working paper

18 16

Technological Innovation in Mortgage Underwriting and the Growth in Credit: 1985–2015

Christopher L. Foote, Lara Loewenstein, and Paul S. Willen

FEDERAL RESERVE BANK OF CLEVELAND

ISSN: 2573-7953

Working Paper 18-16 December 2018

Technological Innovation in Mortgage Underwriting and the Growth in Credit: 1985–2015

Christopher L. Foote, Lara Loewenstein, and Paul S. Willen

The application of information technology to finance, or "fintech," is expected to revolutionize many aspects of borrowing and lending in the future, but technology has been reshaping consumer and mortgage lending for many years. During the 1990s computerization allowed mortgage lenders to reduce loan-processing times and largely replace human-based assessment of credit risk with default predictions generated by sophisticated empirical models. Debt-to-income ratios at origination add little to the predictive power of these models, so the new automated underwriting systems allowed higher debt-to-income ratios than previous underwriting guidelines would have typically accepted. In this way, technology brought about an exogenous change in lending standards, which helped raise the homeownership rate and encourage the conversion of rental properties to owner-occupied ones, but did not have large effects on housing prices. Technological innovation in mortgage underwriting may have allowed the 2000s housing boom to grow, however, because it enhanced the ability of both borrowers and lenders to act on optimistic beliefs about future house-price growth.

Keywords: Mortgage underwriting, housing cycle, technological change, credit boom.

JEL Codes: C55, D53, G21, L85, R21, R31.

Suggested citation: Foote, Christopher L., Lara Loewenstein, and Paul S. Willen. 2018. "Technological Innovation in Mortgage Underwriting and the Growth in Credit: 1985–2015." Federal Reserve Bank of Cleveland, Working Paper no. 18-16. https://doi.org/10.26509/frbc-wp-201816.

Christopher L. Foote is at the Federal Reserve Bank of Boston (Chris.Foote@bos.frb.org). Lara Loewenstein is at the Federal Reserve Bank of Cleveland (Lara.Loewenstein@clev.frb.org). Paul Willen is at the Federal Reserve Bank of Boston (Paul.Willen@bos.frb.org). For helpful comments, the authors thank Ted Durant, Benjamin Keys, Robert Armstrong Montgomery, Elizabeth Murry, Ed Pinto, Eric Zwick, and seminar participants at the Federal Reserve Bank of Cleveland, Freddie Mac, Northwestern, the 2018 AREUEA Mid-Year conference, the Southern Finance Association's 2017 annual conference, the 2018 Federal Reserve Day-Ahead Conference on Financial Markets and Institutions, and the conference on Financial Stability Implications of New Technology at the Federal Reserve Bank of Atlanta. Robert Avery also kindly supplied data.

At first glance, Louise Beyler of Gainesville, GA, might appear as an unlikely candidate for a mortgage to buy a $105,500 home. Self-employed and earning $19,000 a year, Beyler would have to spend nearly 45 percent of her income to cover the mortgage payments. Given her circumstances, many lenders would deny Beyler a mortgage. Thanks to automated underwriting, however, Beyler's application was approved—in just three days.

—from Peter Mahoney and Peter M. Zorn,
"The Promise of Automated Underwriting" (1996)

1 Introduction

How do exogenous changes in mortgage-lending standards affect the housing market? This question is central to understanding the boom-bust housing cycle of the 2000s, but it is difficult to answer because lending standards are endogenous with respect to expected house-price growth. In this paper, we study a specific application of information technology to financial services, or "fintech," which affected mortgage-lending standards and thereby provides some leverage for studying the effects of changes in credit policies. Specifically, we examine how the computerization of mortgage lending during the 1990s led to fundamental changes in the evaluation of credit risk, and then ask how the new credit standards affected the housing market.

Although the recent literature on the housing cycle has focused on the 2000s, the 1990s are widely recognized as a period of intense technical change in mortgage lending (LaCour-Little 2000; Bogdon 2000; Colton 2002). During this decade, mortgage lenders leveraged rapid gains in computing power and data-processing capabilities to originate mortgages faster and at a lower cost than before. The first contribution of the paper is therefore to provide a broad characterization of these changes using some new data and new empirical approaches. We confirm that the 1990s advances in information technology not only sped up mortgage processing but also enhanced the ability of mortgage lenders to compete geographically. The share of mortgages originated by out-of-state lenders rose from around 10–20 percent in 1990 to around 60 percent by the end of the decade. Some of these improvements have recently reversed course, as lenders adapt to new lending regulations intended to forestall another crisis. For example, the time needed to close a standard refinance application fell from around two months in the early 1990s to less than 30 business days by the close of the decade. After falling to near 20 days by 2005, the average required time-to-close has increased to about 50 days in 2018. A qualitatively similar pattern exists for the financial costs of mortgage origination that are borne by borrowers; large declines during the 1990s have been reversed in the years following the financial crisis.

Our second main finding concerns a change that remains in place today: technological advances begun in the 1990s have fundamentally altered the way that mortgage lenders evaluate credit risk. Before 1990, loan officers and originators evaluated mortgage applications by applying so-called knockout rules, which specified maximal cutoffs for variables such as the LTV (loan-to-value) ratio and the DTI (debt-to-income) ratio.[1] This type of rules-based system would seem to be tailor-made for replacement by computers, which have transformed the US economy due to their ability to perform routine tasks efficiently (Levy and Murnane 2004; Acemoglu and Autor 2011). In fact, when the 1990s began many lenders tried to use computers in precisely this way, by encoding their lending rules into formal algorithms that computers could follow. The artificial intelligence (AI) systems that resulted would then evaluate loans just as humans always had, but at a lower cost.

The coders soon discovered that despite the rules-based nature of loan evaluation, intuitive human judgment typically came into play. Making decisions was easy for human underwriters when loan applications cleared all of the relevant hurdles. But for marginal cases the underwriters had to use their own experience and judgment when deciding whether, say, a high DTI ratio could be offset by a low LTV ratio, or by a particularly strong credit history. The developers of one early AI system wrote that:

> Underwriting is considered an art that is taught from one generation of underwriters to the next. It takes years of experience to learn to judge and evaluate the loan information to make an informed decision on the risk of the loan. Underwriters do not follow a systematic, step-by-step approach to evaluate a loan. Instead, the underwriter must look at the strengths and weaknesses of the individual elements in a loan file and evaluate how all the data elements affect one another. The process is intuitive rather than scientific. The challenge for the knowledge engineers was to represent the thought process of an underwriter in a manner that could be implemented as a software system (Talebzadeh, Mandutianu, and Winner 1995, p. 54–55).

Even if human thought processes could be codified, a big problem with basing lending decisions on human judgment is that mortgage defaults are rare. Individual underwriters were therefore unlikely to accumulate enough personal experience with defaults to properly quantify the tradeoffs among a loan's risk characteristics.

Other firms in the mortgage industry used computers differently. They pulled together large datasets of loan-level records and estimated empirical models of mortgage default, which

[1] For mortgage lenders the DTI ratio denotes the ratio of the borrower's monthly *payment* to her monthly income, and does not involve the borrower's entire stock of debt. Although it is not quite accurate, this DTI definition is so ingrained in the mortgage industry that we stick with it here.

could then be used to augment the evaluations of loan files made by underwriters. These modelers soon realized that default predictions could be significantly enhanced by using credit scores, such the FICO score, which were then being developed to predict defaults on unsecured loans. As credit scores worked their way into mortgage-lending decisions, the housing-finance industry was transformed by the same technologies that were revolutionizing the credit-card industry and other forms of consumer lending at the same time (Evans and Schmalensee 2005; Einav, Jenkins, and Levin 2013; Livshits, Gee, and Tertilt 2016).

