 percent below the price index.⁴⁶

⁴⁵ The house price series used are the FHFA annual State and CBSA indices from Bogin, Doerner, and Larson (2019).

⁴⁶ For any house price index that requires the imposition of this 5 percent gap, it is needed, by definition, in the year of the maximum differential, but it can be required in other years too.

4. Calculate the “stress loss” as the difference between the current value of the house price index and the lower bound price level three years in the future.⁴⁷ The use of the lower bound price level three years ahead allows the price shock to play out over a realistic time horizon.
5. Convert the real stress loss to a nominal stress loss, ΔV^s using the CPI.

A visual depiction of this procedure is shown in Figure 14 for the Washington, DC CBSA. The real house price trend is estimated over the entire series as 1.4 percent per year. The maximum differential is achieved in 1999 at 15 percent below the trend, giving a lower bound of 15 percent below the trend in all periods except 1996 through 2000, 2012, and 2013, where the lower bound is set to the minimum of 5 percent below the price index. The maximum stress loss is 41 percent in 2006, which is calculated using the index value for 2006 and the lower bound value for 2009.

This procedure is repeated for each metropolitan statistical area (MSA) and each state. For each loan in the database, the SCLTV is calculated using the house price index for its MSA if available. If the MSA index is unavailable – for instance in small cities where city-specific indices are not available or in rural areas that are not part of an MSA – the maximum stress loss is constructed using the state index.

After calculating loan-level SCLTVs, we implement the procedure outlined in Section 3 to construct stressed default rates for each loan in the database based on SCLTVs substituted for CLTVs at origination. The key difference is the grid cutoffs for SCLTVs, which are now (in percent) 60 or less, 61-80, 81-100, 101-115, 116-130, 131-145, 146-160, and 160+. We then

⁴⁷ Because the real house price trend is log-linear, in practice we calculate the lower bound as the current trend value plus three times the trend slope minus the maximum differential. This value is then set to the house price index minus 5 percent if the value is smaller in magnitude. All variables are in real terms for this calculation.

revise the default tables to use the SCLTV buckets and construct national, state, and MSA stressed default rate by averaging the loan-level stressed default rates by year.

Results

The average SCLTV for the nation as a whole is shown in the top panel of Figure 15. By construction, the SCLTV is higher than the unadjusted CLTV for every loan because the value of the home is at least 5 percent lower than the value used to construct the CLTV. The extent to which the average SCLTV exceeds the CLTV depends on the location-specific house price shocks averaged over the country as a whole. The top panel of the figure shows that the average SCLTV varies much more over time than the unadjusted average CLTV. In 2006, the national average SCLTV reached 138 percent, indicating the enormous collateral risk due to above-trend house prices.

The bottom panel shows that the stressed default rate calculated using the SCLTV exhibits a greater degree of pro-cyclical behavior than when calculated using the unadjusted CLTV. This occurs because the changes in house price risk over the cycle amplify the observed variation in CLTVs at origination, as shown in the upper panel. The stressed default rates converge in the 2006-2007 period by construction, as both sets of default tables reflect the observed performance of loans originated in those years.

There are two periods, in particular, that highlight the usefulness of the stress default measure calculated using SCLTVs. The first is the 1996-2002 period, where this measure shows a significant increase in risk, while the standard stressed default measure shows a much milder rise in risk. The second is the 2013-2018 period, where similarly, the measure based on the SCLTVs shows a more pronounced rise in risk. In both cases, while risk based solely on borrower and loan characteristics is rising slowly, house price risk is steepening the overall risk

trajectory. This brings into focus one of the key insights of this extension of the standard measure. It shows that default risk can rise substantially even when there are only modest overall changes in borrower and mortgage characteristics.

The second main purpose of SCLTV-based default risk estimates for individual loans is that they can be aggregated at fine levels of geography to construct measures of local risk. We calculate stressed default rates for all 50 states, the District of Columbia, and 367 MSAs. These indicators show wide variation in stressed default risk both across areas and over time, reflecting region-specific lending and house price patterns.

Figure 16 shows local stressed default rates by quintile for states and MSAs, averaged over 2006-2007. Areas with the highest stressed default rates include the well-known “Sand States” (Arizona, Nevada, Florida, and California) as well as Maryland, Michigan, and a few less populous states. These states had the highest stressed default risk at the onset of the Great Recession, accounting for the combination of loan characteristics and the potential for severe house price declines. Areas with the lowest risk included the vertical band of states from North Dakota to Texas, along with Iowa, Arkansas, Kentucky, North Carolina, and Vermont.

The MSA panel paints a more nuanced picture of mortgage risk. In some states, including California and Nevada, the MSA patterns match that for the state as a whole, but in other states, there is wide variation. For instance, the state-level indicator for Ohio puts it in the middle quintile of default risk. However, the MSA map shows that the state-level indicator averages relatively low risk in Cincinnati and Columbus with higher risk elsewhere in the state.

Figures 17 and 18 show variation in stressed defaults for both states and MSAs over time, respectively. The states in Figure 17 were chosen to represent a diverse sampling of geographies and populations; the MSAs in Figure 18 are located within those states. These figures illustrate

three elements of the diversity of the state- and MSA-level results. First, there are periods (such as 2010-2018) during which the risk profiles are similar across states or MSAs, yet in other periods (especially around the peak of the housing boom), the risk profiles are remarkably different. Second, in some cases MSAs have higher risk than the state in which they are located (Detroit relative to Michigan), but in others the opposite is true (San Francisco relative to California in the mid-2000s). Finally, the rank-order among MSAs and states in the riskiness of originated home purchase loans varies over time. For example, in Figure 18, Detroit had the highest stressed default rate for the entire period from 1994 to 2005, but it has been in the middle of the pack since 2011.

6. Conclusion

Understanding the evolution of risk in the mortgage market before the financial crisis and after requires a long historical account built from comprehensive and accurate data. Until now, that historical record did not exist, and this paper takes a major step toward filling the gap. We bring together several sources of data, including the entire Enterprise book, to cover essentially the entire market for home purchase loans from 1990 to 2018. We track important loan characteristics and compute a summary measure of risk under stressed conditions.

We use the data to reach a number of conclusions. First, we show that loan risk had already risen by 2000, calling into question the common view that the early 2000s represented a period of normal lending conditions. This conclusion is reinforced when we account for the widening gap between house prices and their longer-term trend levels, which effectively made the CLTVs at origination higher than their stated level. Second, we provide new evidence that credit supply expanded in the early and mid-2000s by documenting a compression of mortgage rate spreads between the riskiest and least risky loans in the PLS market and in lenders’

portfolios. Third, the rise in risk before the financial crisis was very similar for borrowers with low credit scores and those with higher scores. This fact, combined with a downward drift in the low-credit-score share of purchase loans during the housing boom, undercuts explanations of the crisis that focus on low-score borrowers. Fourth, while risky product features accounted for a bit more than half of the rise in risk during the boom years, plain-vanilla factors such as higher DTIs and CLTVs accounted for a large portion as well. Finally, in today's market, overall risk is low, despite high average DTIs and CLTVs, as risky product features have largely disappeared and average credit scores remain elevated. This highlights that the post-crisis regulatory environment succeeded, at least through 2018, in limiting overall mortgage risk, though the high market share for government-guaranteed loans has increased the portion of the risk borne by taxpayers.

Table 1: Share of Home Purchase Loans with Imputed Risk Factors, 1990-2018
(percent)

Risk Factor	Enterprise	Portfolio	PLS	FHA/VA
Credit score	15	39	8	26
DTI	12	63	53	60
Documentation status	0	42	2	0
Amortization status	0	25	0	0
Occupancy status	0	17	0	0

Note: The shares pertain to the loan-level source data before any weighting occurs. Data for PLS loans are available through 2017. Amortization status refers to whether a loan fully amortizes or has less than full amortization through the use of an interest-only period, a period of negative amortization, or a balloon payment.

Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

Table 2: Default Tables

Fixed-rate Mortgages	ARMs
Full doc / Full amortization	Full doc / Full amortization
Full doc / Less than full amortization	Full doc / Less than full amortization
Low or no doc / Full amortization	Low or no doc / Full amortization
Low or no doc / Less than full amortization	Low or no doc / Less than full amortization

Note: “Less than full amortization” indicates that the loan has a feature that reduces principal payments for some period of time. These features include interest-only payments, negative amortization, or a balloon payment at the end of the stated term.

Table 3: Stressed Default Rates for Enterprise 30-year Fixed-Rate Loans with Full Documentation and Full Amortization (in percent)

Home purchase loans, credit scores of 720-769

DTI (%)	CLTV (%)							
	1-60	61-70	71-75	76-80	81-85	86-90	91-95	≥ 96
1-33	1.05	2.34	3.42	4.10	3.93	6.24	7.51	10.70
34-38	1.64	3.65	4.40	6.15	6.01	8.41	9.64	12.95
39-43	1.99	4.36	5.26	7.12	6.75	10.17	11.64	15.57
44-50	2.02	4.39	6.17	8.17	8.48	11.31	13.48	17.99
≥ 51	2.12	4.68	6.44	8.54	9.55	12.75	16.09	24.48

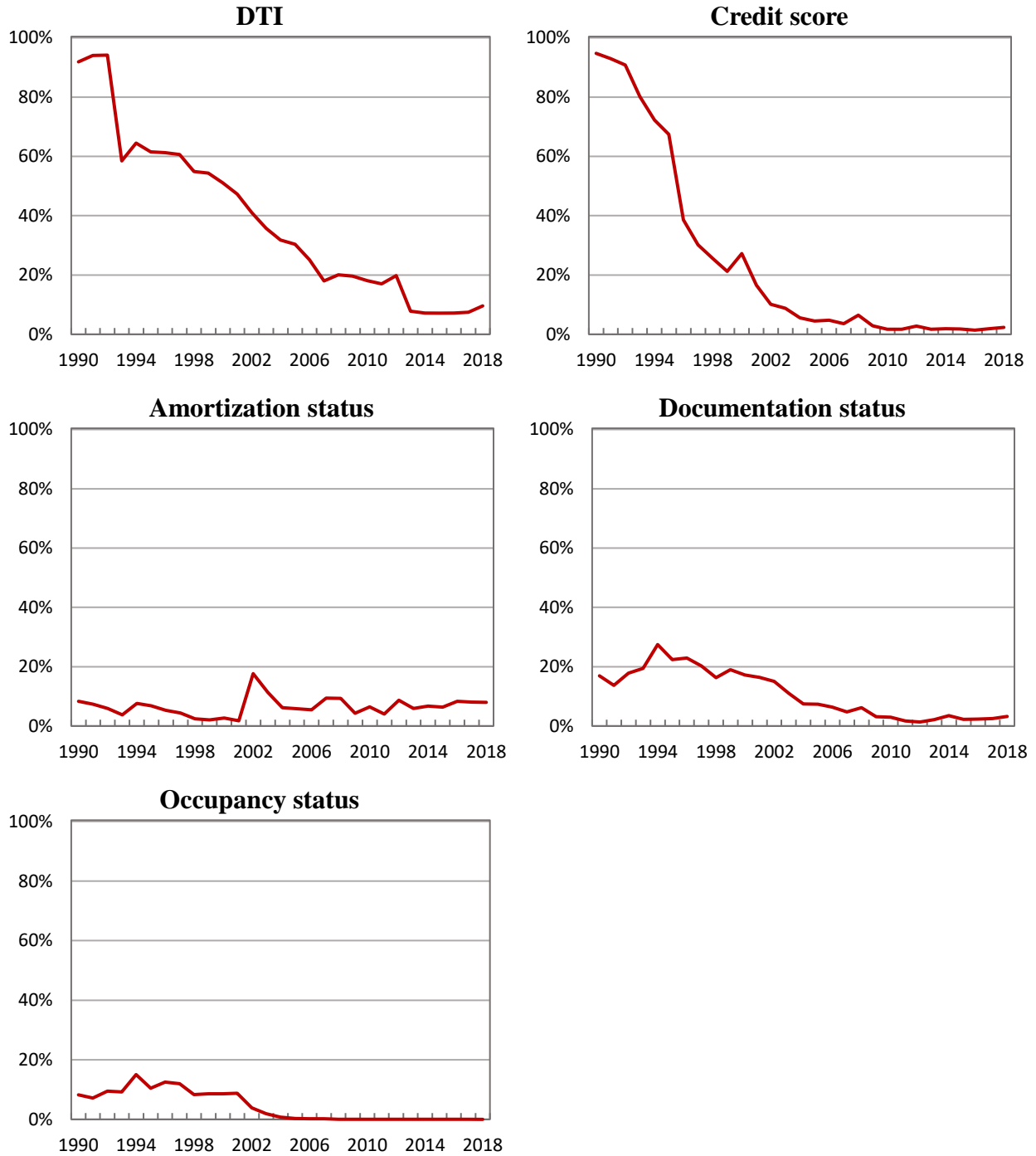
Home purchase loans, credit scores of 660-689

DTI (%)	CLTV (%)							
	1-60	61-70	71-75	76-80	81-85	86-90	91-95	≥ 96
1-33	3.30	6.04	8.82	8.62	8.17	13.06	14.59	22.90
34-38	3.85	7.19	9.05	12.57	13.30	15.80	17.83	26.80
39-43	4.73	10.19	11.54	13.39	14.86	19.15	20.85	30.20
44-50	5.46	9.90	12.77	15.28	16.84	21.30	23.68	32.72
≥ 51	4.89	11.19	14.52	16.40	21.81	23.92	29.26	40.32

Note: Shading: green for stressed default rates of 8% or less, orange for 8.01% to 16%, and red for more than 16%. See the text for details about the calculation of the stressed default rates.

Source: Authors' calculations using FHFA data for Enterprise loans.

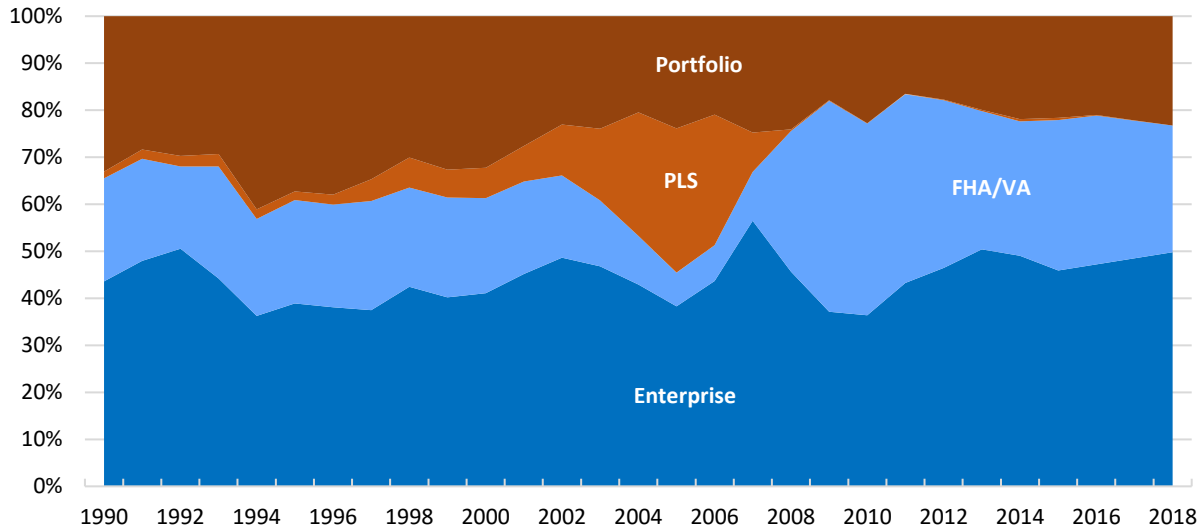
Figure 1: Imputation Share for Home Purchase Loans, Total Market, 1990-2018



Note: Results pertain to 1-4 unit purchase-money mortgages and are based on unweighted loan counts within each market segment, which are then aggregated to the total market using the annual market shares for each segment. Data for PLS loans are available through 2017.

Source: Authors' calculations using data from FHFA, Black Knight, Inc., CoreLogic, and Ginnie Mae data processed by the AEI Housing Center.