In contrast to credit scores, DTI ratios added little to the new mortgage models. As we explain below, modern theories of mortgage default—which are based on income *shocks* rather than initial income *levels*—are consistent with this lack of predictive power. Indeed, the small effect that DTIs have on default probabilities had been foreshadowed by previous, smaller-scale empirical studies, including the ur-study of loan-level default sponsored by the National Bureau of Economic Research in the late 1960s (Herzog and Earley 1970). In the 1990s, however, the unimportance of DTIs began to influence the allocation of mortgage debt. Informed by their empirical models, mortgage lenders discounted the relationship between a borrower's income and her mortgage obligation by allowing larger loan sizes at the lower end of the income distribution.

To illustrate this point, we draw on recent research on mortgage debt and borrower income during the 2000s boom (Mian and Sufi 2009; Adelino, Schoar, and Severino 2016). Because an exogenous relaxation in the debt-income relationship allows lower-income households to take out larger loans, it should "flatten" the upwardly sloping cross-sectional relationship between debt and income at the individual level. We confirm the finding of Adelino, Schoar, and Severino (2016) that such a flattening did not occur in the 2000s; in that decade, larger loans were taken out by borrowers across the income distribution.[2] But extending the sample back in time, we also show that the relationship between debt and income did flatten in the 1990s, when technological advances were changing lending practices. As the quotation at the top of this introduction indicates, automation not only sped up mortgage lending, it also changed the relationship between debt and income in ways that remain relevant today.

The paper's third contribution exploits the exogenous nature of 1990s-era developments by studying their consequences beyond that decade. As we discuss below, the improvements in lending technology that occurred during the 1990s did not come about because the housing market was particularly hot or cold at that time. Rather, mortgage underwriting became increasingly automated after 1990 because mortgage lenders had finally figured out the best way to use computers, and because information technology had finally advanced to the point where large-scale empirical models had become feasible. Because the 1990s developments

[2]See Foote, Loewenstein, and Willen (2016) for an analysis of the stock of mortgage debt across the income distribution during the housing boom, as opposed to the flow of new mortgage debt analyzed here.

were driven by technology, not by the state of the housing market, they provide researchers with a good opportunity to study the effects of exogenous changes in lending standards on housing-market outcomes.

By and large, we find that the effects of these exogenous changes are consistent with economic theory. For example, relaxing DTI limits should have the largest effects on individuals whose expected future incomes are high relative to their current incomes. Young college graduates have relatively steep age-earnings profiles, and we find that the increase in homeownership after 1994 was particularly large for this group. In this respect, the credit expansion of the 1990s played out as a miniature version of the much larger credit expansion between 1940 and 1960. As shown by Fetter (2013), the mid-century credit expansion, driven primarily by the Federal Housing Administration (FHA) and the Veterans Administration (VA), also had its largest effect on the young. Fetter argues that the mid-century policies essentially allowed individuals who would have purchased homes eventually to buy them at an earlier stage in the life cycle. We find the same pattern for the much smaller expansion of mortgage credit that occurred in the 1990s.

We then investigate the link between the exogenous change in lending standards and housing prices. As pointed out by Kaplan, Mitman, and Violante (2017), among others, an exogenous credit expansion raises the potential demand for owner-occupied homes in relation to rented homes and apartments. The relative increase in owner-occupied demand can be satisfied by new construction and by the conversion of rental units into owner-occupied ones. To the extent that rising owner-occupied demand is satisfied by either or both of these supply margins, then the credit expansion will not raise the prices of owner-occupied properties. To investigate the responsiveness of these potential margins, we use the American Housing Survey (AHS), a dataset that is also helpful in our study of mortgage debt and income. Because the AHS follows a panel of the same housing units over many years, we can track the evolution of tenure status for a given unit over time. When we construct "gross flows" of properties in and out of owner-occupied status, we find that the net flow of rented to owner-occupied properties, which is usually negative, becomes positive in the 1990s, consistent with the expansion of supply predicted by theory.

Perhaps because of this supply response, during the late 1990s house prices in lower-income zip codes—where formerly credit-constrained individuals would be expected to buy homes—display no tendency to rise relative to house prices in higher-income areas in the same cities. Low-income zip codes did experience larger-than-average house-price increases after 2002, as the US housing boom gained strength. But this relative price change was likely a natural consequence of the national rise in house prices; as housing became increasingly expensive, affordability concerns led buyers to bid up the prices of homes in cheaper areas. In fact, at the national level housing prices have risen faster than income since 2012, and prices in low-income areas are again rising relative to prices in high-income areas, even though overall credit conditions tightened significantly after the boom (Goodman 2017). The rising prices today argue against the claim that during the boom, price increases in low-income areas were driven primarily by looser credit standards.

Taken together, these findings highlight a potential tension between technological progress at the micro level and boom-bust cycles at the macro level. As we discuss in our historical review, there is ample evidence that harnessing computing power to reduce the importance of human judgment in mortgage-lending decisions led to more efficient predictions of default. The success of model-based lending explains the efforts of today's fintech engineers, who are devising even more sophisticated algorithms to predict default probabilities using machine learning and other techniques (Buchak et al. 2018; Fuster et al. 2017, 2018; Philippon 2018). Although these improvements did not seem to have significant direct effects on housing prices, they could have exacerbated a housing cycle that was driven by some other force, such as overly optimistic expectations for future housing prices at the national level. A central purpose of any financial system is to match people who want to lend with people who want to borrow, and during the 2000s widespread expectations of higher housing prices made borrowers eager to buy homes and lenders eager to lend the money to so.[3] New technologies facilitated these transactions, so they could have played a part in the housing boom by allowing individuals to act more easily on optimistic beliefs about further house-price gains. In the end, the boom-bust housing cycle in the 2000s may have resulted not from a housing-finance system that worked poorly, but one that worked too well.

The remainder of the paper is organized as follows. Section 2 discusses the two main data sources used to quantify productivity improvements in mortgage lending and to show how these changes influenced the relationship between household income and mortgage debt. Section 3 quantifies the significant improvements in mortgage lending during the 1990s, and section 4 focuses on how the debt-income relationship changed over time. Section 5 provides a historical review of the ways that credit scores and the automated lending models transformed the evaluation of default risk in the 1990s. Section 6 traces out the effects of these changes on the homeownership rate and on housing prices in the United States, and section 7 concludes with a discussion of some potential implications of our findings for models of the housing cycle.

[3] For examples of research along these lines, see Gennaioli, Shleifer, and Vishny (2012); **?**; Burnside, Eichenbaum, and Rebelo (2016); Adelino, Schoar, and Severino (2016); Albanesi, De Giorgi, and Nosal (2017); and Gennaioli and Shleifer (2018).

2 Data

2.1 The Home Mortgage Disclosure Act (HMDA)

Our first dataset comes from the Home Mortgage Disclosure Act (HMDA), a 1975 law that requires financial institutions in metropolitan areas to report individual-level information relating to mortgage applications and originations. Variables in the public-use version of HMDA include the dollar amount of each new mortgage; the reported income, race, and gender of the borrower; and the census tract of the house that serves as collateral for the loan.

Some information is not available in HMDA on a consistent basis because of changes in reporting guidelines, including some relatively major changes that occurred in 2004.[4] For example, starting in 2004 we can distinguish between first and second liens. During the 2000s housing boom, first-lien mortgages were often supplemented by second liens at purchase ("piggybacks"), so it is important to account for second liens consistently when making comparisons that span several years. We therefore created an algorithm to identify second liens made before 2004, and then validated this rule using the reported liens available starting in 2004. The algorithm makes use of the application and origination dates of each mortgage, which are available only in a confidential version of the HMDA data to which we gained access.[5] Even when second liens are identified in HMDA, they are not matched to their corresponding first liens. Our algorithm both identifies second liens and matches them to their first liens. We then use the sum of the first and second liens for each borrower in the loan-level regressions.

The HMDA data also include a field that distinguishes owner-occupiers from investors.[6] In the main empirical work below we remove investors from our regressions, in large part because the relationship between income and mortgage size for investors is likely to be very different than that for owner-occupiers. For investment properties, the income backing a mortgage comes not only from the borrower's resources, but also from the rental income the property is expected to generate.

We then clean the HMDA data along the lines suggested by Avery, Brevoort, and Canner (2007), who suggest that analysts drop loans that lack information on race and gender (as these are probably business loans) and combine home improvement loans with refinances.

[4] A good summary of these changes can be found in "HMDA Changes Are On the Way; New Rules Take Effect in 2004," which appears in *Community Dividend*, an online publication of the Federal Reserve Bank of Minneapolis. It is available at https://www.minneapolisfed.org/publications/community-dividend/hmda-changes-are-on-the-way-new-rules-take-effect-in-2004.

[5] Details of the second-liens algorithm appear in Appendix A.1.

[6] This field is based on information supplied by the borrower, so someone purchasing an investment property might intentionally misreport as an owner-occupier in order to get a lower mortgage rate.

Additionally, because miscoded outliers can exert a strong effect in our debt-income analysis, we remove outliers with an algorithm designed to detect inaccurate loan or income entries. The algorithm calculates the implied monthly payment of each loan (assuming that it is a 30-year fixed-rate mortgage at the current interest rate), and then divides this implied payment by the income on the record to generate an implied DTI ratio. We then drop loans with DTI ratios in the bottom and top 1 percent of this distribution.

Figure 1 depicts some summary data from purchase-mortgage originations in HMDA. The top panel shows the total number of purchase loans originated along with the fraction of those loans that have associated second liens. The number of purchase loans originated increases steadily from 1990 until 2005, while the use of piggyback loans grows dramatically in the mid-2000s, near the end of the housing boom. After 2007, both series drop sharply. The bottom panel displays the median total loan amount for owner-occupied purchase mortgages, with this total encompassing both first liens and any associated piggybacks. The panel also shows the share of purchase mortgages made to owner-occupiers, which declines by almost 10 percentage points during the boom. Note also that there is a hump shape in the median loan amount at the peak of the boom, when the piggyback share and the investor share were highest.

A potential problem with income information in HMDA is that it reflects only what the mortgage lender verified in order to qualify the borrower for the loan, which may not necessarily reflect the household's total income. Mortgage lenders have long favored income that can be documented and reasonably expected to continue into the future. If the borrower is purchasing a relatively inexpensive property, then the mortgage broker may not go to the trouble of documenting the stability of any income that is not needed for the loan to be approved.[7] Consequently, when house prices rise, incomes reported on HMDA records may also rise, because borrowers require additional income sources to qualify for the larger loans. This fact makes individual-level income as reported on HMDA records potentially endogenous with respect to housing values. Additionally, as pointed out by Mian and Sufi (2017), borrowers or mortgage brokers may misrepresent income to lenders if they cannot document enough income to qualify for loans through legitimate means.[8] In light of potential

[7]Page D-10 of the 2013 Guide to HMDA Reporting states that: "An institution reports the gross annual income relied on in evaluating the credit worthiness of applicants. For example, if an institution relies on an applicant's salary to compute a debt-to-income ratio but also relies on the applicant's annual bonus to evaluate creditworthiness, the institution reports the salary and the bonus to the extent relied upon."

[8]Some evidence on the potential endogeneity of incomes reported in the HMDA data comes from a comparison of HMDA incomes to census income data. The decennial US Census and the American Community Survey (ACS) include data on the incomes of homeowners with a mortgage who have recently moved. Avery et al. (2012) find that average incomes reported on HMDA records were up to 30 percent higher than incomes reported in the ACS in 2005 and 2006 in five states: California, Hawaii, Massachusetts, Nevada, and New York. By comparison, incomes in HMDA in 2000 were no more than 10 percent above those in the 2000 decennial census, and HMDA incomes from 2009 and 2010 were no more than 10 percent above those in the

discrepancies in HMDA incomes, in our main HMDA regressions we instrument for income using the median household income by census tract from the decennial census and American Community Survey (ACS).

2.2 The American Housing Survey (AHS)

Another source of both income and mortgage-debt data is the American Housing Survey (AHS), which began in 1973 and has been conducted in each odd-numbered year since 1981. We use AHS data after a significant redesign in 1985.[9] The AHS is a joint project of the US Department of Housing and Urban Development (HUD) and the Census Bureau that is designed to measure the size, quality, and composition of the American housing stock. The survey also measures monthly housing costs for US residents and collects information on demographics and income of sampled households.

The AHS can be used to analyze the flow of new purchase-mortgage debt because the survey includes information about the size of any existing mortgages *at the time those mortgages were originated.* Our sample consists of homeowners who moved into their residences since the last AHS survey, for which the current mortgage is likely to be the purchase mortgage. Although there is no question on the AHS asking whether the current mortgage on the house is the actual purchase mortgage, we do know whether the mortgage was taken out in the same year that the house was purchased, and we only use those mortgages. For income, the AHS includes both wage income and the household's total income. We use total income in our regressions, but our results are essentially identical when only wage income is used.[10] Because the AHS income measures come from surveys, not mortgage applications, they are much less likely to be influenced by housing prices.

As we do with the HMDA data, we take precautions regarding spurious outliers in the AHS, where problems can arise for two reasons. First, consider a new homeowner appearing in the 1991 AHS who moved into his residence in 1990, and then retired or significantly cut back his hours in 1991. His 1991 income would be much lower than the 1990 income used to qualify for his mortgage. Some recent homeowners who report only a few thousand of dollars of income in the survey year likely fit this scenario. A second problem arises from the topcoding of income and debt. Like many household surveys, the AHS reports a

ACS. These time-series and geographical patterns are consistent with a positive relationship between house prices and reported HMDA incomes, although it is unclear how much of this correlation comes from the need to document additional income as opposed to outright fraud.

[9] For a general description of the AHS see U.S. Department of Housing and Urban Development (2017). For discussions of the AHS as a data source for housing-finance issues, including new mortgage debt, see Lam and Kaul (2003), Eggers (2007), and Eggers (2008).

[10] The similarity of results using either total or wage income is not surprising in light of the exclusion of capital gains from the total-income measure. Because capital gains are rarely expected to continue year after year, this source of income is typically not used to quality for mortgage loans.

household's income level only if it falls below some upper limit, which changes across survey years. Mortgage debt is treated analogously, and income and debt levels above the topcode limits are replaced with allocated values.[11] By truncating the top and bottom 5 percent of the debt and income distributions, we found that we could exclude all topcoded debt and income values and also generate reasonable minimum income levels for home purchasers. As we show in the appendix, however, our results are materially unchanged (albeit somewhat noisier) when we run our regressions on the non-truncated data instead. Table 1 displays the unweighted sample counts from the AHS, with the last column showing that our baseline sample includes about 35,000 households in total.

Figure 2 compares median levels of new mortgage debt and income from HMDA and the AHS.[12] The top panel shows that the time-series patterns of median mortgage debt line up well across the two datasets, especially when we subset on metropolitan areas, where HMDA reporting is concentrated. The lower panel, however, shows that income reported to HMDA rises relative to AHS income during the height of the housing boom. The income pattern thus confirms a lesson from Avery et al. (2012), who note that income reported to HMDA potentially overstates a borrower's true income in areas or periods where house prices are rising rapidly.