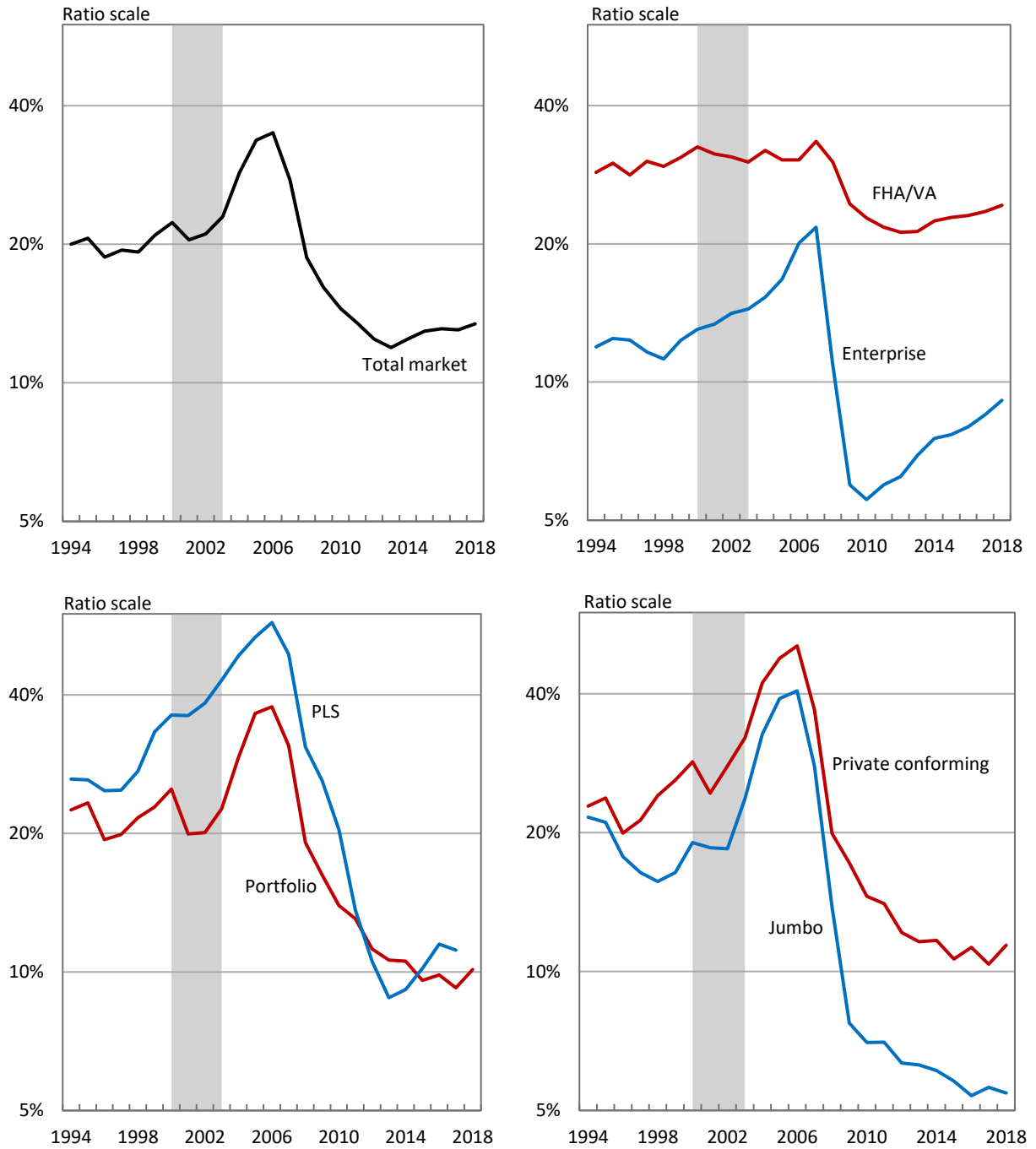
Figure 2: Market Shares of Home Purchase Loans



Note: Shares pertain to 1-4 unit purchase-money mortgages and are based on loan counts. The share for portfolio loans is calculated as a residual given shares for Enterprise loans, PLS loans, and FHA/VA loans. The counts for the total market and FHA/VA loans are based on HMDA data and are grossed-up by an estimate of the undercount in HMDA. The Enterprise count comes from FHFA data and is assumed to cover the universe of such loans. The PLS count from CoreLogic is believed to cover nearly the entire market and thus is also not grossed up. Data for PLS loans are available through 2017; for 2018, we assume the PLS share is zero. Because HMDA only began to collect information on lien status in 2004, the shares for 1990-2003 include first liens and subordinate liens. Starting in 2004, the shares are based on first liens only.

Source: Authors' calculations using data from FHFA, CoreLogic, and HMDA.

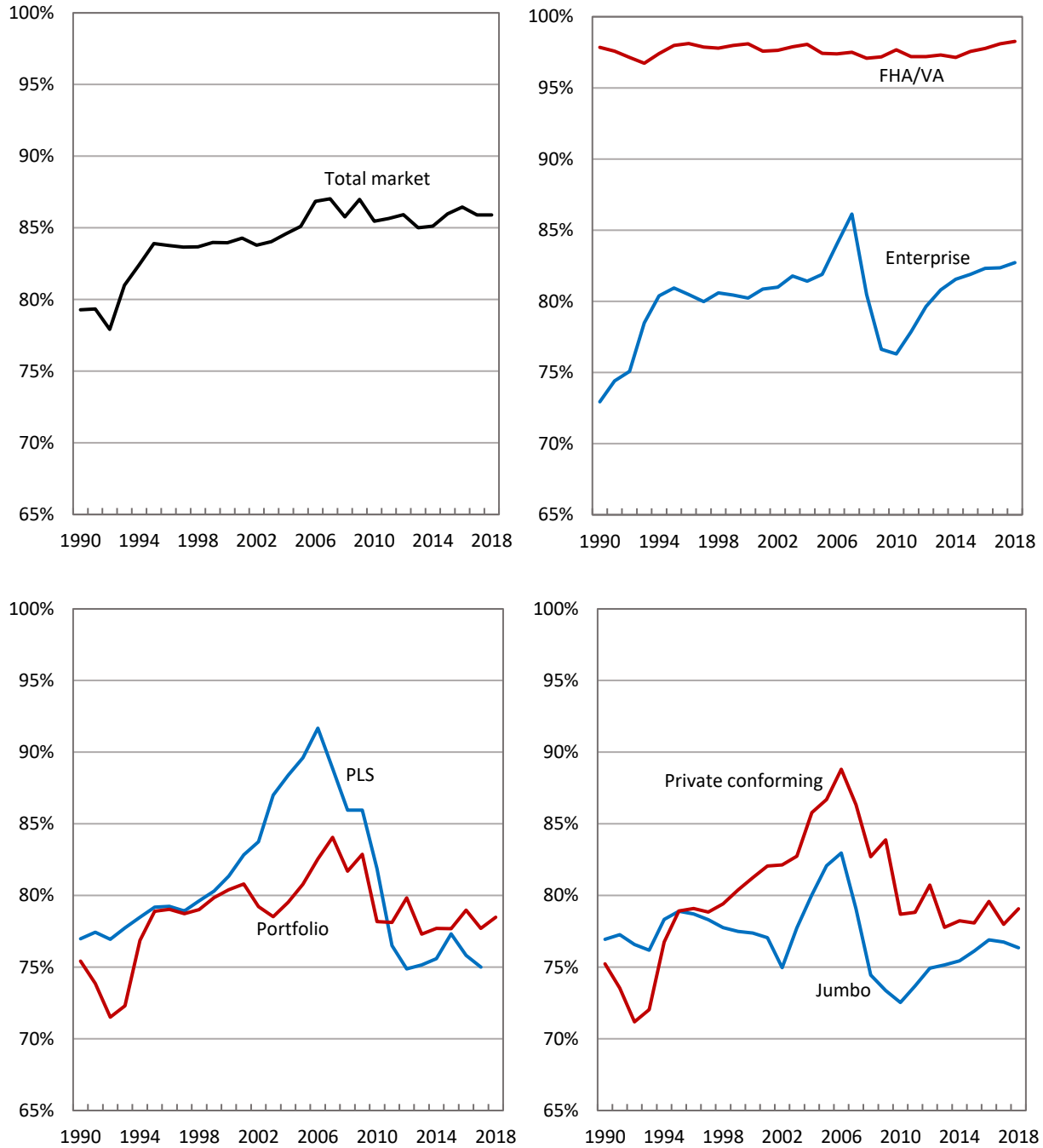
Figure 3: Stressed Default Rate for Home Purchase Loans, 1994-2018



Note: Results pertain to 1-4 unit purchase-money mortgages. Private conforming loans include portfolio and PLS loans with initial loan amounts below the applicable Enterprise conforming loan limit. Data for PLS loans are available through 2017; for 2018, PLS loans are omitted from the results for the total market, private conforming loans, and jumbo loans. Shading is for 2000-2003.

Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

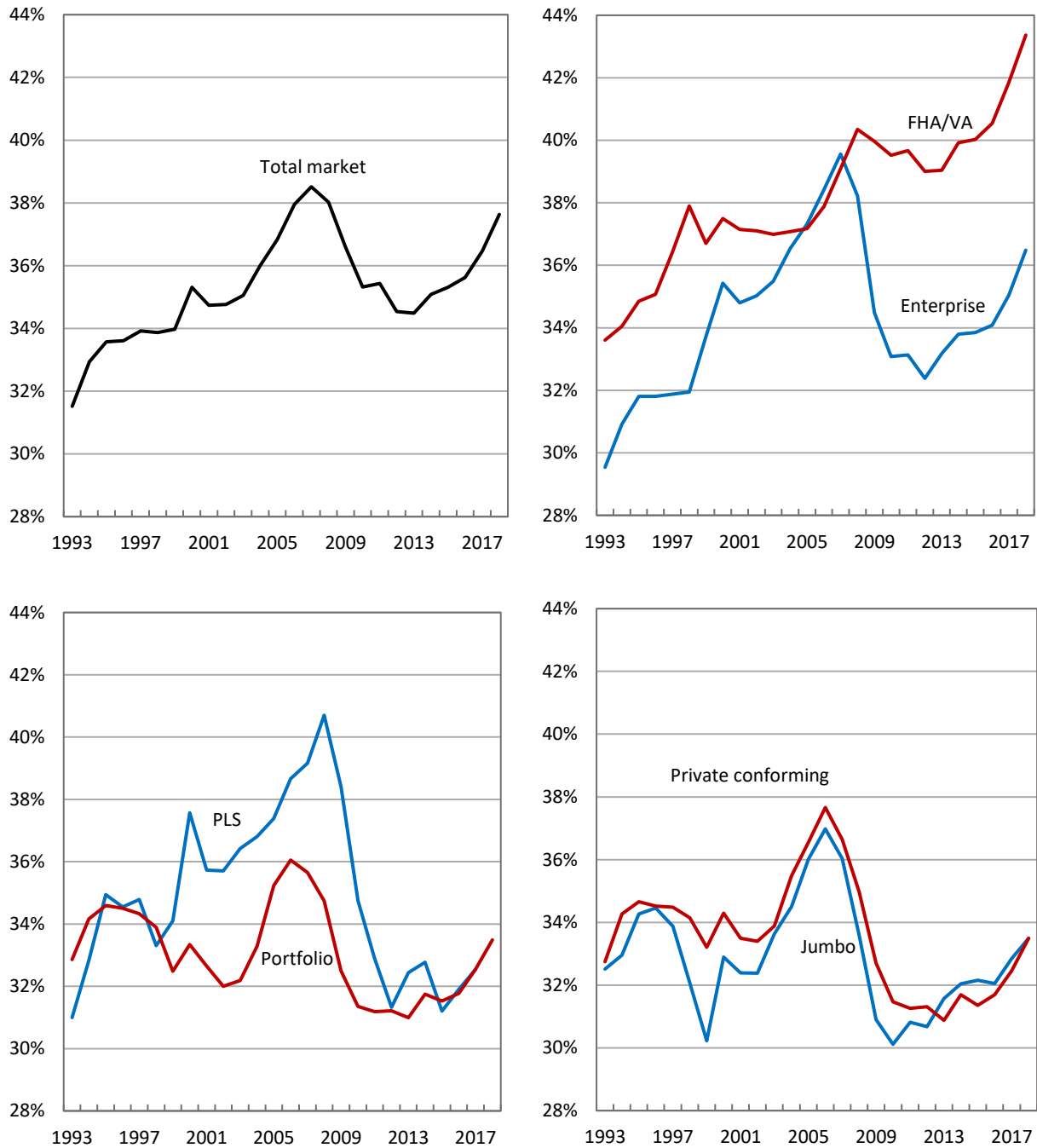
Figure 4: Average CLTV for Home Purchase Loans, 1990-2018



Note: Results pertain to 1-4 unit purchase-money mortgages. Private conforming loans include portfolio and PLS loans with initial loan amounts below the applicable Enterprise conforming loan limit. Data for PLS loans are available through 2017; for 2018, PLS loans are omitted from the results for the total market, private conforming loans, and jumbo loans.

Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

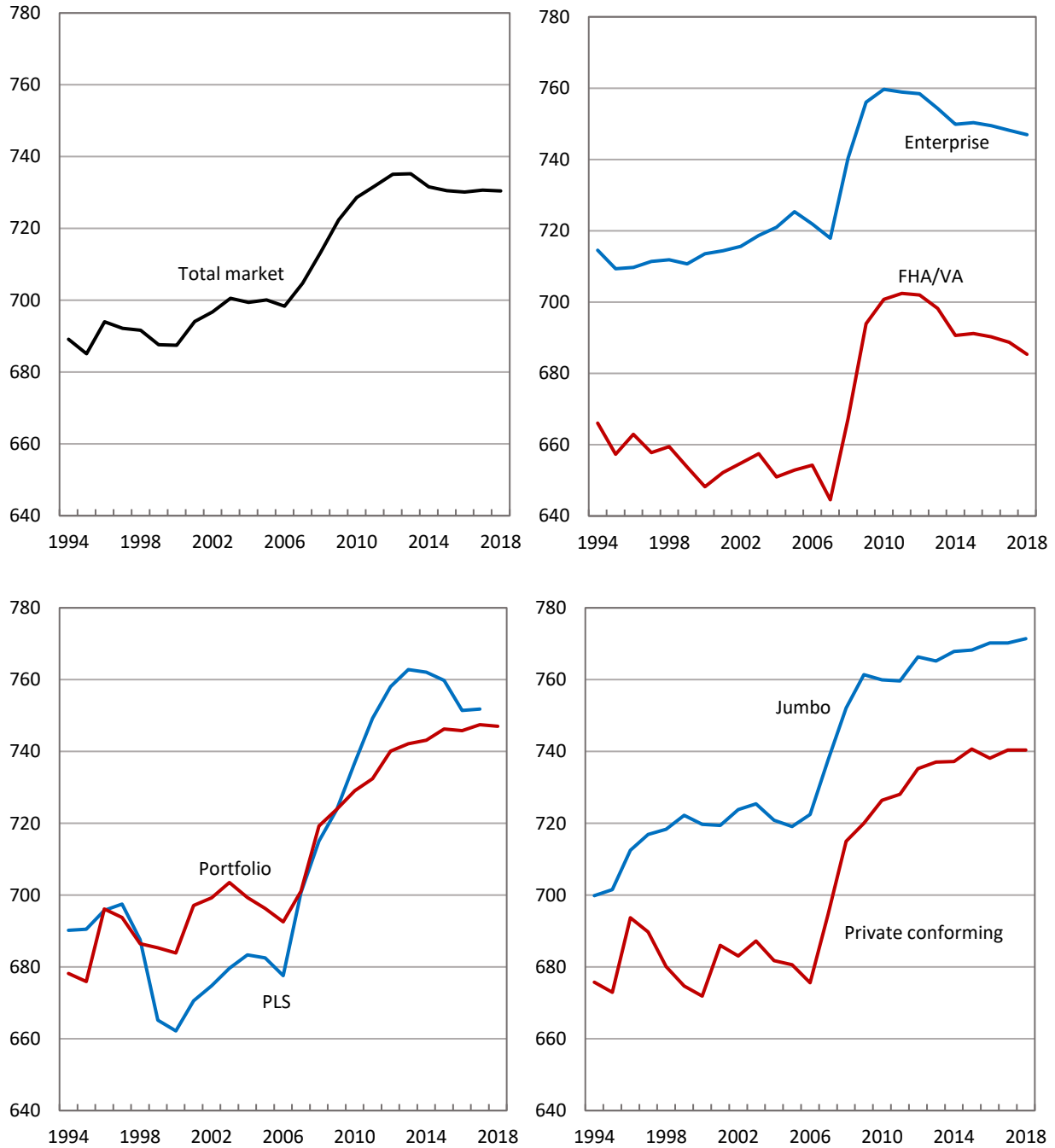
Figure 5: Average DTI for Home Purchase Loans, 1993-2018



Note: Results pertain to 1-4 unit purchase-money mortgages. Private conforming loans include portfolio and PLS loans with initial loan amounts below the applicable Enterprise conforming loan limit. Data for PLS loans are available through 2017; for 2018, PLS loans are omitted from the results for the total market, private conforming loans, and jumbo loans. Missing DTIs are imputed using regression techniques.

Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

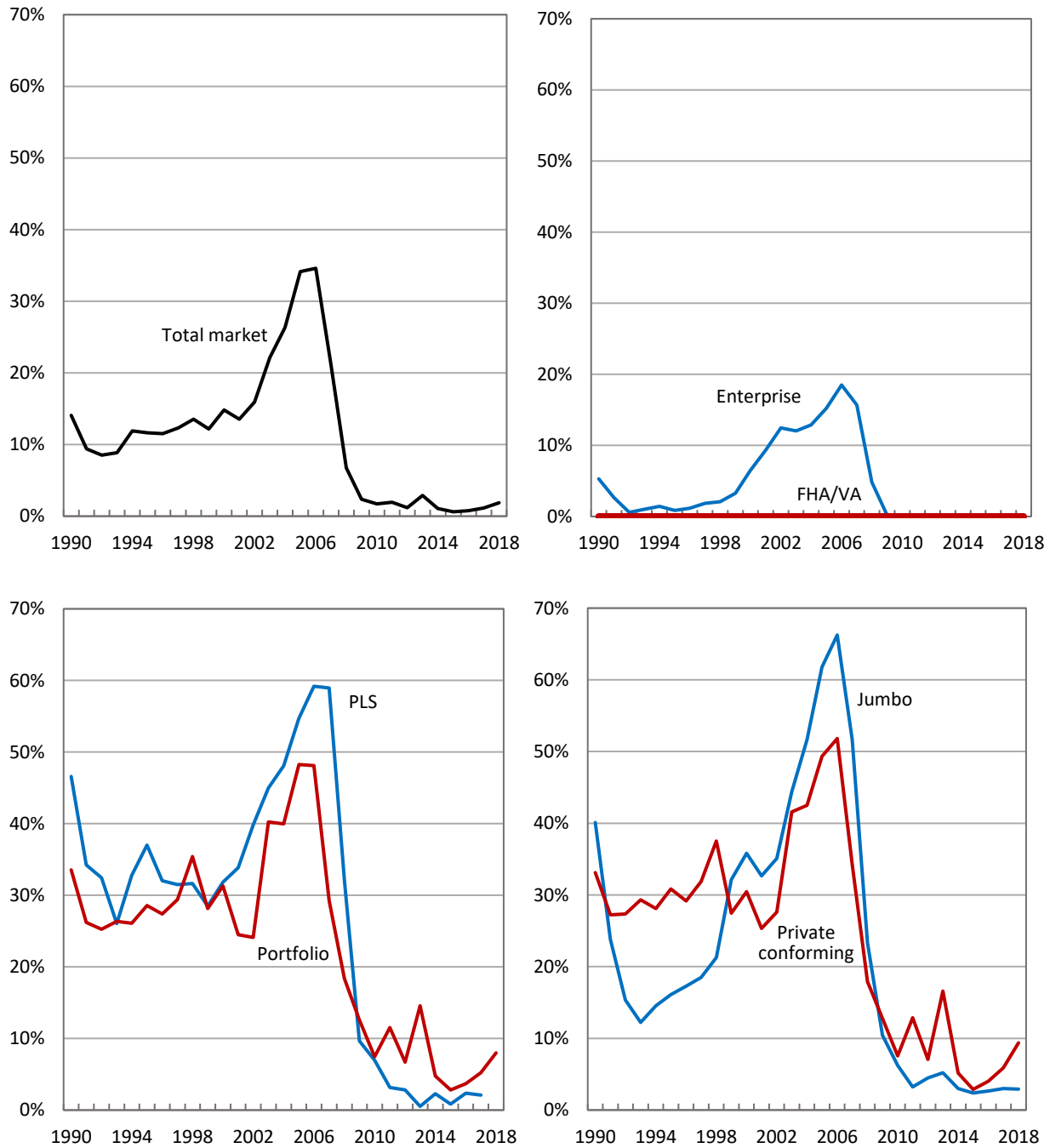
Figure 6: Average Credit Score for Home Purchase Loans, 1994-2018



Note: Results pertain to 1-4 unit purchase-money mortgages. Private conforming loans include portfolio and PLS loans with initial loan amounts below the applicable Enterprise conforming loan limit. Data for PLS loans are available through 2017; for 2018, PLS loans are omitted from the results for the total market, private conforming loans, and jumbo loans. Missing credit scores are imputed using regression techniques.

Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

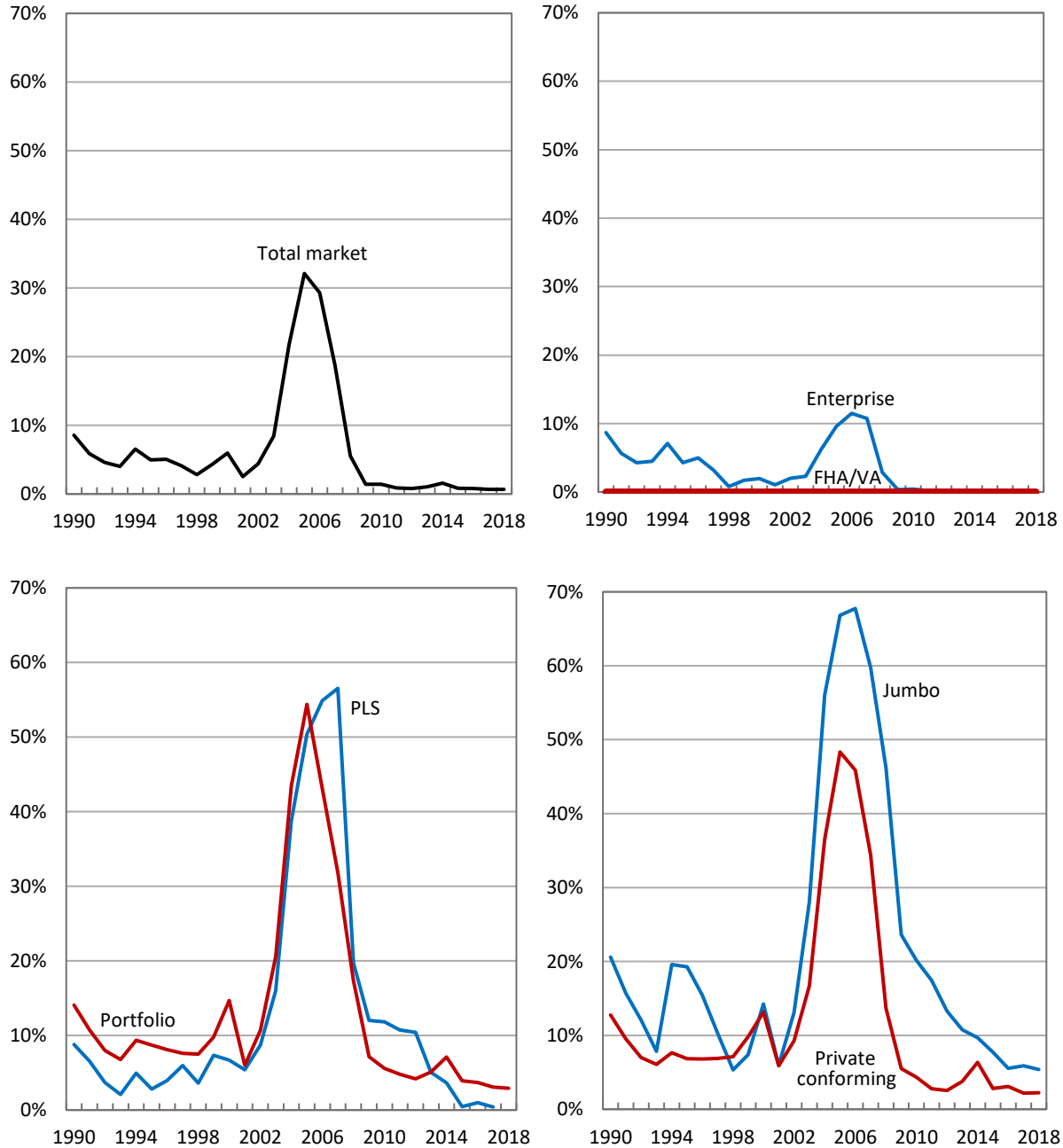
Figure 7: Share of Home Purchase Loans with Low or No Documentation, 1990-2018



Note: Results pertain to 1-4 unit purchase-money mortgages. Private conforming loans include portfolio and PLS loans with initial loan amounts below the applicable Enterprise conforming loan limit. Data for PLS loans are available through 2017; for 2018, PLS loans are omitted from the results for the total market, private conforming loans, and jumbo loans. Missing documentation status for portfolio and PLS loans is imputed using regression techniques.

Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

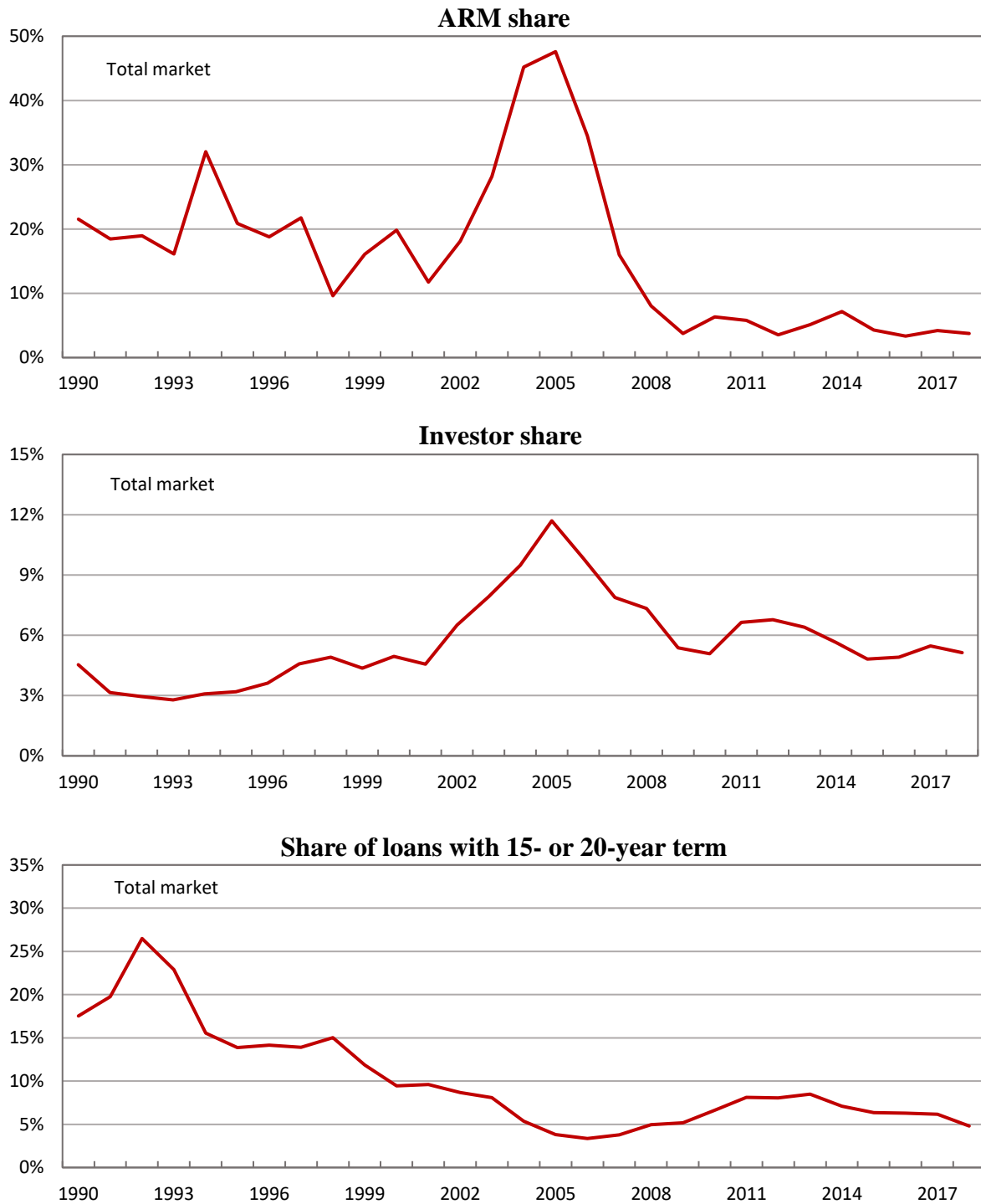
Figure 8: Share of Home Purchase Loans with Less Than Full Amortization, 1990-2018



Note: Results pertain to 1-4 unit purchase-money mortgages. Private conforming loans include portfolio and PLS loans with initial loan amounts below the applicable Enterprise conforming loan limit. Data for PLS loans are available through 2017; for 2018, PLS loans are omitted from the results for the total market, private conforming loans, and jumbo loans. A loan is classified as having less than full amortization if it has an interest-only period, negative amortization, and/or a balloon payment. Missing information on whether the loan has such features is imputed using regression techniques; this imputation is necessary only for portfolio loans.

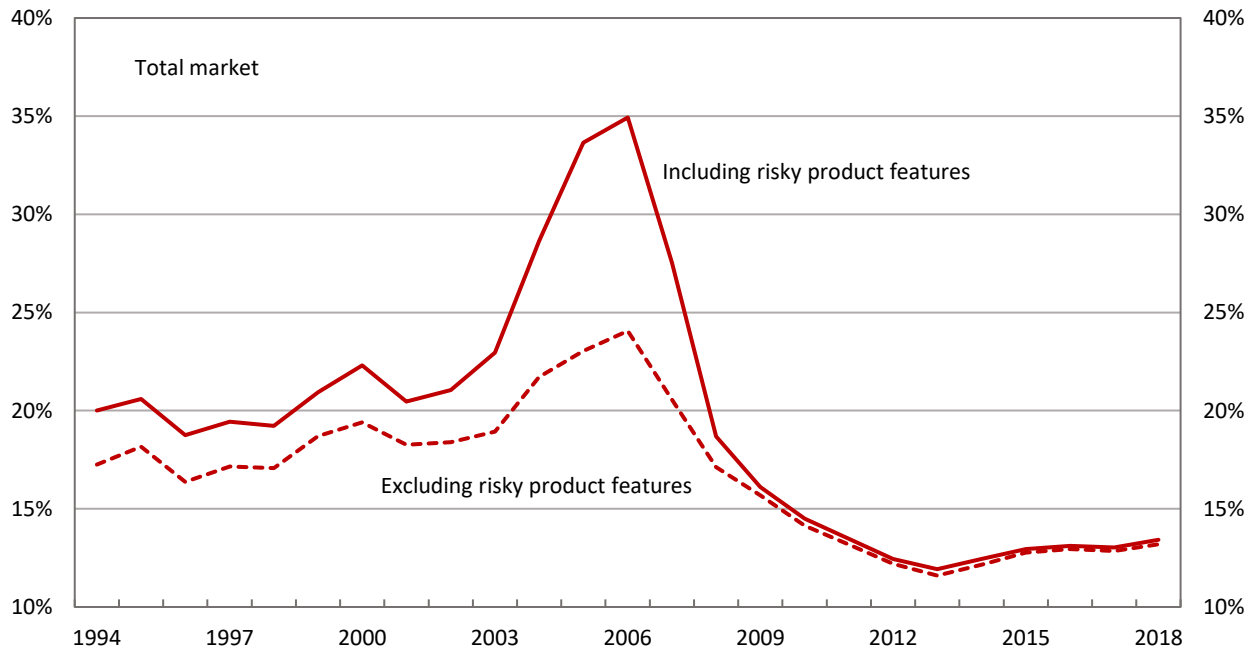
Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

Figure 9: Other Characteristics of Home Purchase Loans, 1990-2018



Note: Results pertain to 1-4 unit purchase-money mortgages. Data for PLS loans are available through 2017.
Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

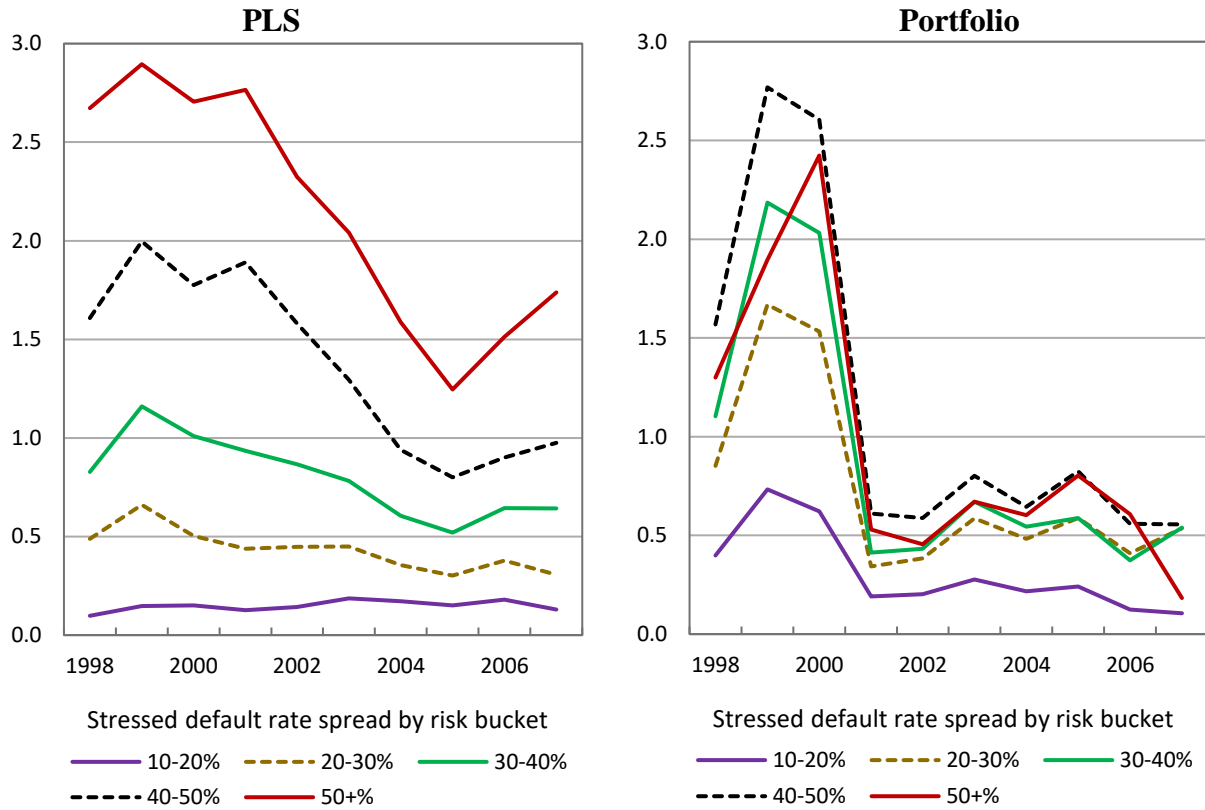
Figure 10: Influence of Risky Product Features on the Stressed Default Rate for Home Purchase Loans, Total Market, 1994-2018



Note: Results pertain to 1-4 unit purchase-money mortgages. Risky product features include low/no documentation, less than full amortization, and a loan term greater than 30 years. The stressed default rate without risky product features uses the same loans as the rate with risky product features but runs those loans with those features through the default tables as full doc, full amortization, 30-year loans. Data for PLS loans are available through 2017.