3 Quantifying Technical Progress in Mortgage Lending

3.1 Cross-State Mortgage Originations

In this section, we quantify three areas in which technological improvements in mortgage lending are reflected. The first is that since 1990, lenders have been increasingly able to handle loan applications from areas outside of their normal markets. Using HMDA data, we calculate the percentage of new loans in each year for which the lender has no physical presence in the state in which the mortgaged property is located. To make this calculation, the HMDA data are matched with data on each lender's parent company and/or bank holding company (which we refer to as the lender's topholder), using institutional information from the National Information Center (NIC) and bank-branch information from the FDIC's Summary of Deposits (SOD).[13] Mortgage companies and credit unions are excluded from this

[11] Early in the sample, the allocated values are the topcode limits themselves; later in the sample, the allocated values are the mean values of income or debt above the respective topcode limits.

[12] Because medians are not affected by the truncation rule, the AHS series are generated using the non-truncated sample of homeowners who recently moved.

[13] The NIC is a repository of financial data and institutional characteristics collected by the Federal Reserve System (http://www.ffiec.gov/nicpubweb/nicweb/SearchForm.aspx). Institutions are identified in the NIC by tags known as RSSD IDs. To match lenders to their parent companies, bank holding companies and branches, we use a dataset kindly supplied by Robert Avery that matches HMDA lender identifiers with

analysis.[14] We also do not have information on the locations of loan production offices (LPs), which process loans but do not take deposits and therefore are not considered branches. However, LPs cannot originate loans, only process them. Loans are approved and funds are dispersed by other branches or offices in the bank. For LP offices to exist any significant distance from a branch of office that is able to approve applications requires a certain level of technology, such as fax machines or internet access, to transfer documents between locations with sufficient speed.

The results are displayed in Figure 3. The two panels depict the percentage of loans originated out-of-state from the perspective of the individual lender (left panel) and the lender's topholder (right panel). By construction, the percentage of loans originated out-of-state must be lower when considering all brick-and-mortar locations from the perspective of the topholder. Both panels show an increase in out-of-state lending during the 1990s. In 1990 fewer than 20 percent of purchase loans in the HMDA records were made on properties located in different states compared to the lender's location. By 2000 this fraction had increased to more than 60 percent at the lender level, and more than 55 percent from the perspective of the topholder. The percentage of out-of-state loans has remained high ever since.

3.2 Mortgage-Processing Timelines

The 1990s also saw a significant decline in the time that mortgage originators required to process loans. Using the confidential version of HMDA, we can calculate the time to process the loan as the number of business days between the application date and the date the loans are either denied or originated. A number of factors determine loan-closing timelines, including the volume of applications processed by the lender, the lender's size, and whether the borrower is applying alone or with a co-applicant. To account for as many of these determinants as possible, we run loan-level regressions that project the time required to process a loan on the assets of the lender (in logs), the type of the lender (credit union, thrift, mortgage company, and so on), the race and gender of the borrower, whether the

RSSD IDs back to 1990. This dataset also provides information for the lender's parent company or bank holding company. HMDA provides RSSD IDs for lenders and their topholders starting in 2004. The file provided by by Robert Avery includes information back to 1990 and includes some additional checks on the information available publically in HMDA.

[14] Credit unions are not included in the SOD. The locations of credit union branches are available from the National Credit Union Association only from 2010 onward. Mortgage companies are firms that do not have branch locations but often fund loans using lines of credit obtained from commercial banks. Unfortunately, given our data it is not always feasible to see connections between mortgage companies and commercial banks. For example, a mortgage company may be owned by a bank holding company, but this ownership would not be taken into account when considering the locations where the bank holding company has a physical presence.

borrower has a co-applicant, and a concurrent measure of mortgage application volume.[15]

Figure 4 depicts the results of these regressions. There is a dramatic decline in the average processing time for refinances between 1995 and 1998, which then continues to drift lower until 2005. The timeline increases after 2007, but on average it remains about 10 business days faster today than before 1995. There is no such pattern for purchase loans. This is not surprising, because closing dates for purchase loans are often chosen to accommodate the borrower and seller as they move to new residences. Consequently, the time between a purchase-loan application and the closing date can be lengthy, even if the borrower has been pre-approved and has provided much of the necessary documentation before making an offer on the house.

3.3 Declining Cost of Intermediation

As lenders began to compete for loans across state lines and were able to close loans more quickly, the costs of financial intermediation borne by borrowers declined. Such declines have been noted in some previous research, which references the steep decline in initial points and fees paid by borrowers. As seen in the top panel of Figure 5, this series, which is now maintained by the Federal Housing Finance Agency (FHFA), declined from just under 3.0 percent in the early 1980s to about 0.5 percent in the mid-2000s. Although this decline is partly determined by the reduced costs of mortgage intermediation, the decline is also influenced by other factors, such as the fraction of borrowers who choose to "buy down" their interest rate by paying higher points when their loans are originated.

The lower panels of Figure 5 plots a measure of mortgage-originator profits and unmeasured costs (OPUCs) that accounts for many of these other factors. The details of its calculation is described in Fuster et al. (2013), who measure the portion of the revenues earned by originating a loan that is retained by the originator.[16] It is important to keep in mind that this measure is comprised of both costs and profits. A decline could reflect a fall in costs, while an increase could reflect higher profits.[17] Despite this caveat, Fuster et al. (2013) provide a much cleaner measure of the intermediation costs borne by borrowers than the points-and-fees series shown in the top panel.

[15] Our measure of lending volume is the average number of mortgage applications per business day in each month using the HMDA data. To account for seasonality, we only consider loans originated in the second and third quarter of each year; other methods of accounting for seasonality give similar results.

[16] The calculation assumes that the loan is a 30-year fixed rate mortgage sold into a Fannie Mae or Freddie Mac mortgage-backed security (MBS). This value is the sum of profits from origination and the present value of servicing rights. The profits from origination include revenues earned from closing costs plus any margin the originator makes from selling the loan on the secondary mortgage market, less guarantee fees and loan-level price adjustments paid to the GSEs.

[17] For example, during the refinancing boom of 2000–2003, intermediation costs rose, but loan-processing timelines stayed the same. This is consistent with a story in which to keep processing times low, people were willing to pay more for intermediation services.

This bottom left panel shows that, consistent with the growth of cross-state lending and declines in processing timelines, OPUCs declined sharply during the 1990s. Costs then fluctuated during the housing boom and have increased dramatically in recent years. Because intermediate costs are influenced by capacity constraints, the bottom-right panel controls for application volume using the same technique as was used in the analysis of mortgage timelines; specifically, a regression that includes the number of applications per business day per month. The results change very little.

3.4 Reversal of Trends after the Great Recession

An interesting facet of both the decline in refinance timelines and the drop in intermediation costs during the 1990s is that both movements have at least partially reversed. In fact, the measure of intermediation costs is higher today than it was two decades ago. In this subsection, we discuss some new policies and regulatory changes that help explain the timeline and cost reversals, even though the underlying technology of mortgage origination permanently changed in the 1990s. Then, in section 4, we study how technology changed the relationship between income and debt during the 1990s—a change that was more stable over time.