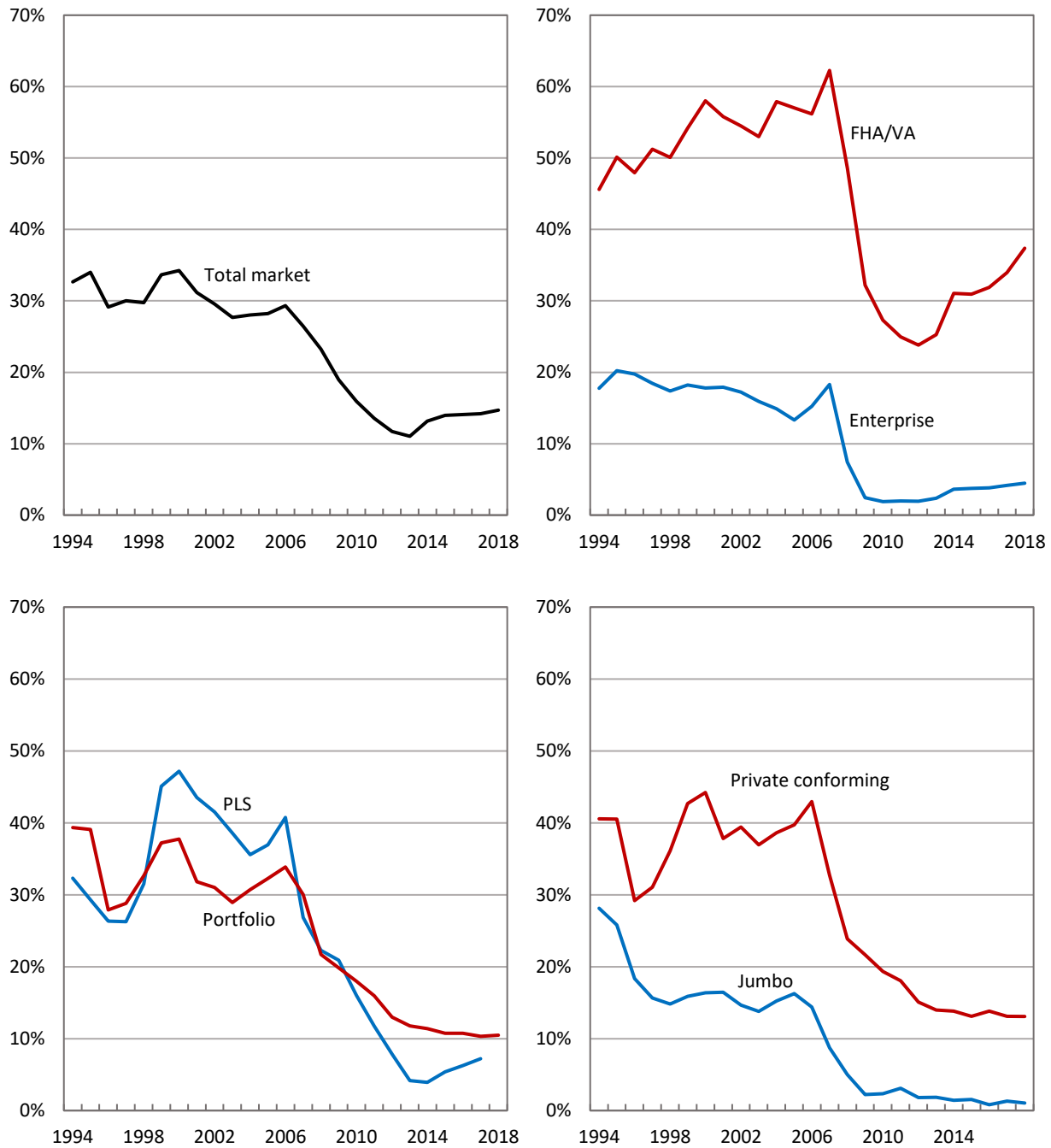
Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

Figure 11: Average Mortgage Rate Spread within a Market Segment, by Stressed Default Rate, 1998-2007
(in percentage points)



Note: The spreads in each panel pertain to 30-year, fixed-rate 1-4 unit purchase-money mortgages and are calculated relative to the average mortgage rate for the lowest-risk group in that market segment (the group with stressed default rates of 0-10 percent). We exclude loans with private mortgage insurance to obtain spreads for lenders and investors that fully bear the credit risk.
Source: Authors' calculations using data from Black Knight, Inc. (for portfolio loans) and CoreLogic (for portfolio and PLS loans).

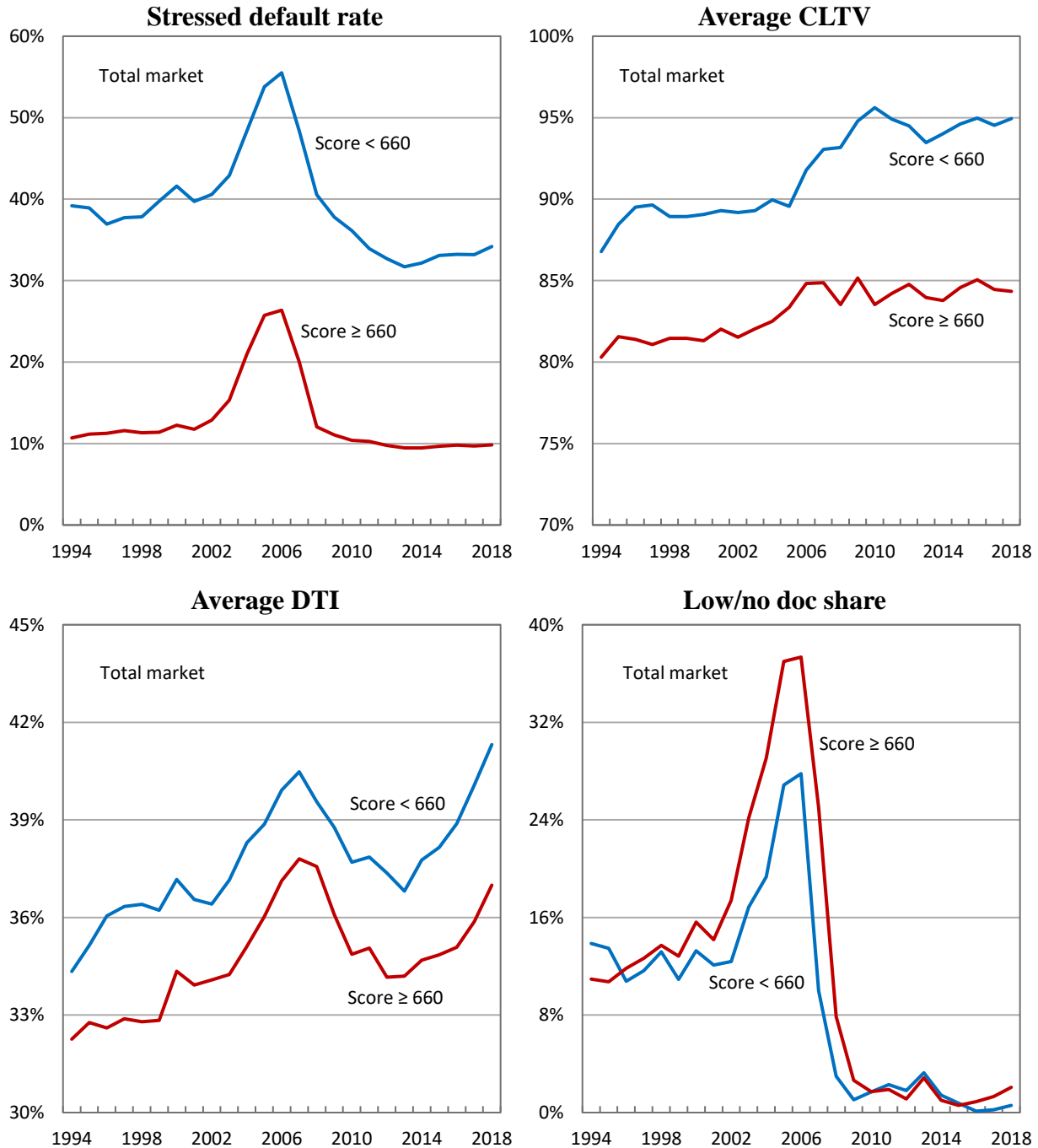
Figure 12: Low-Credit-Score Share of Home Purchase Loans, 1994-2018



Note: Results pertain to 1-4 unit purchase-money mortgages. Private conforming loans include portfolio and PLS loans with initial loan amounts below the applicable Enterprise conforming loan limit. Data for PLS loans are available through 2017; for 2018, PLS loans are omitted from the results for the total market, private conforming loans, and jumbo loans. Low-credit-score loans are defined as those with credit scores below 660. Missing credit scores are imputed using regression techniques.

Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

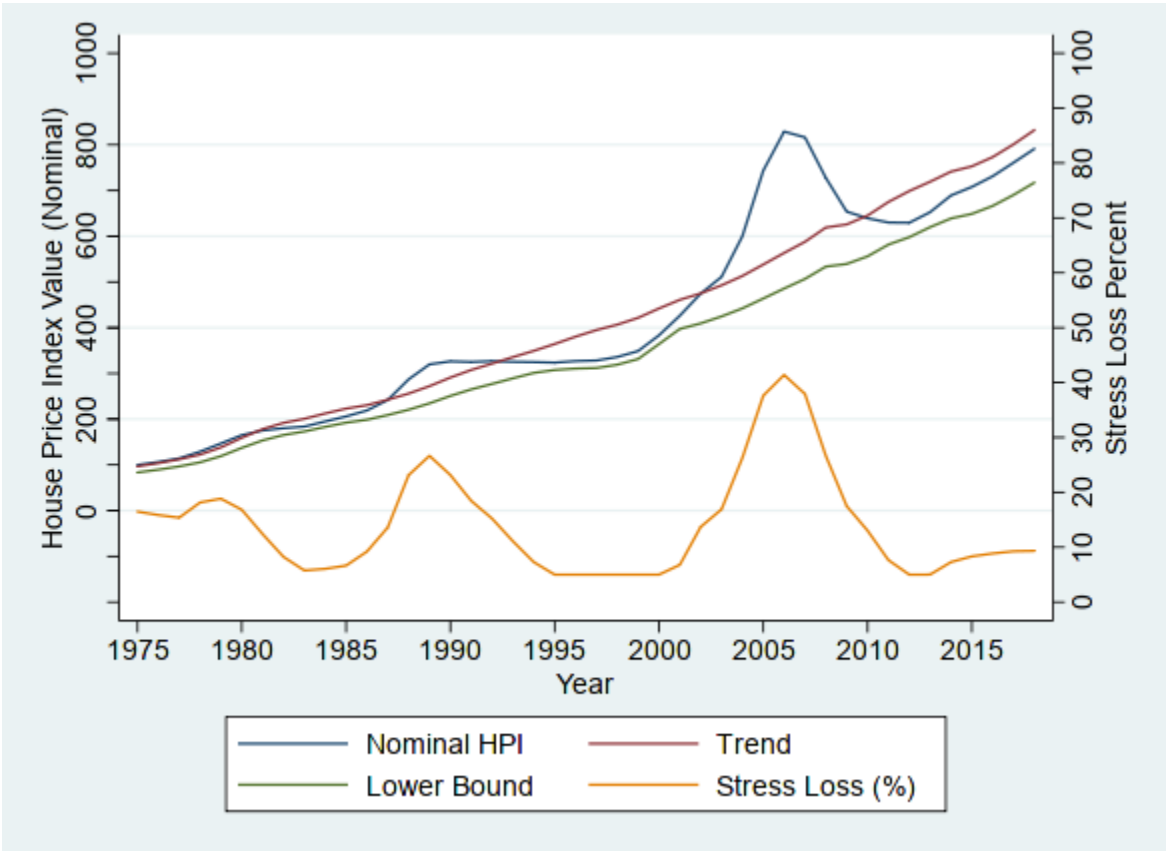
Figure 13: Stressed Default Rate and Key Risk Factors for Low-Credit-Score vs. Higher-Score Home Purchase Loans, Total Market, 1994-2018



Note: Results pertain to 1-4 unit purchase-money mortgages. Data for PLS loans are available through 2017. Low-credit-score loans are defined as those with credit scores below 660; higher-score loans have scores greater than or equal to 660. Missing credit scores are imputed using regression techniques.

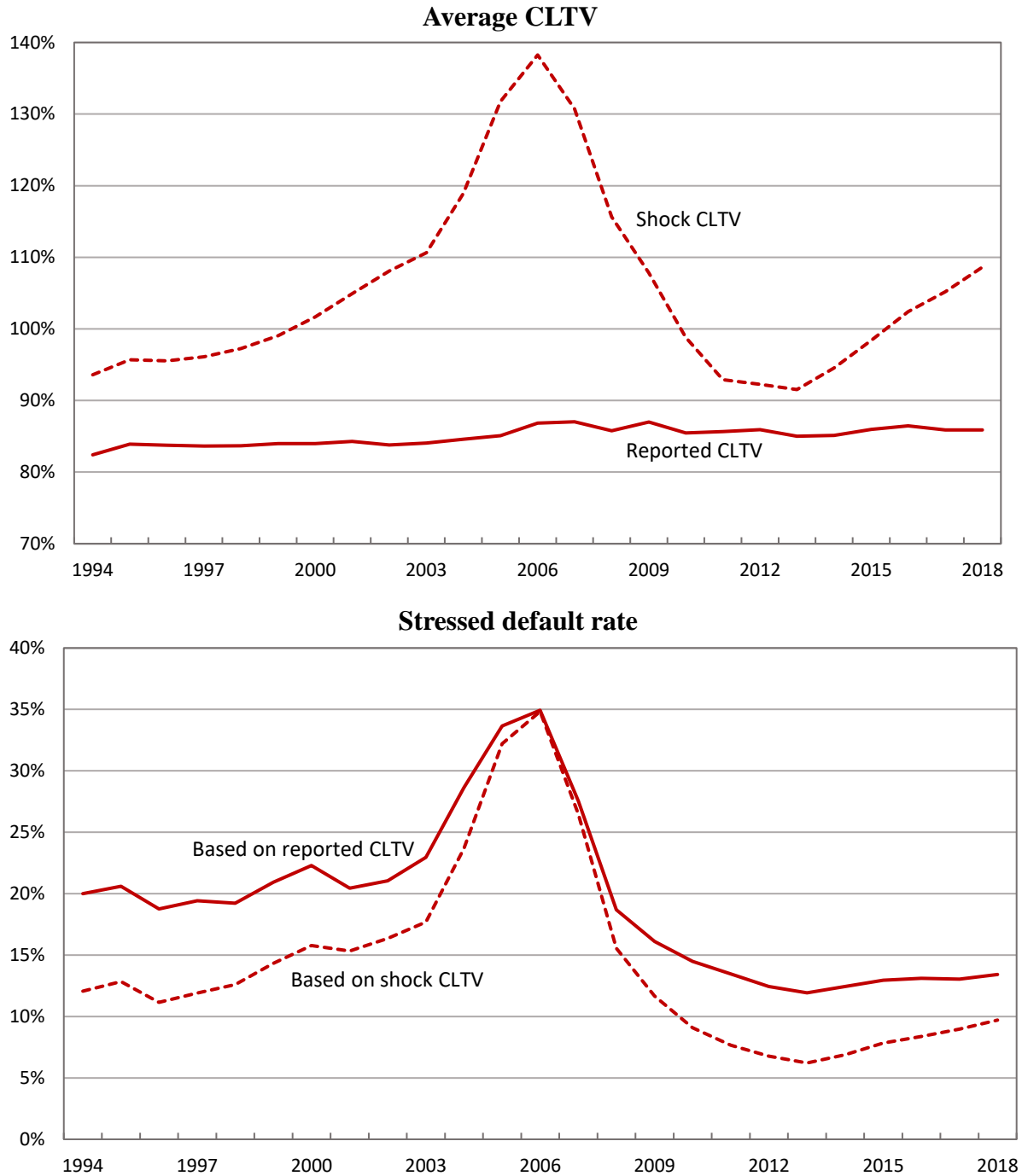
Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

Figure 14: Construction of House Price Shock Series for Washington, DC CBSA



Source: Authors' calculations using the FHFA house price index for CBSA 47900 in Bogin, Doerner, and Larson (2019).