One of the most significant factors to affect mortgage originators after the housing boom of the 2000s was a change in the repurchase policies of the large government-sponsored enterprises (GSEs) that securitized loans, Fannie Mae and Freddie Mac. These agencies occasionally require mortgage originators to repurchase loans that do not meet the agencies' guidelines in some way. The top-left panel of Figure 6 depicts the number of loans repurchased by Fannie Mae and Freddie Mac, taken from a sample of single-family agency loans analyzed by Goodman, Parrott, and Zhu (2015). Not surprisingly, the agencies required originators to repurchase a relatively large number of loans that were originated near the peak of the housing boom, because this cohort of loans had a large number of defaults. More recently, a loan has not had to become distressed before being "put back" to the originator. The top-right panel depicts the fraction of repurchased loans that had always been current at the time they were repurchased. Most of the problematic boom-era loans were distressed before being put back, but in recent years, few repurchased loans had ever missed a payment. The implication is that after the housing boom, originators selling loans to the GSEs had to be much more careful on each loan that they processed. Failing to follow an agency's rules to the letter could result in a costly repurchase, and a loan's perfect repayment history no longer protected the originator from that penalty.

The lower panels of Figure 6 illustrate the consequences of another change in the post-crisis lending landscape: new disclosure rules for mortgage originators. The Dodd-Frank Act of 2010 instructed the new Consumer Finance Protection Bureau to propose rules that would combine and integrate mortgage disclosures under the Truth in Lending Act (TILA) and the

Real Estate Settlement Procedures Act (RESPA). The final rule, called the TILA-RESPA Integrated Disclosure (TRID) rule, became effective in 2013. One goal of the new rule was to give borrowers more information about mortgage offers at the start of the origination process, so that they could better shop around for the best deal. Additional disclosures near the end of the process were intended to ensure that borrowers were not surprised at the loan closing about any features of the mortgages they ultimately chose.

As illustrated by the lower left panel of Figure 6, lenders report that adapting to the TILA-RESPA rule has entailed significant costs. The panel shows the results of a survey conducted by Federal Reserve System and the Conference of State Bank Supervisors that asked community bankers about the impact of various banking regulations, including regulations unrelated to mortgage lending. The bankers reported that the new TILA-RESPA rule was the most confusing regulation they faced, that it accounted for the highest share of compliance costs, and that it was the regulation they would most like to change. The lower right panel illustrates the consequences of TILA-RESPA changes on various aspects of mortgage origination. Among the most important consequences were a slowdown in the pace of origination, delayed closings, and increased staffing costs.

The new disclosure regulations may well be a net plus for the housing market, because they provide potential borrowers with important information about their loans. Additionally, any effects of the new regulations or GSE practices on origination costs and timelines may fade as lenders adapt to them and the rules themselves evolve. Indeed, as discussed in Goodman (2017), the GSEs began instituting new policies for loan repurchases in September 2012; these policies include time limits on repurchase requests for newly originated loans and additional guidance on the types of defects that might prompt a repurchase.[18] But the post-crisis rules do explain why timelines and costs have risen recently, even though mortgage origination remains a highly automated process. The impact of automation on the debt-income relationship has remained more stable, as we show in the next section.

4 Income and Mortgage Debt in the 1990s and 2000s

4.1 The Canonical Debt-Income Regression

After the housing boom ended, the relationship between mortgage debt and income became the focus of a large body of empirical research, most of which is based on some variant of what can be called the canonical specification for the debt-income relationship:

$$D_{it} = \alpha_t + \beta_t I_{it} + \epsilon_{it}, \tag{1}$$

[18]See also the *Housing Wire* story titled "Fannie Mae, Freddie Mac Announce New Mortgage Buyback Rules," by Ben Lane, Oct. 7, 2015.

where D_{it} is the log of the value of a new mortgage originated for individual i in year t, I_{it} is log income, and the coefficients α and β have subscripts because they can change over time. Empirical estimates of β, the partial correlation between new debt and income, are positive because more-affluent people tend to live in more expensive houses and take out larger mortgages. Additionally, low-income borrowers might face limits on the amount of mortgage debt they can take on, via ceilings on permissible DTI ratios. Relaxing these limits allows low-income households to borrow more, which increases the amount of debt at the bottom of the income distribution and causes a decline in the positive cross-sectional relationship between debt and income that is summarized by β.

Figure 7 motivates the regression analysis with binned scatter plots of log debt and income using individual-level data from HMDA and the AHS. In both cases, the cross-sectional relationship between income and debt is nearly linear in logs. The plots also show that for both data sources, the positive relationship between debt and income flattens from the beginning to the end of the sample periods. By running the canonical regression and examining the yearly βs that result, we can determine precisely when this decline takes place.

4.2 Regression Results

We first use the canonical specification to study individual-level debt and income from HMDA. Some limited demographic variables are available in that dataset, and we are also able to include CBSA-year fixed effects to control for local housing market characteristics, including area-wide fluctuations in housing demand.[19] The CBSA-year fixed effects also correct for any over-reporting of income that is consistent across a local market in a given year. We modify Equation 1 to

$$D_{it} = \alpha_{ct} + \beta_t I_{it} + \gamma X_{it} + \epsilon_{it}, \tag{2}$$

where the vector X_{it} includes dummy variables denoting the borrower's race and gender, while c indexes the CBSA. Note that α_{ct} is indexed by CBSA as well as by time, so this term now represents CBSA-year fixed effects.[20] Because HMDA income can potentially be overstated, as described in Section 2, we instrument for individual-level income reported to HMDA with median tract-level household income from the decennial census and ACS.

The top panel of Figure 8 displays the β_ts from this regression, which trend downward from 1990 through the early 2000s. This trend is consistent with a steady decline in the

[19]CBSAs, or core-based statistical areas, consist of at least one urbanized core with a population of 10,000 or more, along with adjacent counties that have a high level of social and economic integration. CBSAs have been the main way that the government classifies metropolitan areas since 2003.

[20]In this regression, we will cluster the standard errors by CBSA, but the results remain robust when clustering by state.

partial correlation of income and debt until the housing boom begins, at which point the β_t's rise. The bottom panel displays expected mortgage amounts, which are calculated from a regression that includes only yearly intercepts α_t rather than CBSA-fixed effects α_{ct}. The estimates of these intercepts are added to the yearly estimates of $\beta_t \times \overline{I}_{.t}$, where $\overline{I_{.t}}$ denotes the mean of income in the entire sample. These expected amounts increase sharply after 2000, when the well-documented increase in aggregate US mortgage debt to aggregate US income takes place.[21]

Figure 9 presents analogous results using the AHS. Recall that in this dataset, there is no need to instrument for income, which comes from survey data, not loan applications. Unfortunately, small sample sizes in the AHS mean that we cannot run regressions with CBSA-year fixed effects.[22] The top panel mimics the HMDA results in that the debt-income relationship summarized in β_t declines throughout the 1990s and flattens out in the 2000s. The bottom panel shows the estimates of expected mortgage amounts. As with the HMDA data, these amounts increase rapidly during the boom.

Both the HMDA and AHS regressions confirm a central finding in Adelino, Schoar, and Severino (2016): the debt-income relationship remained stable during the 2000s housing boom. Because we have treated second liens consistently over the sample period, and because we have addressed the possibility that individual-level HMDA income can be overstated, our findings are immune to some criticisms of Adelino, Schoar, and Severino (2016) advanced by Mian and Sufi (2017). More important for this paper, however, is the decline in the debt-income coefficients that we do find during the 1990s. Were technological advances responsible for that decline, and if so, what effects did those advances have on the housing market in the 2000s? We take up those questions in the next section.