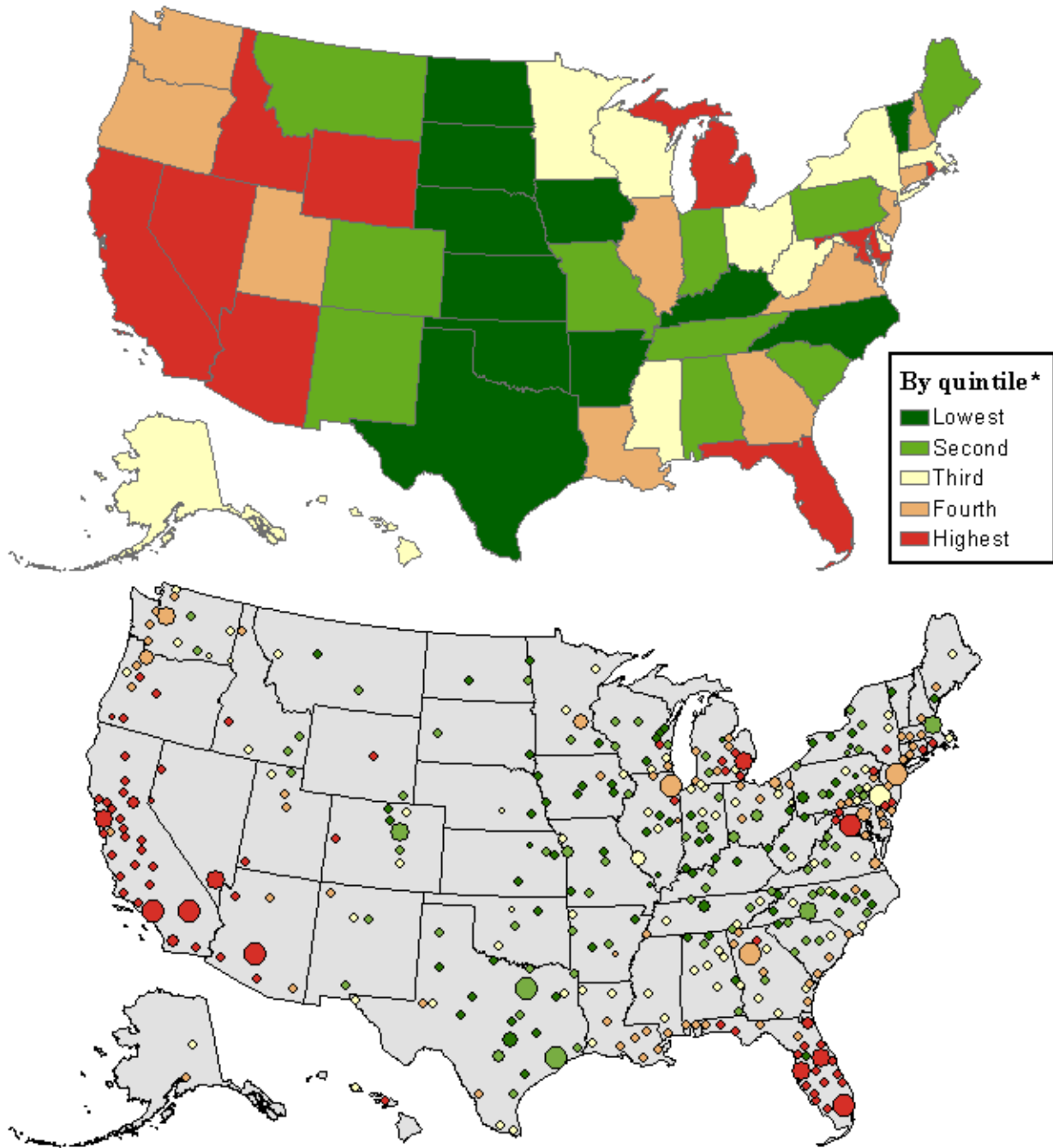
Figure 15: Two Versions of the Average CLTV and Stressed Default Rate for Home Purchase Loans, Total Market, 1994-2018



Note: Results pertain to 1-4 unit purchase-money mortgages. Data for PLS loans are available through 2017. The shock CLTV adjusts the reported CLTV to reflect an estimate of the decline in home value that would occur in a severely stressed scenario.

Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

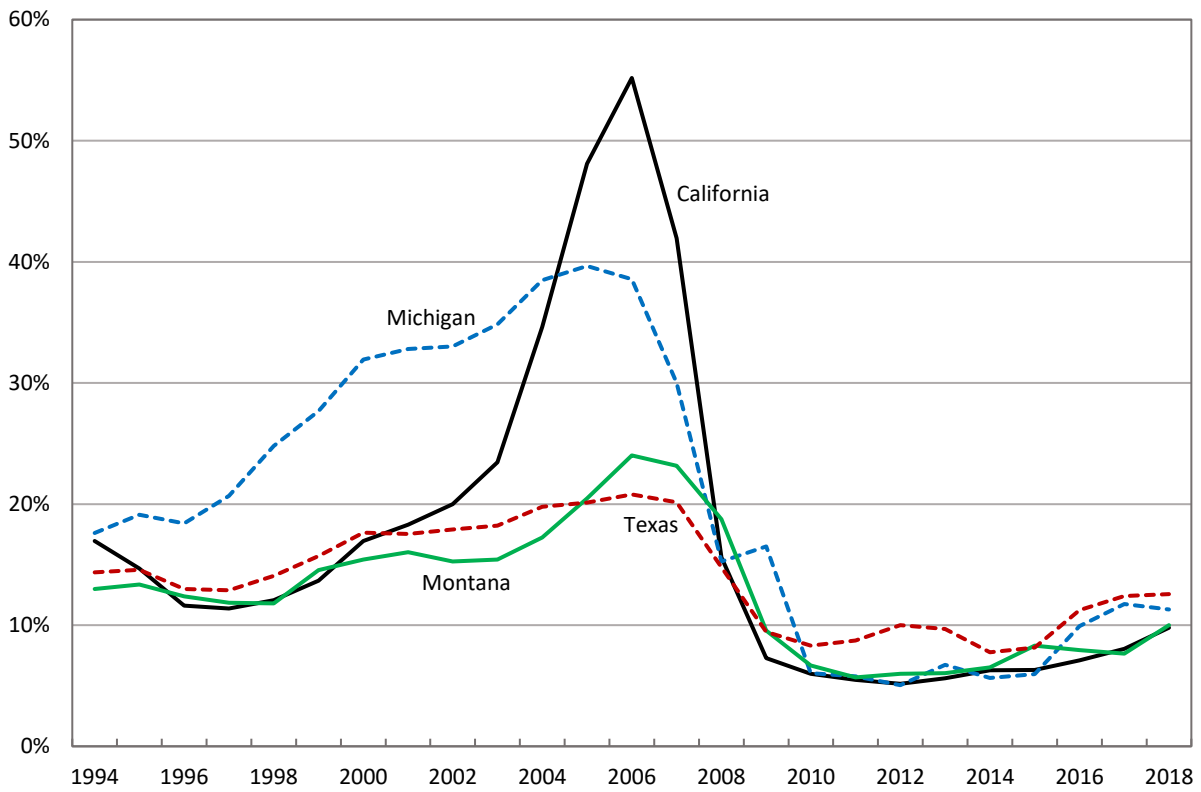
Figure 16: Stressed Default Rate using Shock CLTV, by State and MSA, 2006-2007



Note: Results pertain to 1-4 unit purchase-money mortgages. Range of state stressed default rates by quintile (rounded to the nearest 1%) is 15% to 20%, 20% to 22%, 22% to 26%, 26% to 31%, and 31% to 61%. Range of MSA stressed default rates by quintile (rounded to the nearest 1%) is 8% to 18%, 18% to 21%, 21% to 25%, 25% to 34%, and 34% to 70%. Circle sizes represent estimated mortgage lending volume in 2006 and 2007.

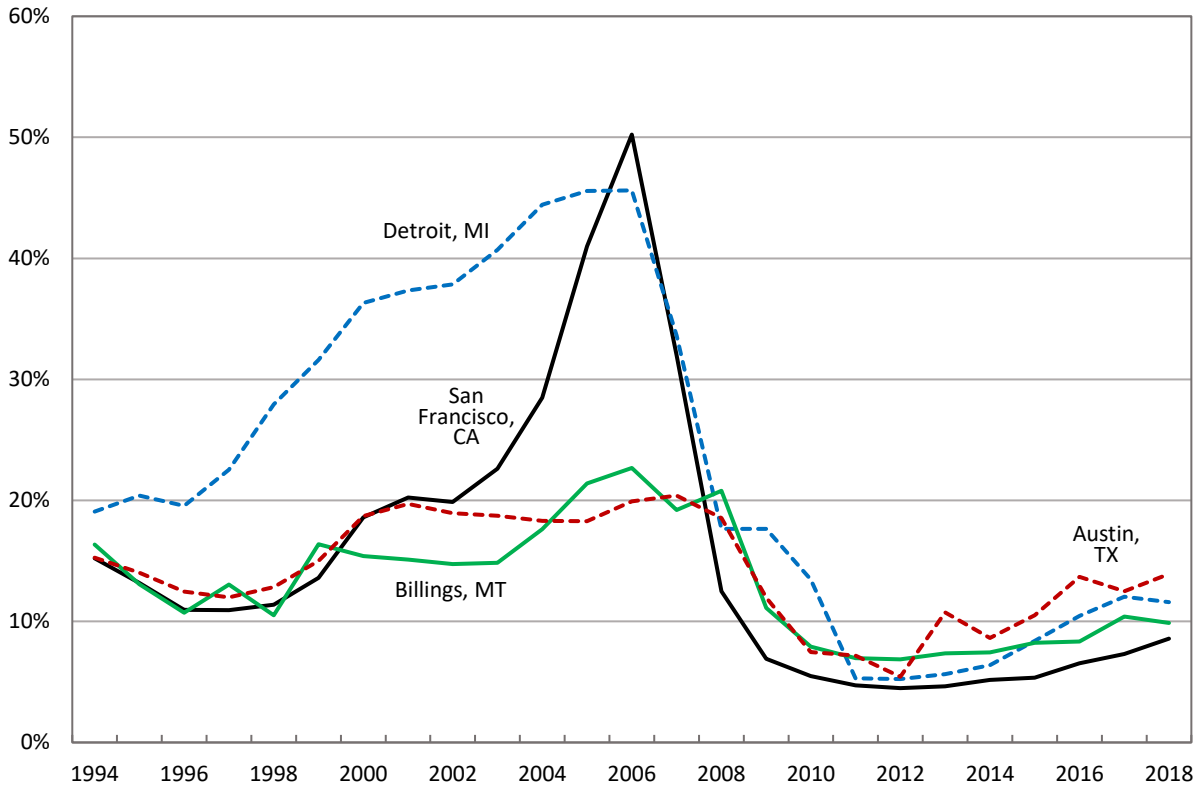
Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

Figure 17: State-Level Stressed Default Rate for Home Purchase Loans, using Shock CLTV, 1994-2018



Note: Results pertain to 1-4 unit purchase-money mortgages. Data for PLS loans are available through 2017.
Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

Figure 18: MSA-Level Stressed Default Rate for Home Purchase Loans, using Shock CLTV, 1994-2018



Note: Results pertain to 1-4 unit purchase-money mortgages. Data for PLS loans are available through 2017.
Source: Authors' calculations using data from FHFA (for Enterprise loans), Black Knight, Inc. (for portfolio and FHA/VA loans), CoreLogic (for portfolio, PLS, and FHA/VA loans), and Ginnie Mae data processed by the AEI Housing Center (for FHA/VA loans).

References

- Adelino, Manuel, W. Scott Frame, and Kristopher Gerardi.** 2017. "The Effect of Large Investors on Asset Quality: Evidence from Subprime Mortgage Securities." *Journal of Monetary Economics* 87: 34-51.
- Adelino, Manuel, Antoinette Schoar, and Felipe Severino.** 2016. "Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class." *Review of Financial Studies* 29(7): 1635-70.
- Adelino, Manuel, Antoinette Schoar, and Felipe Severino.** 2017. "Dynamics of Housing Debt in the Recent Boom and Bust." *NBER Macroeconomics Annual* 32: 261–311.
- Albanesi, Stefania, Giacomo De Giorgi, and Jaromir Nosal.** 2017. "Credit Growth and the Financial Crisis: A New Narrative." NBER Working Paper 23740.
- Amromin, Gene, Jennifer Huang, Clemens Sialm, and Edward Zhong.** 2018. "Complex Mortgages." Forthcoming in *Review of Finance*.
- An, Xudong, Raphael W. Bostic, Yongheng Deng, Stuart A. Gabriel.** 2007. "GSE Loan Purchases, the FHA, and Housing Outcomes in Targeted, Low-Income Neighborhoods." *Brookings-Wharton Papers on Urban Affairs*: 205-40.
- Anenberg, Elliot, Aurel Hizmo, Edward Kung, and Raven Molloy.** 2018. "Measuring Mortgage Credit Availability: A Frontier Estimation Approach." Working paper, August 24. <https://sites.google.com/site/edwardkung/>
- Avery, Robert B., Raphael W. Bostic, Paul S. Calem, and Glenn B. Canner.** 1996. "Credit Risk, Credit Scoring, and the Performance of Home Mortgages." *Federal Reserve Bulletin* 82 (July): 621-48.
- Avery, Robert B., Kenneth P. Brevoort, Tim Critchfield, Ian H. Keith, Ismail E. Mohamed, Forrest W. Pafenberg, Saty Patrabansh, Jay D. Schultz, and Claudia E. Wood.** 2017. "National Mortgage Database Technical Report 1.2." NMDB Technical Report Series, October 30. https://www.fhfa.gov/PolicyProgramsResearch/Programs/Documents/NMDB-Technical-Report_1.2_10302017.pdf
- Bhutta, Neil, and Benjamin J. Keys.** 2018. "Eyes Wide Shut? The Moral Hazard of Mortgage Insurers during the Housing Boom." NBER Working Paper 24844, July.
- Bogin, Alexander N., William M. Doerner, and William D. Larson.** 2019. "Local House Price Dynamics: New Indices and Stylized Facts." *Real Estate Economics* 47(2): 365-98.
- Breiman, Leo.** 2001. "Random Forests." *Machine Learning* 45(1): 5-32.
- Breiman, Leo, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone.** 1984. *Classification and Regression Trees*. Chapman and Hall/CRC: Boca Raton, FL.

- Bubb, Ryan, and Alex Kaufman.** 2014. "Securitization and Moral Hazard: Evidence from Credit Score Cutoff Rules," *Journal of Monetary Economics* 63: 1-18.
- Conklin, James, W. Scott Frame, Kristopher Gerardi, and Haoyang Liu.** 2018. "Villains or Scapegoats? The Role of Subprime Borrowers in Driving the Housing Boom." Federal Reserve Bank of Atlanta Working Paper 2018-10, August.
- CoreLogic.** 2017. "Housing Credit Index: First Quarter 2017". CoreLogic Insights Blog, June 20. <https://www.corelogic.com/blog/2017/06/housing-credit-index-first-quarter-2017.aspx>
- Courchane, Marsha J., Leonard C. Kiefer, and Peter M. Zorn.** 2015. "A Tale of Two Tensions: Balancing Access to Credit and Credit Risk in Mortgage Underwriting." *Real Estate Economics* 43(4): 993-1034.
- DeFusco, Anthony A., Stephanie Johnson, and John Mondragon.** 2017. "Regulating Household Leverage." Kellogg School of Management, Guthrie Center for Real Estate Research Working Paper, October.
- Demyanyk, Yuliya, and Otto Van Hemert.** 2011. "Understanding the Subprime Mortgage Crisis." *Review of Financial Studies* 24(6): 1848–80.
- Dunsky, Robert M., Austin Kelly, and Ken Lam.** 2013. "Impacts of Down Payment Underwriting Standards on Loan Performance – Evidence from the GSEs and FHA Portfolios." FHFA Working Paper 13-3, December. https://www.fhfa.gov/policyprogramsresearch/research/paperdocuments/2013-12_workingpaper_13-3-508.pdf.
- Ferreira, Fernando, and Joseph Gyourko.** 2015. "A New Look at the Foreclosure Crisis: Panel Data Evidence of Prime and Subprime Borrowers from 1997 to 2012." NBER Working Paper 21261.
- Firestone, Simon, Robert Van Order, and Peter Zorn.** 2007. "The Performance of Low-Income and Minority Mortgages." *Real Estate Economics* 35(4): 479-504.
- Foote, Christopher, Kristopher S. Gerardi, Lorenz Goette, and Paul S. Willen.** 2010. "Reducing Foreclosures: No Easy Answers." In *NBER Macroeconomics Annual 2009*, edited by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, University of Chicago Press, pp. 89-138.
- Foote, Christopher L., Lara Loewenstein, and Paul S. Willen.** 2018. "Technological Innovation in Mortgage Underwriting." Working paper, October 19.
- Foote, Christopher L., and Paul S. Willen.** 2018. "Mortgage-Default Research and the Recent Foreclosure Crisis." *Annual Review of Financial Economics* 10: 59-100.

- Fout, Hamilton, Grace Li, Mark Palim, and Ying Pan.** 2018. "Credit Risk of Low Income Mortgages." Forthcoming in *Regional Science and Urban Economics*.
- Garmaise, Mark J.** 2018. "Alternative Mortgage Contracts and Affordability – Overview." Forthcoming in *Regional Science and Urban Economics*.
- Goldman Sachs.** 2014. "Are Mortgage Lending Standards Easing?" Goldman Sachs Credit Strategy Research, July 11.
- Goodman, Laurie, Ellen Seidman, and Jun Zhu.** 2014. "VA Loans Outperform FHA Loans. Why? And What Can We Learn?" Urban Institute working paper.
<https://www.urban.org/sites/default/files/publication/22776/413182-VA-Loans-Outperform-FHA-Loans-Why-And-What-Can-We-Learn-.PDF>
- Haughwout, Andrew, Richard Peach, and Joseph Tracy.** 2008. "Juvenile Delinquent Mortgages: Bad Credit or Bad Economy?" *Journal of Urban Economics* 64(2): 246-57.
- Herzog, John P., and James S. Earley.** 1970. *Home Mortgage Delinquency and Foreclosure*. National Bureau of Economic Research General Series No. 91. New York: Columbia University Press. <https://papers.nber.org/books/herz70-1>.
- Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti.** 2017. "The Mortgage Rate Conundrum." NBER Working Paper 23784. (decline in mortgage spreads)
- Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vikrant Vig.** 2010. "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans." *Quarterly Journal of Economics* 125(1): 307–362.
- Keys, Benjamin J., Amit Seru, and Vikrant Vig.** 2012. "Lender Screening and the Role of Securitization: Evidence from Prime and Subprime Mortgage Markets." *The Review of Financial Studies* 25(7): 2071-2108.
- LaCour-Little, Michael, Charles A. Calhoun, and Wei Yu.** 2011. "What Role Did Piggyback Lending Play in the Housing Bubble and Mortgage Collapse?" *Journal of Housing Economics* 20: 81-100.
- Leventis, Andrew.** 2014. "The Relationship between Second Liens, First Mortgage Outcomes, and Borrower Credit: 1996-2010." FHFA Working Paper 14-3, September.
https://www.fhfa.gov/PolicyProgramsResearch/Research/PaperDocuments/WP_14-3_Second_Liens.pdf
- Li, Wei, and Laurie Goodman.** 2014. "Measuring Mortgage Credit Availability Using Ex-Ante Probability of Default." Urban Institute working paper, November.
<https://www.urban.org/sites/default/files/publication/33581/2000018-Measuring-Mortgage-Credit-Availability-Using-Ex-Ante-Probability-of-Default.pdf>

- Little, Roderick J. A.** 1988. "Missing-Data Adjustments in Large Surveys." *Journal of Business Economics and Statistics* 6(3): 287-96.
- Mahoney, Peter E., and Peter M. Zorn.** 1997. "The Promise of Automated Underwriting: Freddie Mac's Loan Prospector." *Housing Finance International* 11(4): 13-21.
- Mayer Christopher, Karen Pence, and Shane M. Sherlund.** 2009. "The Rise in Mortgage Defaults." *Journal of Economic Perspectives* 23(1): 27-50.
- Mian, Atif, and Amir Sufi.** 2009. "The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis." *Quarterly Journal of Economics* 124(4): 1449-96.
- Mian, Atif, and Amir Sufi.** 2014. *House of Debt: How They (and You) Caused the Great Recession, and How We Can Prevent It from Happening Again*. University of Chicago Press.
- Mian, Atif, and Amir Sufi.** 2017a. "Household Debt and Defaults from 2000 to 2010: The Credit Supply View." In *Evidence and Innovation in Housing Law and Policy*, edited by Lee Anne Fennell and Benjamin J. Keys, 257-88. Cambridge University Press.
- Mian, Atif, and Amir Sufi.** 2017b. "Fraudulent Income Overstatement on Mortgage Applications during the Credit Expansion of 2002 to 2005." *Review of Financial Studies* 30(6): 1832-64.
- Mian, Atif, and Amir Sufi.** 2018. "Finance and Business Cycles: The Credit-Driven Household Demand Channel." *Journal of Economic Perspectives* 32(3): 31-58.
- Newberger, Harriet.** 2011. "FHA Lending: Recent Trends and Their Implications for the Future." Federal Reserve Bank of Philadelphia, Community Development Studies and Education Discussion Paper, December. https://www.philadelphiafed.org/-/media/community-development/publications/discussion-papers/discussion-paper_fha-lending-trends-and-implications.pdf
- Pacelle, Mitchell.** 1991. "Quick Home Loans Have Quickly Become Another Banking Mess." *Wall Street Journal*. July 5.
- Rubin, Donald B.** 1976. "Inference and Missing Data." *Biometrika* 63(3): 581-92.
- Rubin, Donald B.** 1986. "Statistical Matching Using File Concatenation with Adjusted Weights and Multiple Imputations." *Journal of Business & Economic Statistics* 4: 87-94.
- Rubin, Donald B.** 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.
- Quercia, Roberto G., and Michael A. Stegman.** 1992. "Residential Mortgage Default: A Review of the Literature." *Journal of Housing Research* 3(2): 341-79.

- Scheessele, Randall M.** 1998. "HMDA Coverage of the Mortgage Market," U.S. Department of Housing and Urban Development, Office of Policy Development and Research, Working Paper No. HF-007, July. https://www.huduser.gov/portal/Publications/pdf/wkgpapr_7.pdf
- Sichelman, Lew.** 1990. "Fannie Mae Sets 30% Down on Low-doc Loan." Chicago Tribune. September 13.
- Smith, Scott, Debra Fuller, Alex Bogin, Nataliya Polkovnichenko, and Jesse Weiher.** 2016. "Countercyclical Capital Regime Revisited: Tests of Robustness." *Journal of Economics and Business* 84: 50-78.
- Smith, Scott, and Jesse Weiher.** 2012. "Countercyclical Capital Regime: A Proposed Design and Empirical Evaluation." Federal Housing Finance Agency Working Paper 12-2, April. <https://www.fhfa.gov/PolicyProgramsResearch/Research/Pages/Working-Paper-12-2.aspx>
- StataCorp.** 2013. "Stata multiple-imputation reference manual." College Station, TX, StataCorp, LP. <https://www.stata.com/manuals13/mi.pdf>
- Szymanoski, Edward, Randall Scheessele, Christian Malagon, and Ishmail Mohammed.** 2011. "A Look at the FHA's Evolving Market Shares by Race and Ethnicity," *U.S. Housing Market Conditions, 1st Quarter 2011*, May: 6-12. U.S. Department of Housing and Urban Development, Office of Policy Development and Research. <https://www.huduser.gov/portal/periodicals/ushmc/spring11/ch1.pdf>
- Urban Institute.** 2018. Housing Credit Availability Index. <https://www.urban.org/policy-centers/housing-finance-policy-center/projects/housing-credit-availability-index>
- White, Ian. R., Rhian Daniel, and Patrick Royston.** 2010. "Avoiding Bias Due to Perfect Prediction in Multiple Imputation of Incomplete Categorical Variables." *Computational Statistics and Data Analysis* 54(10): 2267-75.
- Williams, Michael.** 2015. "HMDA Coverage of the Mortgage Market," MPP Professional Paper, University of Minnesota, May 18.

Appendix A: Data Preparation

Scope of the dataset

As noted in the main text, our analysis uses data on first-lien purchase-money mortgages originated in 1990-2018 to purchase 1-4 unit residential properties in the 50 states and the District of Columbia. The purchases include second homes and investment properties in addition to primary owner-occupied homes. The dataset covers all major segments of the mortgage market, including loans guaranteed by the Enterprises, FHA, and VA as well as loans without a government guarantee that are bundled into private-label securities (PLS) or held in the portfolios of banks and other lenders.

Identifying portfolio loans

The portfolio loans used in our analysis are drawn from two servicer datasets, the Loan-Level Market Analytics (LLMA) dataset from CoreLogic and the McDash dataset from Black Knight, Inc. Both datasets identify conventional home purchase loans, but neither one fully identifies the subset of such loans that lenders hold in portfolio (rather than selling to the Enterprises or PLS securitizers). As the first step in identifying these portfolio loans, we include all conventional home purchase loans that are ever reported as private and that are never reported as having been acquired by the Enterprises; we also include loans that are never reported to be private but have loan amounts at origination that are above the applicable Enterprise conforming loan limit. We obtain this information on private/Enterprise status by checking the entire monthly sequence of investor codes for every conventional loan in LLMA and McDash. This initial set of potential portfolio loans still needs to be purged of PLS loans, duplicate copies of true portfolio loans, and loans judged likely to be Enterprise loans after further scrutiny. Before

undertaking that task, we apply a set of data cleaning filters to the potential portfolio loans and to Enterprise, PLS, and FHA/VA loans.

Filtering out loans from the source datasets

Table A.1 summarizes the exclusion of loans from the source datasets. We first remove loans for which any of the following important characteristics are not reported: loan amount, LTV, interest rate, state in which the property is located, and ZIP code. We then remove duplicates of loans that appear more than once in the source data.⁴⁸ Duplicates are common for both the potential portfolio loans and FHA/VA loans because both types of loans are drawn from the combination of LLMA and McDash (except for FHA/VA loans originated after 2012, which come from a separate dataset with no duplicates). Many potential portfolio loans and FHA/VA loans appear in both LLMA and McDash, and some appear more than once within each dataset. We also remove loans from the potential portfolio dataset that we flag as likely to be Enterprise or PLS loans (detailed provided below). After these exclusions, we further clean the dataset by removing loans for which the term, product type (fixed rate versus adjustable rate) or property type is either not reported or reported with incomplete or inconsistent information. In the PLS dataset, we remove the very small share of loans with missing occupancy status; in contrast, in the FHA/VA dataset, we assume loans with missing occupancy status are owner-occupied given that this is the norm for FHA/VA loans, and in the potential portfolio dataset, we impute occupancy status for the loans that are missing this information. We also remove loans with LTVs below 25 percent, which we suspect are second liens incorrectly reported as first liens, and loans with CLTVs above 135 percent, which likely are erroneous. As shown in the row labelled

⁴⁸ Loans with missing information on loan amount, LTV, interest rate, and property location are excluded before we check for duplicates because those loan characteristics are needed for the matching algorithm that identifies duplicates.

“Total percent excluded,” these exclusions taken together remove only about 1-2 percent of Enterprise and PLS loans. A much larger fraction of potential portfolio and FHA/VA loans are excluded, but most of these are duplicate loans or potential portfolio loans we believe to be Enterprise or PLS loans. After netting out these exclusions, the data filters remove about 19 percent of potential portfolio loans and about 5 percent of FHA/VA loans. Overall, the table shows that we use the vast majority of unique loans in the source data.

Detail on removing duplicate loans from the servicer datasets

In our FHA and VA datasets, we attempt to match every FHA loan with every other FHA loan, and do the same for VA loans. If the loans in a given pair do not match on specified characteristics (described below), we declare them not to be duplicates. Because some data are erroneously reported, we allow some differences across the match fields based on extensive testing to validate whether these differences represent truly unique loans or duplicates. When a pair of loans satisfies the match criteria, we deem that pair to be duplicates and exclude one of the loans from the dataset.

Our matching proceeds in two steps:

1. Declare a non-match for a given pair of loans when any of the following is true: the 3-digit ZIP code is different, the loan type is different, the origination date differs by more than one month, the interest rate differs by more than five basis points, or the rounded original loan amount differs by more than \$1,000. These non-matching pairs are deemed not to be duplicates.
2. For all remaining pairs of loans, score their similarity across a given set of fields. The fields used to score the loans are LTV, CLTV, DTI, credit score, product type, occupancy type, loan term, loan performance, origination month, original loan amount, 5-digit ZIP code, and

interest rate. We declare pairs that match perfectly to be duplicates, and use a scoring algorithm to declare that additional pairs are duplicates if they match on a large majority of the fields, prioritizing the key match variables used in step 1.

The procedure for removing duplicates from the dataset of potential portfolio loans is very similar to that for FHA and VA loans. We repeat step 1 described above but modify step 2 to include the presence of private mortgage insurance as an additional matching field.

When we remove a duplicate from a pair of loans, we use all available information on loan characteristics across the two loans. For example, if one loan in the pair has a reported DTI but the other does not, we keep the reported DTI. In other cases, duplicate loans can have different values for one or more loan characteristics. The differences generally are small – otherwise the two loans would not have had a high match score in step 2 above. In these cases, we have a set of decision rules for which value to keep. When the value for a particular characteristic has been rounded in one loan but not the other, we keep the more specific value. In addition, we keep the highest CLTV as some servicers may not record all liens against the property. Otherwise, the decision rules involve essentially arbitrary selection.

Detail on removing PLS loans from the potential portfolio dataset

Although the previous steps have removed duplicate loans from the potential portfolio dataset, the dataset still contains PLS loans. Thus, the next step is to remove PLS loans as fully as possible by matching to the separate CoreLogic PLS dataset.

We consider two types of matches: one-to-one matches and one-to-many matches. To establish one-to-one matches between loans in the two datasets, we use steps 1 and 2 above and remove loans that match to the PLS dataset. Because our goal is to remove essentially all PLS loans – even if that means losing true portfolio loans – we perform additional one-to-many matches between individual loans in the PLS dataset and multiple loans in the potential portfolio

dataset. For this one-to-many matching, we perform step 1 described above, which looks for matches based on origination month, original loan amount, 3-digit ZIP code, interest rate, and loan type. We know there will not be one-to-one matches at this stage because those loans were removed in the one-to-one matching just completed. For one-to-many matches that involve up to four loans in the potential portfolio dataset, we remove all the matched loans. For matches that involve more than four loans, we do additional diagnostics to identify the four most likely matches and remove those four loans.

Detail on removing Enterprise loans from the potential portfolio dataset

We follow a similar procedure to remove loans from the potential portfolio dataset that likely were acquired by the Enterprises. Earlier, we removed loans that were affirmatively shown by the investor information in LLMA and McDash to have been acquired by the Enterprises. We now look for matches between loans that remain in the potential portfolio dataset and the affirmatively Enterprise loans that were removed before. In effect, we are searching for duplicates of these loans that may have had incomplete investor information in one of the servicer datasets. To establish matches, we use the same process as with PLS loans to remove potential portfolio loans that are one-to-one matches to Enterprise loans.⁴⁹ As additional safeguards against the inadvertent retention of Enterprise loans, we remove conforming loans with investor information that records the loan as private in less than half of the months of its total lifespan,⁵⁰ as well as jumbo loans that became conforming the year after origination.

⁴⁹ Because McDash only reports 3-digit ZIP codes for Enterprise loans, we require loans matching only on a 3-digit ZIP basis to have a higher match score than loans matched on a 5-digit ZIP basis.

⁵⁰ Loans with performance files that fail to track a loan until the end of its lifespan are assumed to have a total lifespan through the most recent performance month.

Weighting of loan data

Because our FHA, VA, and portfolio datasets do not constitute the universe of such home purchase loans, we construct two sets of weights so that the data will be nationally representative at the state and MSA level using adjusted loan counts for originated purchase loans from the Home Mortgage Disclosure Act (HMDA) Loan Application Register. Our national results are then aggregations of the state-level results.

For FHA and VA loans, we partition both our dataset and the HMDA data by year, loan type (FHA or VA), loan amount buckets, and geography. Because HMDA does not include all mortgage loans, we increase the HMDA loan counts to approximate the universe of loans.⁵¹ For FHA and VA loans, we use the estimate of the undercount through 1995 in Scheessele (1998). We then assume HMDA coverage increases monotonically from 93 percent in 1996 to 100 percent in 2010 where it remains through 2018, based on the coverage results for FHA loans in Scheessele (1998) and Szymanoski et al. (2011) and the results for FHA and VA loans together in Williams (2015). For each cell in the partition described above, we use the ratio of the adjusted HMDA loan count to the count in our dataset as the sampling weight.⁵²

For purchase loans held in lenders' portfolios, we use the same partition as for FHA and VA loans. We then estimate the universe of portfolio loans through a series of steps as the estimated universe count of conventional loans minus the counts of Enterprise and PLS loans. The key component of these steps is estimating the universe count of conventional loans as a

⁵¹ For information about the regulatory changes affecting the undercount of mortgages reported in HMDA, see <https://www.ffiec.gov/hmda/history2.htm>

⁵² For a small number of cells, the count in our dataset exceeds the adjusted HMDA count. In those instances, we use our loan count rather than the HMDA-based count, in effect setting the weight to one.

gross-up of the conventional loans reported in HMDA, where the gross-up factor is based on the HMDA undercount of Enterprise loans.

We begin with the count of Enterprise purchase loans reported in HMDA.⁵³ Note that loans sold to the Enterprises after the year of origination are not counted as Enterprise loans in HMDA. To correct for this undercount, we increase the HMDA count of Enterprise purchase loans by the share of such loans acquired after the year of origination according to FHFA data. We then estimate the true HMDA undercount of Enterprise purchase loans as the ratio of the count in the FHFA dataset to the HMDA count that includes the post-origination-year acquisitions. We calculate this ratio for each cell in the partition and assume that the proportional undercount in HMDA for all conventional purchase loans is the same as for Enterprise loans. Figure A.1 shows the estimated annual share of total conventional purchase loans reported in HMDA; we divide the reported annual HMDA counts by the annual shares in Figure A.1 to estimate the total count of conventional purchase loans. Finally, we calculate the universe count of portfolio loans for each cell in the partition as the adjusted HMDA count of conventional loans minus the Enterprise count from the FHFA data and the PLS loan count from the CoreLogic data. Our sampling weight for each cell in the partition equals the ratio of the estimated universe count for portfolio loans to the count in our dataset.

For Enterprise and PLS loans, the underlying source data represent either a universe count (Enterprise loans) or close to a universe count (PLS loans). Our data cleaning removes a small fraction of these loans. To restore the original loan counts, we weight the loans in each cell of the partition by the ratio of the pre-cleaning to post-cleaning count.

⁵³ The Enterprise count includes both loans that are sold directly to the Enterprises and loans that are first sold to another HMDA reporter and then sold to the Enterprises.

Imputation of missing values for risk factors

As described in Section 2 of the paper, we use the multiple imputation approach (Rubin, 1976) to impute missing data. A relevant question is why impute at all: dropping observations with missing information is conceptually much simpler. The key issue is whether the missing information is correlated with other variables. If missing information is not correlated with any observed or unobserved variables, it is classified as missing completely at random (MCAR), in which case observations with missing information can be safely dropped without introducing any bias. However, if missing information is correlated with other observable variables, then dropping observations will result in biased statistics.⁵⁴

If the data were MCAR, we should observe no systematic variation in the likelihood of missing values across market segments (Enterprise, PLS, etc.) or risk characteristics (DTI, credit scores, etc.). However, we do find substantial evidence that missing information in our mortgage data is not MCAR. For instance, Enterprise loans with missing credit scores tend to have lower reported incomes. Based on this evidence, we proceed with an imputation strategy rather than dropping observations to avoid a known bias in the reported statistics. We elect to impute credit score, DTI, documentation status, occupancy status, and the three components of amortization status (balloon status, negative amortization status, and interest-only status).⁵⁵

⁵⁴ The two alternative missing data structures to MCAR are missing at random (MAR) and missing not at random (MNAR). Under MAR, the missing values for a variable x are correlated with other observed variables, but whether an observation of x is missing does not depend on the value of x itself. In contrast, under MNAR, the missingness of x does depend on the value of x . Multiple imputation is suitable for the MAR structure, but not for MNAR, which requires the researcher to model the process that generates missing values. It is not possible to test whether the data are MAR versus MNAR because that would require the missing data to be known. We assume missing data in our datasets are MAR in order to implement multiple imputation. Under MAR, it is important to point out that the fraction of missing values only affects the uncertainty of the imputation.