5 Credit Scores, Automated Underwriting and the Debt-Income Relationship

A potential explanation for the 1990s decline in the debt-income relationship is the downward trend in nominal interest rates, but interest rates are unlikely to explain much of what we have found. To be sure, the decline in nominal interest rates has been significant; the top-left panel of Figure 10 shows the mean and median of nominal interest rates of AHS movers, along with the conventional 30-year rate for US mortgages published by Freddie Mac. The top-right panel shows that dispersion in interest rates declines from 1985 on, consistent with

[21] Figure A.2 in the appendix contains results from parallel regressions not including CBSA-year fixed effects and without using census tract-level median household income as an instrument.

[22] We can, however, include yearly fixed effects for the country's four census regions: Northeast, Midwest, South and West. These results along with additional robustness checks are included in Figure A.3 in the appendix.

the increased geographic competition among lenders discussed earlier.[23] But the lower left panel shows that the β_ts from our canonical regression display the same pattern even after controlling for individual-level interest rates, by allowing for yearly interest-rate effects as well as yearly income effects (the interest-rate coefficients themselves appear in the lower right panel).

Government policy has also been suggested as a reason for looser lending standards during the 1990s (Wallison and Pinto 2012; Morgenson and Rosner 2011; Rajan 2010). This literature often cites the Clinton Administration's National Homeownership Strategy of 1995, a policy initiative that encouraged housing-market participants from the private and public sectors to raise the number of US homeowners by 8 million by 2000. The varied nature of the 100 action items in this strategy make it difficult to assess this effort's direct effects. These items ranged from building-code reform (item 8), home mortgage loan-to-value flexibility (item 35), subsidies to reduce down payment and mortgage costs (item 36), flexible mortgage underwriting criteria (item 44), and education on alternative forms of homeownership (item 88).

Other regulations designed to increase homeownership included the Community Reinvestment Act (CRA), which was passed in 1977 but strengthened in the 1990s, and a 1992 act that encouraged the GSEs to increase credit availability in low-income or underserved areas. Part of the GSE's efforts to expand lending to underserved borrowers were conducted through affordability programs, such as Freddie Mac's Affordable Gold program, which began in 1992. Among other things, the Affordable Gold program relaxed front-end and back-end DTI limits to 33 and 38 percent, respectively, and also allowed smaller down payments.[24] Government policies that rewarded lenders for making loans to underserved areas have been subject to a number of empirical tests, but regression-discontinuity studies fail to show that either the CRA or the 1992 GSE act had much of an effect (Bhutta 2011, 2010; Avery and Brevoort 2015). In any case, the fact that the debt-income relationship has remained stable throughout the 2000s and 2010s, even as government policies have changed, suggests that these policies do not explain our results.

[23] There is a long literature on the integration of the mortgage market with national capital markets, which limits the dispersion in mortgage interest rates paid by US households. See, for example, Rudolph and Griffith (1997) and Hurst et al. (2016).

[24] In some cases, the back-end ratio could rise to 42 percent. The normal limits for these ratios was 28 and 36 percent. The program also allowed borrowers to contribute less than the full 5 percent down payment from their own funds, and required participants to take a financial counseling course before purchasing a home. Fannie Mae had a similar program, the Community Homebuyer Program, which began in 1990. See U.S. Department of Housing and Urban Development (1996, p. 90) for details.

5.1 Credit Scoring and Mortgage Lending in Theory and Practice

In contrast to interest rates and government policies, technological improvement stands out as a better explanation of our empirical findings. As noted in the introduction, computerization in the mortgage-lending industry was initially expected to take the form of AI algorithms that would replicate mortgage underwriters' decisions about the risk of default. But the intuitive decision rules proved difficult to code into formal algorithms, and the development of AI systems in the early 1990s eventually turned out to be a dead end. As one industry professional wrote, these algorithms "gave speed and consistency to the underwriting process, [but] by 1995 their accuracy was seriously questioned. These systems were built to reproduce manual underwriting without much consideration of whether the manual-underwriting thought process was optimal" (Makuch 1998, p. 4).

At the opposite end of the spectrum was a method that would prove much more accurate: numerical credit-scoring algorithms. An example of such an algorithm is a linear default regression that projects binary default indicators on variables from borrowers' credit histories and the risk characteristics of their loans. A more complicated algorithm might be estimated via machine learning and then predict default risk using nonlinear or hierarchical relationships among these variables. In either case, the resulting default model could then be used to construct predicted probabilities of default, which could then inform a lender's decisions regarding whether to make a loan and the interest rate that should be charged. Credit-scoring algorithms have been used in one form or another at least since the 1970s; the first modern version of the FICO score appeared in 1989. Previous researchers have shown that the distillation of high-dimensional information from credit records into a single score had a substantial impact on many types of consumer lending, particularly credit cards (Evans and Schmalensee 2005).

Yet even as credit scoring transformed consumer lending in the 1990s, many lenders doubted that credit scores could help them predict mortgage default. Unlike consumer debt, mortgages are secured loans, and the borrower's equity stake in the property plays a critical role in the default decision. In fact, the central project for academic economists studying default in the 1980s and 1990s was building the so-called frictionless option model (FOM), in which borrower-specific variables, including credit scores, play no role in the default decision. In the FOM, the borrower's default decision is fully characterized by the level of negative equity at which the borrower should "ruthlessly" or "strategically" default. This threshold equity level is a complex and time-varying function of borrower equity, house prices, and interest rates. Individual-level adverse life events such as job loss and illness do not lead to default in the FOM, because the model assumes that borrowers can ride out these problems with unsecured loans at the risk-free rate.

In reality many borrowers are liquidity constrained, so adverse shocks can lead to default when borrowers are underwater on their mortgages. The mortgage-default literature is now attempting to blend insights from the FOM with those of the "double-trigger" model, which links default to the simultaneous occurrence of negative equity and an adverse life event.[25] In these models, borrowers who suffer adverse life events often lack the liquid funds needed to remain current on their mortgages and are unable to take take out unsecured loans to tide themselves over. Those borrowers who also have negative equity are unable to sell their homes for enough to discharge their loans, so default occurs.

Another relevant finding from the mortgage-default literature concerns those negative-equity borrowers who do not suffer adverse life events. For these unconstrained borrowers, most calibrations of the FOM generate optimal default triggers in the neighborhood of 10–25 percent negative equity (Kau, Keenan, and Kim 1994; Bhutta, Dokko, and Shan 2017). In empirical data, however, negative equity typically exceeds this level without the borrower defaulting. This result is relevant for mortgage-default modeling because it suggests that empirical models do not have to predict the relevant negative-equity default thresholds in a way that is consistent with the FOM.

The overall characterization of the default decision that emerges from this research suggests that initial LTV ratios and credit scores should be included in any empirical default model. Low initial LTV ratios (that is, high down payments) reduce the probability of future negative equity, so they reduce the probability of double-trigger and strategic defaults. Borrowers with high credit scores should also be less likely to experience double-trigger defaults if these scores reflect low probabilities of experiencing a liquidity shock, either because the borrower has a stable job, or because he has ample liquid wealth.[26]

Although income is critical in the new generation of default models, the role of future income shocks as opposed to initial income levels suggests that DTI ratios at origination should not affect default very much. In the double-trigger model, default occurs when an income shock raises the borrower's current DTI to very high levels. Origination DTIs should matter only to the extent that low DTIs make it less likely that a borrower will experience a shock large enough to trigger default. The variance of income shocks at the individual level is so high that setting a low DTI at origination may not buy the lender much insurance in that regard (Foote et al. 2010). After all, in the case of a job loss that halts income completely, the borrower's DTI rises to an infinite level no matter what the origination DTI.

[25] Examples of this work include Corradin (2014), Campbell and Cocco (2015), Laufer (2018), and Schelkle (2018). For a survey of recent research in this area, see Foote and Willen (2018).

[26] High credit scores may also reflect high "stigma" costs to the borrower for any type of default, which also supports their inclusion in mortgage-default regressions.

5.2 Integrating Credit Scores into Mortgage Lending

Empirical models estimated with individual mortgages confirm these theoretical predictions. In a very early study sponsored by the NBER, Herzog and Earley (1970) used data from 13,000 individual loans and found that initial LTV ratios were good default predictors, and in the following decades subsequent studies confirmed this early result. Later researchers found that the new credit scores developed in the 1990s also entered mortgage-default regressions significantly. Mahoney and Zorn (1997) discuss Freddie Mac's modeling work in the early-to-mid 1990s, noting that borrowers with FICO scores less than 620 were found to be more than 18 times more likely to experience foreclosure than borrowers with scores greater than 660. Yet DTI ratios proved to be much weaker default predictors. An influential Federal Reserve study (Avery et al. 1996) summarized the existing consensus in both industry and academia by noting that, "[p]erhaps surprisingly, after controlling for other factors, the initial ratio of debt payment to income has been found to be, at best, only weakly related to the likelihood of default" (p. 624). The Fed study also presented original research showing that credit scores were good predictors of default and could be used to improve lending decisions: "[t]he data consistently show that credit scores are useful in gauging the relative levels of risk posed by both prospective mortgage borrowers and those with existing mortgages" (p. 647).

Armed with this information, the GSEs began encouraging loan originators to use credit scores in their lending decisions. In July 1995, Freddie Mac's executive vice president for risk management, Michael K. Stamper, sent a letter to Freddie's sellers and servicers encouraging them to use credit-score cutoffs when underwriting loans. Loans with FICO scores over 660 should be underwritten with a "basic" review, while borrowers with scores between 620 and 660 should get a more "comprehensive" review. For borrowers with FICO scores below 620, underwriters should be "cautious," in that they should:

> [P]erform a particularly detailed review of all aspects of the borrower's credit history to ensure that you have satisfactorily established the borrowers' willingness to repay as agreed. Unless there are extenuating circumstances, a credit score in this range should be viewed as a strong indicator that the borrower does not show sufficient willingness to repay as agreed. (Stamper 1995, p. 2)

The letter also explicitly permitted lenders to use high credit scores to offset high DTI ratios: "A FICO bureau score of 720 or higher... will generally imply a good-to-excellent credit reputation. If your underwriter confirms that the borrower's credit reputation is indeed excellent, then it could be used a compensating factor for debt-to-income ratios that are somewhat higher than our traditional guidelines ..." (Stamper 1995, p. 13). Within a few months, Fannie Mae followed suit by encouraging lenders to use identical credit-score cutoffs.[27]

In addition to supporting the general use of credit scores, the empirical default models were used to develop numerical scorecards that could weigh all the data in a loan application. One scorecard produced by Freddie Mac was the Gold Measure Worksheet, which was designed to assist underwriters in evaluating applications for the Affordable Gold program. This worksheet, which also appeared in Avery et al. (1996) and which we reprint as Figure 11, allocates risk units to a loan applications based on the borrower's LTV ratio, DTI ratio, credit score, and other credit information.

The implicit weights in the worksheet reflect the lessons of empirical default models by assigning high importance to equity and credit scores and low importance to origination DTIs. The table on the following page illustrates this fact by using the worksheet to evaluate three hypothetical loans. Loan A has an LTV of 90 percent, a FICO score of more than 790, and a DTI ratio exceeding 50.5 percent.[28] Although the DTI ratio is very high, the high FICO score offsets this penalty enough to reduce to the total risk-unit score to 14, one unit below Freddie's cutoff. The application for Loan B is generated by a hypothetical borrower with a low credit score and carries a much smaller DTI ratio. This loan turns out to be too risky, as it scores one risk unit above Freddie's cutoff. Finally, Loan C has a very high LTV and DTI ratios (99.5 percent and 50.5 percent, respectively), as well as an adjustable interest rate. But the borrower also has a credit score of more than 790 and five months of liquid financial reserves. The latter factors are enough to offset the high DTI and LTV ratios so that the loan falls below the risk-unit cutoff.

Although the Gold Measure Worksheet fit on a single page and took only minutes to complete, it proved far more accurate in predicting default than human underwriters, who also took much longer to evaluate each loan file. The superior speed and accuracy of scorecard-based evaluations were illustrated powerfully in a head-to-head comparison between human underwriters and the computer-based scorecard that is described in Straka (2000). Sometime after 1994 Freddie Mac purchased about 1,000 loans from a major lender through the Affordable Gold program. After Freddie Mac received these loans, the agency's quality-control investigators indicated that the loans' risk characteristics fell outside Freddie Mac's guidelines, so a sample of 700 loans was scored using the Gold Measure Worksheet. This exercise, which took only a few hours, indicated that only about half the loans were of "investment quality" and thus eligible for purchase by Freddie Mac. At that point, human underwriters then re-evaluated all 1,000 mortgages, a process that took six months. The human underwriters also found that about half of the loans were good enough to be purchased by Freddie Mac. Yet while there was some overlap, the humans and the worksheet differed substantially

[27] See Poon (2009, p. 663) and Dallas (1996) for details of Fannie Mae's instructions.

[28] This DTI ratio is the back-end ratio, so it encompasses not only the borrower's mortgage payment but also car loans and other regular installment payments.

	LTV Ratio	FICO Score	DTI Ratio	Months of Reserves	Fixed or Adjustable Rate	Total Risk Units
Loan A	90%	Over 790	Over 50.5%	2–3	Fixed	
Risk Unit Increment	0	–16	+30	0	0	14
Loan B	90%	640	Below 32.6%	2–3	Fixed	
Risk Unit Increment	0	+17	0	0	0	17
Loan C	99.5%	Over 790	50.5%	5	Adj.	
Risk Unit Increment	+13	–16	+18	–6	+6	15
Freddie Mac Guideline for Loan Acceptance						**15**

Evaluating Alternative Loans Using the Gold Measure Worksheet
Note: These calculations correspond to single-family, 30-year mortgages for which no special cases apply (for example, the borrower is not self-employed). The adjustable-rate mortgage in Loan C is a rate-capped ARM (not a payment-capped ARM).
Source: Authors' calculations using the Gold Measure Worksheet in Figure 11.

on the set of mortgages that met Freddie Mac's standards.

By following the mortgages over time, Freddie Mac could conduct a horse race between the worksheet and human underwriters regarding their respective abilities to predict mortgage default. As Straka writes, "the race was not very close." During the first three years after origination—a period when underwriting differences tend to have the strongest effect on default—the worksheet ratings proved to be powerful predictors of future distress. The foreclosure rate on loans placed in the worksheet's noninvestment category was almost three times higher than the rate for its investment category, and the 30-day delinquency rate was nine times higher. But the two categories as determined by the human underwriters performed almost exactly the same, despite their extra cost: "[r]eview underwriting ratings that took six months to complete performed not much better (if better) than flipping coins" (Straka 2000, p. 219).

5.3 Automated Underwriting

The next step was for the GSEs to leverage their central place in the mortgage industry (and their substantial financial resources) by developing software that incorporated their default-prediction scorecards and could be distributed directly to loan originators.[29] By

[29]Freddie Mac was generally ahead of Fannie Mae in this effort. In fact, Fannie Mae started out as a leader in developing computerized AI algorithms to mimic human decisions. But Fannie Mae officials eventually realized that it was critical to incorporate lessons from empirical default models into any computerized

1995, both GSEs had developed AU systems: Loan Prospector at Freddie Mac and Desktop Underwriter at Fannie Mae. These proprietary software packages allowed loan originators to enter borrower and loan characteristics into a desktop computer, which would then report how the GSEs would treat the loan. For example, an "accept" rating from Loan Prospector meant that Freddie Mac would purchase the loan without additional analysis. Ratings of "caution" or "refer" required the originator to perform additional screening before submitting the mortgage for purchase to the GSE.