⁵⁵ The theory of multiple imputation is based on having access to infinite imputations, but according to Rubin (1987), an average of just two imputations has 90% asymptotic efficiency with a 50% rate of missing values. For higher missing rates, the consequence is reduced efficiency but not bias. To assess the accuracy of the imputations for the first and second moments, we have performed a cross-validation exercise for PLS imputations and found them to be closely aligned with observed moments. This allayed our concerns about potential bias in the mean or variance of the imputations.

Within the multiple imputation framework, we use predictive mean matching (PMM) as developed in Little (1988) and Rubin (1986). The PMM procedure uses regression-based predictions to find the nearest “donors” of a given variable for loans with missing values. We randomly select one of the eight closest donors to provide the missing value. We use the PMM procedure for two reasons. First, PMM is well suited for imputing variables that are not normally distributed, such as credit scores and DTIs. And second, PMM avoids the convergence problems that can affect other methods, such as logit models, for predicting categorical variables. For the imputation of occupancy status, we use a conditional imputation model that first imputes if a loan is primary owner-occupied, and then imputes whether it is an investor loan or a second home loan only if it is found not to be primary owner-occupied. To allow the estimated parameters to vary over time, every PMM model is run over the shortest time windows that we judged to have sufficient data for reliable estimation, typically one month, and by loan purpose. When there are insufficient observations in one month for a given loan purpose, we use a mix of quarterly, half-yearly, annual, and grouped-year windows depending on loan counts. See Table A.2 for details.

We use multiple imputation chained equations (MICE) to iteratively impute all of the variables with missing data in our dataset. An iterative procedure is necessary because the structure of our missing data is not monotone (i.e., the variables with missing data cannot be ordered such that the loans with missing data for the first variable are a subset of the loans with missing data for the next variable, and so on.) The imputation equations include a large set of explanatory variables: dummy variables for the state, product type, property type, loan term, documentation status, occupancy status, balloon status, negative amortization status, interest-only status, default status, month of origination, and market segment, as well as fourth-degree

polynomial functions of the interest rate, loan amount, DTI, credit score, reported CLTV, shocked CLTV, and borrower income. We perform ten iterations of the model and then draw the set of imputed values.⁵⁶

To assess the robustness of the imputed values, we repeated the imputation procedure 40 times with both the PLS and portfolio datasets (we did not attempt the same exercise with the much larger Enterprise dataset due to computational limitations). We used these results to construct 95 percent confidence bands for the risk factors imputed for those loans and for the implied stressed default rate. The confidence bands pertain to the full respective sets of loans, including those with reported values for the variable shown. Figures A.2 and A.3 display the results for PLS and portfolio loans, respectively. For PLS loans, the confidence bands are fairly tight except for the average DTI in 1990, 1991, and 1992. Confidence intervals for portfolio loans are narrower than for PLS loans during the 1990s for the average DTI and average credit score, the two heavily imputed risk factors.⁵⁷ Overall, Figures A.2 and A.3 indicate that the model uncertainty associated with loan-level imputations has little effect on the summary statistics we report in the paper.

However, as we discuss in the text, even when there is little model uncertainty, the resulting summary statistics still could be biased if the available source data used to estimate the imputation models are not a random sample. Because of this concern, we elected not to report results for DTIs or credit scores until these risk factors have good coverage in the Enterprise

⁵⁶ In the FHA and VA datasets from 2013 onward, we use number of units rather than property type and the loan's origination channel rather than default status, as property type and default status are not available in the Ginnie Mae data we use for those years. Borrower income is available only in the Enterprise dataset.

⁵⁷ This may seem surprising given that a larger share of portfolio loans require imputations over 1990-2017 as a whole. However, in the early 1990s, when the confidence bands are the widest, the reporting of DTIs and credit scores was worse for PLS loans than for portfolio loans. In addition, in the early and mid-1990s, the number of portfolio loans far exceeded the number of PLS loans, with the greater sample size working to narrow the confidence intervals for the estimated sample means.

data, which helps anchor the results for the purchase loan market as a whole. We also make use of external benchmark data to correct potential biases, as described in the next section.

External benchmarking

The National Mortgage Database (NMDB) serves as our primary source of external benchmark information. The NMDB is a 5 percent random sample of credit files back to 1999 provided by Equifax, one of the national credit repositories; these credit files are linked to other data sources to fill out loan and borrower characteristics beyond those contained in the credit files. Because the NMDB is still undergoing internal quality checks at FHFA, we use only the data for credit scores, which are pulled directly from the credit files, and the data for DTIs, which – when missing – are calculated by NMDB staff based in part on loan information in the credit files. Even for credit scores and DTIs, we benchmark to the NMDB only when the data line up with other data sources known to be reliable or accord with well documented features of the mortgage market. For FHA loans, we also benchmark to information in An et al. (2007) and Newberger (2011). The details of the benchmarking for each market segment are as follows.

Portfolio loans

We benchmark to the annual NMDB distribution of credit scores for portfolio purchase loans from 1999 to 2018. This benchmarking is accomplished by reweighting the portfolio loans in our dataset to replicate the annual NMDB distribution of credit scores across the buckets used for our stressed default tables (i.e., less than 580, 580-619, and so on). We do not utilize the NMDB data on portfolio loan DTIs because of concerns that the NMDB may overstate the share of loans with DTIs above 43 percent in recent years.

PLS loans

As noted in the text, PLS pools after the housing bust included resecuritizations of defaulted PLS loans that were originated in earlier years. The dataset we created from the CoreLogic PLS data likely contains such loans, which could misrepresent the characteristics of newly originated PLS loans. To correct for such distortions, we weight our PLS loans to replicate the annual NMDB bucket distribution for credit scores for PLS loan originations starting in 2008. We do not benchmark to the joint distribution of credit scores and DTIs because the sample of PLS loans in the NMDB is too small to fill out the grid with sufficient loan counts.

FHA and VA loans

We benchmark our DTI results for FHA purchase loans to distributions reported in An et al. (2007) for 1992 and 1996 and to the NMDB for 1999-2018. For 1993 to 1995, we linearly interpolate year by year between the 1992 and 1996 DTI bucket distributions, benchmarking to the interpolated distributions; for 1997 and 1998, we create benchmark distributions by interpolating between the 1996 An et al. distribution and the parallel 1999 NMDB distribution.

For FHA credit scores, we use the annual bucket distributions for 2004-2009 for purchase loans in Newberger (2011).⁵⁸ For those years, we combine the credit score buckets in Newberger with the DTI buckets in the NMDB to create annual benchmark distributions for both risk factors. We reweight our FHA data for those years to match the combination of the credit score and DTI distributions.

⁵⁸ We do not use the NMDB data for FHA purchase-loan credit scores because we found these data did not match the distribution of scores in the Ginnie Mae data for 2013-2018. This difference does not necessarily imply any error in the NMDB data. It could be that the NMDB convention for calculating the credit score for loans with more than one borrower differs from that in the administrative data provided by Ginnie Mae.

We found no benchmark information for VA loans prior to 1999, and we elected not to use the NMDB data for later years, as neither the DTI nor the credit score distributions matched their Ginnie Mae counterparts for the overlap period of 2013-2018. However, the Ginnie Mae data and earlier VA data for 1999-2003⁵⁹ show that the median DTI for VA purchase loans has been on average 2 percentage points below the median for FHA purchase loans, and we adjust the median DTI for VA loans to maintain this gap in other years as well. We make no adjustments to the credit scores for VA purchase loans.

⁵⁹ See table 5-25 in https://www.benefits.va.gov/homeloans/documents/docs/final_report.pdf

Table A.1: Loan Counts and Exclusions of Home Purchase Loans

	Type of Loan			
	Enterprise	Potential Portfolio	PLS	FHA/VA
Initial count (millions)	52.34	19.38	8.57	43.22
<i>Percent excluded for reason shown</i>				
Loan amount not reported	0.06	0.01	<0.01	0.01
LTV not reported	<0.01	4.74	0.26	6.50
Interest rate reported	<0.01	0.99	1.07	1.02
State not reported	<0.01	0.49	0.18	0.29
ZIP code not reported or not valid	<0.01	1.06	1.11	0.39
Duplicate loans	<0.01	24.04	0.00	34.82
Removal of PLS Loans	NA	15.96	NA	NA
Removal of Enterprise Loans	NA	8.59	NA	NA
Term not reported	0.01	0.15	0.21	0.07
Product type not reported or reported inconsistently	<0.01	0.63	0.04	0.36
Property type not reported or not specified	0.03	7.56	0.33	4.64
Occupancy status not reported	0.00	NA	0.11	NA
LTV < 25 percent	0.65	9.24	0.14	0.05
CLTV > 135 percent	<0.01	0.37	0.28	0.06
Total percent excluded	0.75	67.69	2.68	39.72
Final count (millions)	51.96	6.26	8.34	26.05

Note: The counts represent first lien mortgages originated in 1990-2018 used to purchase 1-4 unit properties. The percent of loans excluded is rounded to the nearest 0.01. The total percent excluded will be less than the sum of the percent excluded in the preceding rows because some loans are excluded for more than one reason. NA in the row for occupancy status indicates that we did not exclude these loans when occupancy status was missing. As noted in the text, we assumed all FHA/VA loans with missing data were primary owner-occupied, and we imputed missing occupancy status for potential portfolio loans.

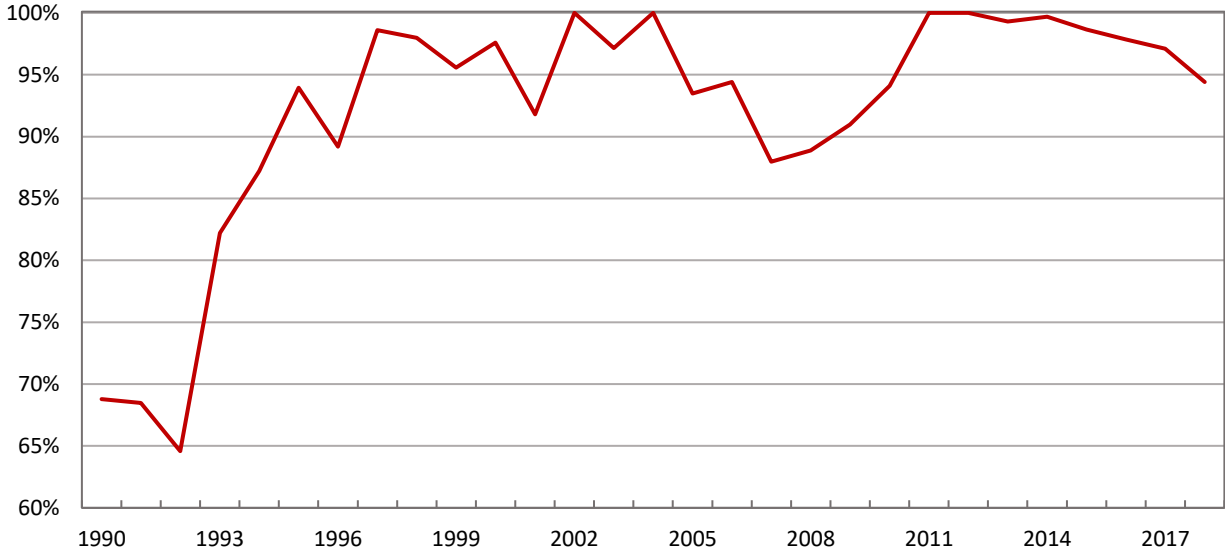
Source: Authors' calculations using data from Black Knight, Inc., CoreLogic, FHFA, and Ginnie Mae data processed by the AEI Housing Center.

Table A.2: Estimation Periods for Imputation Regressions

Risk factors	Estimation periods
Enterprise	
Credit score & DTI	Grouped months: 1990-1994:Jan., 1990-1994:Feb., ... , 1990-1994:Dec.
	Monthly: 1995:Jan. to 2018:Dec.
PLS	
Credit score, DTI, documentation status	Monthly: 1999:Jul. to 2007:Jun.
	Quarterly: 1998:Q3, 1998:Q4, 1999:Q1, 1999:Q2
	Half-yearly: 1997:H1, 1997:H2, 1998:H1, 2007:H2
	Yearly: 1994
	Grouped years: 1990-93, 1995-96, 2008-2012, 2013-2014, 2015-2017
Portfolio	
Credit score, DTI, documentation status, amortization status, occupancy status	Monthly: 2003:Oct. to 2007:Jun.
	Quarterly: 2002:Q3 to 2003:Q3, 2007:Q3 to 2008:Q2, 2014:Q3 to 2018:Q4
	Half yearly: 1995:H1 to 2002:H1, 2008:H2, 2011:H1 to 2014:H1
	Yearly: 1990, 1991, 1992, 1993, 1994, 2009, 2010
FHA/VA	
Credit score & DTI, 1997-2018	Monthly: 1998:Jan. to 2018:Dec.
	Quarterly: 1997:Q1 to 1997:Q4
Credit score, 1990-96	Quarterly: 1992:Q1 to 1996:Q4
	Yearly: 1990, 1991
DTI, 1990-96	Grouped years: 1990-1996*

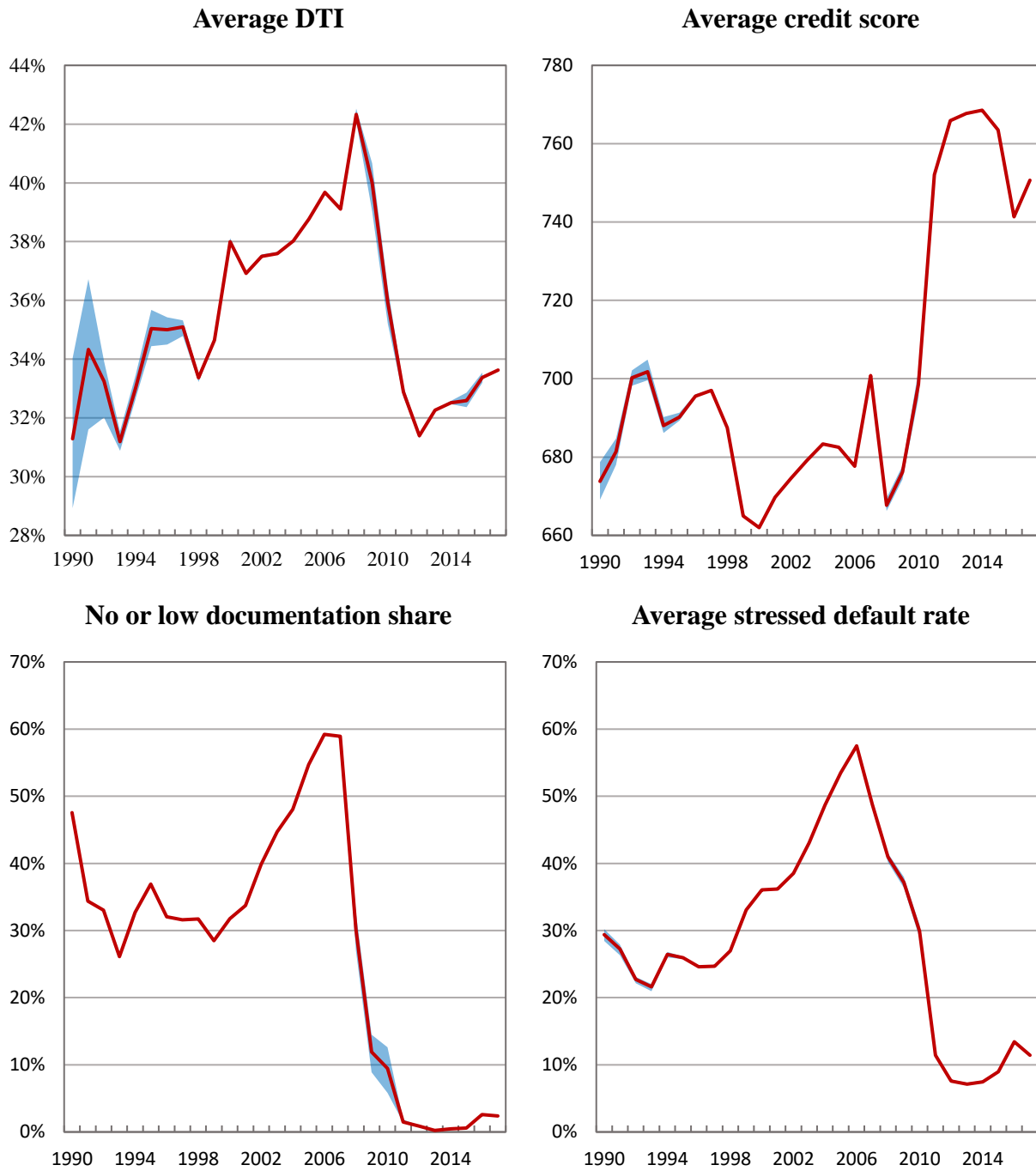
* For both FHA and VA loans, a large share of the reported DTIs from 1990 through 1995 have the same value, which seems implausible. Thus, we estimate the DTI imputation regression with data for 1996 and use the 1996 loans as potential donors to impute DTIs for loans originated in 1990-1995.

Figure A.1: Estimated Share of Total Conventional Purchase Loans Reported in HMDA (percent)



Source: Authors' calculations using data from HMDA and FHFA.

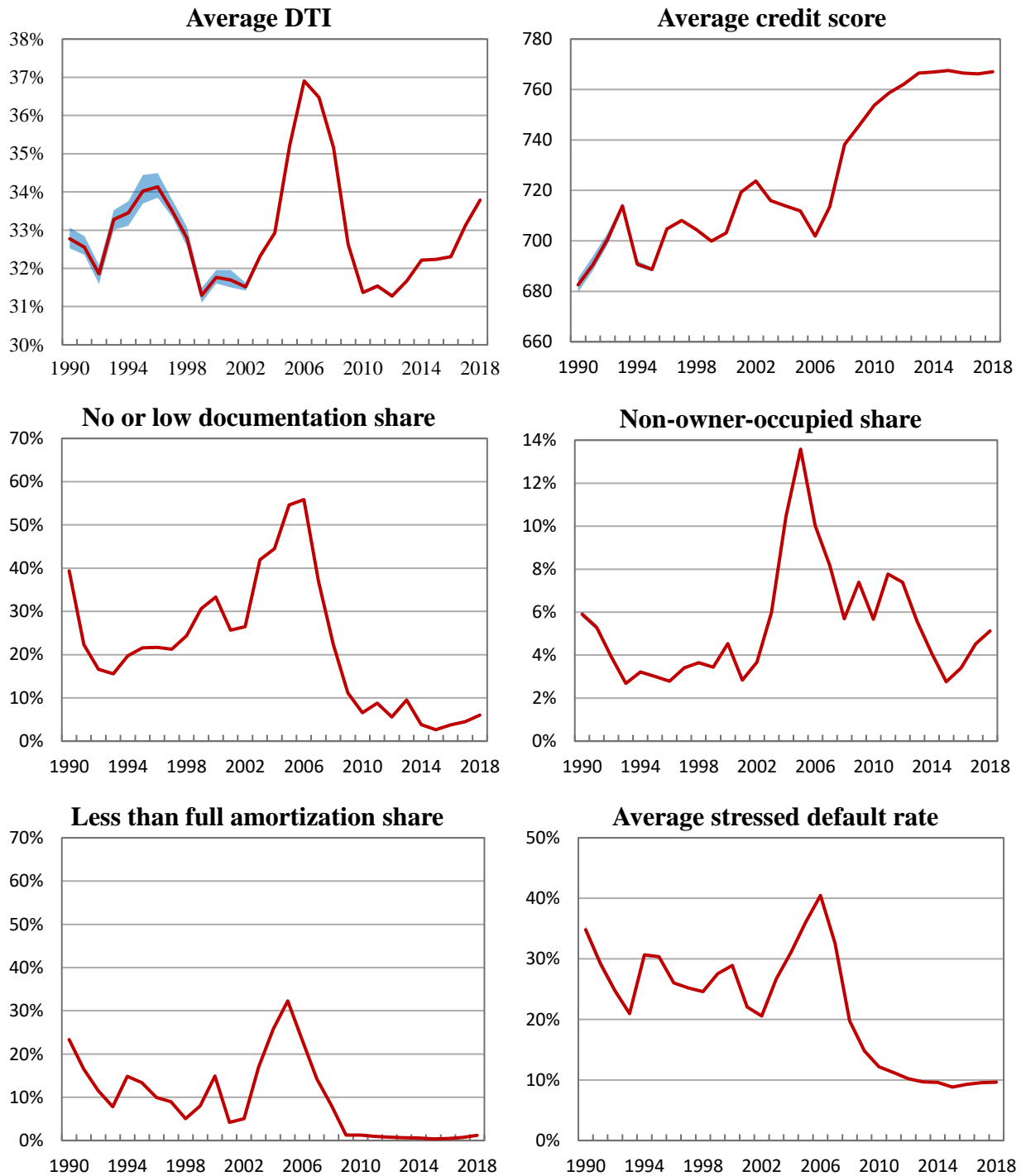
Figure A.2: Estimated Characteristics of PLS Purchase Loans, 1990-2017



Note: Results pertain to 1-4 unit purchase-money mortgages. Data for PLS loans are available through 2017. Area in blue represents the 95% confidence interval that accounts for the imputation of missing values for risk factors.

Source: Authors' calculations using data CoreLogic.

Figure A.3: Estimated Characteristics of Portfolio Purchase Loans, 1990-2018



Note: Results pertain to 1-4 unit purchase-money mortgages and are calculated from the unweighted source data. Area in blue represents the 95% confidence interval that accounts for the imputation of missing values for risk factors.

Source: Authors' calculations using data from Black Knight, Inc. and CoreLogic.

Appendix B: Robustness Tests

To test the robustness of our baseline default-table methodology, we construct three additional indicators of mortgage risk for PLS and Enterprise loans using a logit model and two machine learning models. For all three models, we use the same set of risk factors as in the default tables and implement the models using home purchase loans originated in 2006 and 2007.⁶⁰

We specify the logit model as

$$d_i = \Phi(\boldsymbol{\beta}\boldsymbol{\delta}_{im} + \varepsilon_{im})$$

where Φ is the logit function; d_i is a binary variable capturing whether loan i defaulted⁶¹; $\boldsymbol{\delta}_{im}$ is a matrix of explanatory variables that includes the borrower's credit score, CLTV, and DTI as numerical values, and loan term, occupancy status, amortization status, income documentation status, and loan type (fixed-rate versus adjustable rate) as categorical dummy variables; $\boldsymbol{\beta}$ is a vector containing the coefficient estimates; and ε_{im} is the error.

The second model uses the same specification of $\boldsymbol{\delta}_{im}$ to build a regression tree based on the idea first proposed in Breiman, Friedman, Olshen, and Stone (1984).⁶² The model, which performs functionally like a decision tree, compares splits of the data on all eight risk characteristics to find which one produces the most productive split and then repeats this step until only 500 loans are left in each bucket.⁶³ Then, the model builds 20 additional trees using 95 percent of the data, with the remaining 5 percent kept for cross-validation. The results from the

⁶⁰ The results reported in this appendix have not been updated from the March 2019 version of the paper. An update will be included in the next version.

⁶¹ The definition of default for these models is the same as in the baseline default-table approach: we deem a loan to have defaulted if it was ever 180 days' delinquent or was terminated with the borrower forfeiting the property.

⁶² The package `rpart` was used to create these trees. For more on recursive partitioning, see: <https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf>.

⁶³ For an overview of productive splits, see chapter 3 of the link in the previous footnote.

20 cross-validations determine the amount of tree complexity associated with the maximum predictive power. We then apply this complexity parameter to the initial decision tree, which “prunes” branches from the tree that are caused by spurious correlation. In the end, the final regression tree built using PLS loans has 438 decision nodes and 469 terminal nodes. The 469 terminal nodes are the counterpart in the regression tree to the 2,560 cells in the default tables.

Figure B.1 depicts the first several splits in the PLS regression tree. The ordering of the splits provides information about the relative importance of the various risk factors for explaining defaults in the PLS market. The initial split is based on the loan’s credit score, indicating that it is the single most informative risk factor. In all, credit scores account for 41 percent of the tree’s fit, followed by CLTVs at 28 percent, amortization status at 10 percent, income documentation status at 8 percent, with the remaining risk factors making up the rest. While regression trees can be used to impute missing variables, we choose not to do this as simulations showed that loans run through the model from the 1990s have different correlations among risk factors than loans used to build the model from 2006 and 2007. Instead, we use the imputation methodology as described in Appendix A.

We also build a high-performance random forest to capture further relationships between variables in our dataset.⁶⁴ Random forests, proposed in Breiman (2001), are an extension of regression trees. We construct our forests as follows: first, we randomly select three of our eight risk factors; second, we construct a training sample by picking observations with replacement such that approximately two-thirds of the total sample is drawn to construct a given tree; and finally, we build a regression tree using those observations and variables. We repeat these steps 500 times to construct a proverbial “forest” of decision trees. By using only a subset of our risk

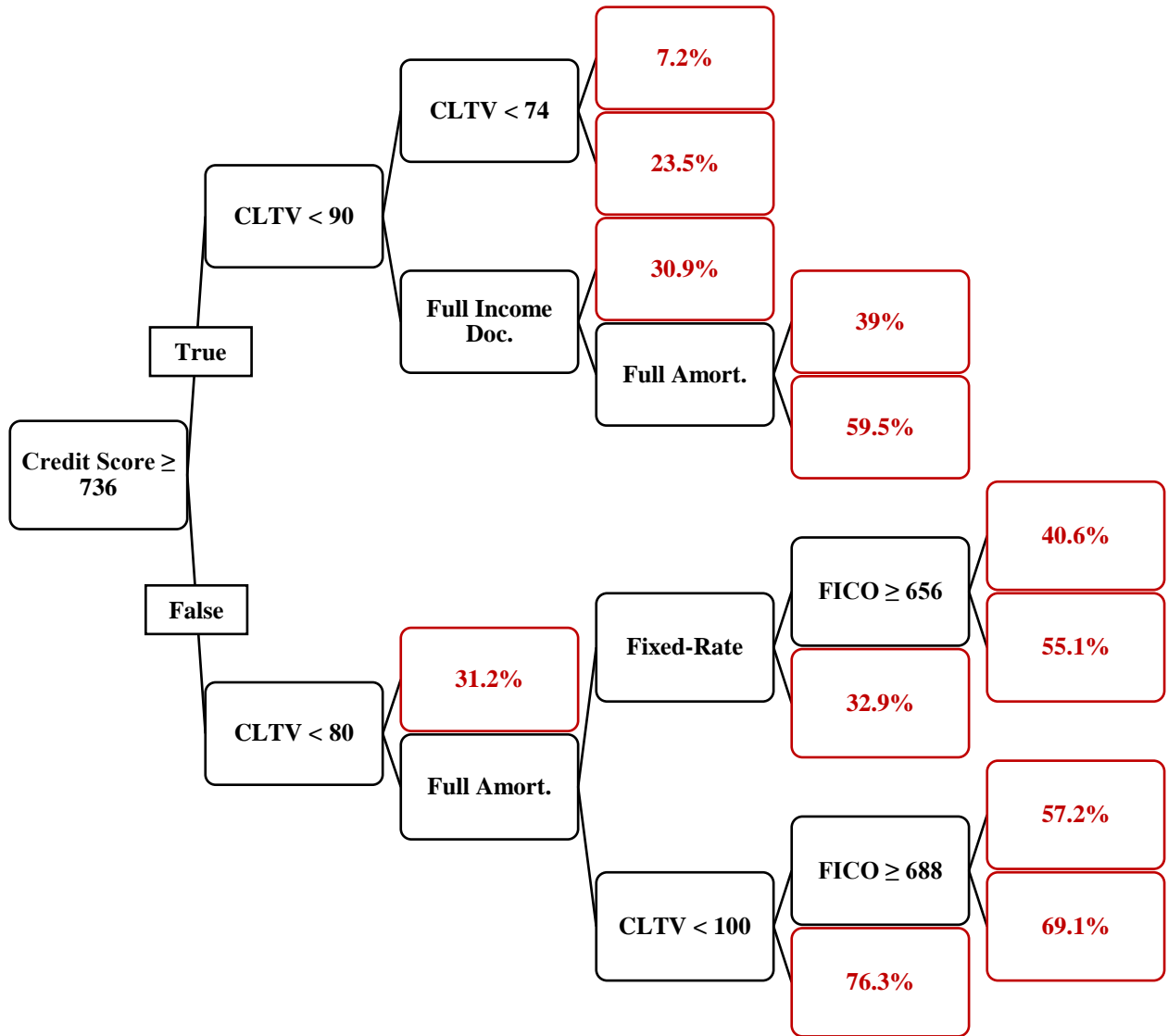
⁶⁴ The package Rborist was used to implement this model. For more on random forests, see: <https://cran.r-project.org/web/packages/Rborist/index.html>.

factors and observations in every tree, each one captures a unique relationship among the selected risk factors. Additionally, selecting only a portion of loans to build each decision tree allowed us to validate our model using the out-of-sample observations. The random forest model built using PLS loans confirms many of the findings from the regression tree, notably that credit scores and CLTVs are the most highly predictive risk factors for default.

For each of the three alternative models (logit, regression tree, and random forest), we generated stressed default rates that can be compared to those from our baseline methodology. To create these alternatives default series, we calculate the predicted default rate for every loan originated from 1990 to 2017 using the results for each model. We then compute the average predicted default rate for each origination year, just as in our baseline methodology.

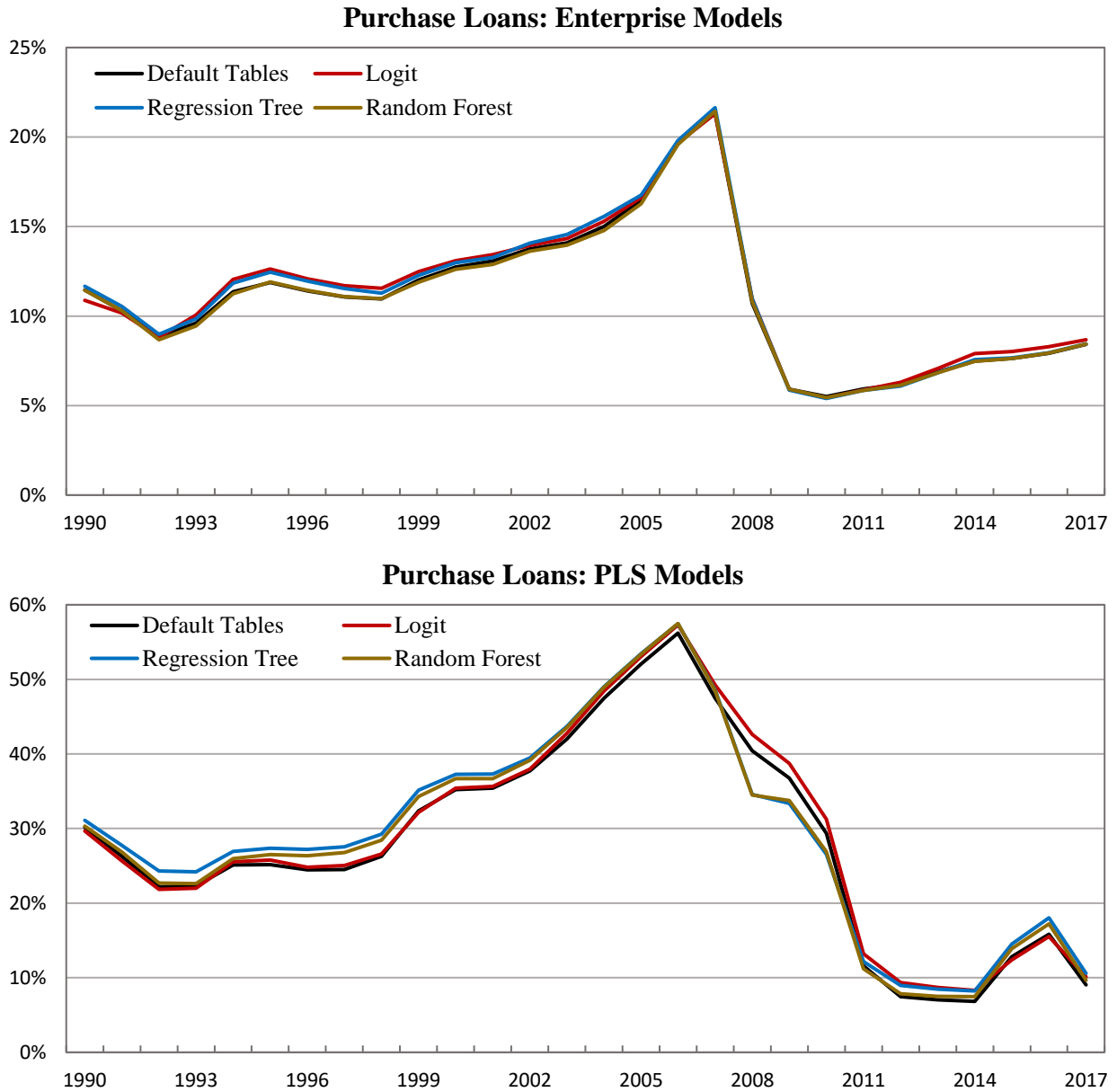
Figure B.2 presents the results from this comparison. The figure shows a high correlation between our baseline results and the three alternatives. For Enterprise loans, the four stressed default series are almost indistinguishable from one another. Although there are wider differences among the PLS series – most notably in 2008 and 2009 when the regression tree and random forest models find a faster decline in mortgage risk – they all trace the same contour over 1990-2017. The bottom line from this exercise is that our baseline results are robust to the use of alternative default models.

Figure B.1: Visualization of the First (Most Important) Splits in the PLS Regression Tree



Note: For each split, the upward arm indicates that the condition shown in the decision node is true, while the downward arm indicates that the condition is false. Black represents decision nodes, while red represents terminal nodes. The value in every terminal node represents the default rate for loans in that node. Results pertain to 1-4 unit purchase-money mortgages in PLS originated in 2006 and 2007.
Source: Authors' calculations using data from CoreLogic.

Figure B.2: Stressed Default Rates from Different Models, Enterprise and PLS Purchase Loans, 1990-2017



Note: Results pertain to 1-4 unit purchase-money mortgages.
Source: Authors' calculations using data from FHFA and CoreLogic.