As the GSEs gained more confidence with these AU systems' abilities to evaluate risk, they began to expand the credit box. Evidence on this point comes from data in Gates, Perry, and Zorn (2002), which we use to construct Table 2. The table reports the results of two additional horse races that also use the set of Affordable Gold mortgages referenced above. The two races pit the human underwriters against the 1995 and 2000 calibrations of Freddie Mac's Loan Prospector AU system. The bracketed numbers in the table report the 90-day delinquency rate for each group, relative to the delinquency rate for the entire sample; a rate of 1.00 indicates that the group defaulted at the same rate as the entire sample. The non-bracketed numbers table refer to group shares, as percentages of the entire sample of evaluated loans.

The first column reports the results from groups as classified by the human underwriters. As noted above, the underwriters took six months to designate about half of the loans as acceptable for purchase by Freddie Mac, although this half ultimately performed about the same as did the noninvestment-quality half.[30] The next two columns show how Freddie Mac's 1995 model evaluated the sample. The bottom row of these columns indicate that this model was slightly more conservative than the human underwriters, with only 44.8 percent of the mortgages labeled as acceptable by the 1995 AU model. But the 1995 model accepted many of the mortgages that the human underwriters rejected (a group comprising 20.8 percent of the sample), while it rejected many mortgages that the human underwriters accepted (27.5 percent). And the model appeared to pick the right mortgages on which to disagree, as its accepted mortgages defaulted at only about one-fifth the rate of the sample as a whole.

Gates, Perry, and Zorn (2002) write that over time, "Freddie Mac rapidly expanded accept rates as the tool became more accurate and the company gained experience with and confidence in the new technology" (p. 380). This expansion is shown in the last two columns, which depict accept rates and relative performance according to the 2000 version of Loan Prospector. The model now accepts more than 87 percent of the sample, but the relative delinquency rate of this group is still better than the 51.6 percent accepted by the humans (0.70 vs. 1.04).

underwriting system. See McDonald et al. (1997) for details.

[30]Indeed, the "accept" group had a relative default rate of 1.04, a bit higher than the default rate for rejected half (0.96).

How much of this credit-box expansion involved an increase in permissible DTIs? Once lending policies have been encoded into a proprietary automated underwriting system, we can no longer evaluate them with comparisons of hypothetical loans, as we did for the Gold Measure Worksheet. But relaxed DTI standards were no doubt an important part of the credit expansion. As the use of AU systems grew in the late 1990s, some borrowers and lenders became frustrated by their black-box nature, so the GSEs provided limited information about their underlying scorecards in mid-2000. Freddie Mac, for example, reported that DTI ratios (both front-end and back-end) were one of 14 factors that its algorithm considered. But this ratio was not one of the three most important factors, which were the borrower's total equity, loan type, and credit scores. As for Fannie Mae, a well-known syndicated real estate journalist wrote in mid-2000 that the most critical component of the Desktop Underwriter system was the credit score, and the last of ten factors listed was the "payment shock." This shock was not the the DTI ratio itself, but the *difference* between the new monthly mortgage payment and the amount that the homeowner was already paying for housing. And mortgage brokers did not view even this weaker income test as a "major tripper-upper."[31]

More recently, the GSEs evaluation of DTI has taken on additional importance after the passage of the Dodd-Frank Act, which imposes ability-to-pay rules for mortgages. In general, these requirements are automatically met for loans acceptable for purchase by the GSEs, and the GSEs often allow higher DTIs for mortgages underwritten through their AU systems. For exmaple, in its selling guide published on October 2, 2018, Fannie Mae reported that for manually underwritten loans, the maximum permissible back-end DTI ratio (which includes nonhousing debt) is 36 percent of the borrower's "stable monthly income" in most cases. Manual underwriters may permit DTIs up to 45 percent, however, as long as the borrower meets certain requirements related to her credit score and available liquid reserves. Yet for loans underwritten through Desktop Underwriter, Fannie's AU system, the maximum allowable DTI ratio is 50 percent.[32]

All told, the historical record indicates that DTI ratios at origination became less important in lending decisions during the 1990s, consistent with modern theories of mortgage default as well as the debt-income analysis presented in section 4. Data also show that downplaying the importance of human judgment in default decisions improved the evaluation of credit risk, which explains why numerical methods continued to be developed today, via machine learning and other methods. To the extent that the information used in these methods is racially neutral, the models also deliver racially unbiased lending decisions, something that

[31]See "Building Blocks of Automated Underwriting," by Lew Sichelman, United Feature Syndicate, June 4, 2000.

[32]See Fannie Mae Selling Guide, Part B3-6-02, "Debt-to-Income Ratios." Available at https://www.fanniemae.com/content/guide/selling/b3/6/02.html, accessed October 17, 2018.

could not be taken for granted at the start of the 1990s (Munnell et al. 1996). The decreased cost, increased accuracy, and unbiased nature of AU systems help explain why they were embraced so quickly by mortgage lenders.

A deeper point is that the computerization of mortgage lending during the 1990s has close parallels with technological change in other industries. When the decade began, it was not clear exactly how computerization could enhance mortgage lending. Initially, attention focused on computerized AI systems designed to replicate human decisions, but lenders eventually realized that information technology could bring about a deeper transformation of their industry by reducing the importance of human decisionmaking. There is an obvious parallel here with the introduction of electricity into manufacturing in the early twentieth century (David 1990). US manufacturers took some time to realize that electricity could do more than simply replace the central steam-power sources in their multistory factories. Because sources of electricity could be distributed more easily throughout factories than steam power could, electricity allowed manufacturers to essentially turn their factories on their sides by constructing new single-story factories, through which materials could move more easily. As David (1990) and others have argued, in the modern era computerization has also led to delayed productivity gains, since it takes businesses some time to figure out the fundamental changes that computerization allows. The transformation of mortgage lending during the 1990s provides a good example of this phenomenon.

6 Consequences of the 1990s Credit Expansion

6.1 Effects on Homeownership Rates

So far, this paper has argued that changes in mortgage lending during the 1990s were determined by technology and therefore these changes were exogenous with respect to the cyclical state of the housing market. A valid question is whether and how these changes contributed to the 2000s housing boom. In this section, we discuss the potential impacts of the 1990s developments on the homeownership rate and on housing prices.

According to the Census Bureau, the US homeownership rate rose from 64.0 percent in 1994 to 69.0 percent in 2004. To the extent that this increase was driven at least in part by changes in current-income requirements for mortgages, we would expect homeownership increases to be especially large for households with higher future or permanent incomes relative to their current incomes. The official homeownership rate is generated by the Current Population Survey/Housing Vacancy Survey (CPS/HVS), so it is straightforward to test this prediction by disaggregating the CPS/HVS microdata by age and education. Young college graduates are known to have steeper age-earnings profiles than individuals of the same age without college degrees, and Figure 12 shows that the post-1994 increase in the homeownership rate was particularly large for younger persons with at least some years of college (that is, for college graduates and for persons with some college attendance but no degrees). The middle panel of the top row shows that among households headed by a 25-29 year-old, homeownership rates rose sharply among the more educated households, but barely moved for less-educated households. Qualitatively similar but less pronounced patterns are evident for older households as well.[33] For all but the youngest age groups, homeownership rates for the college-educated are substantially higher than for the less-educated. Putting all the pieces together, we see that the credit expansion of the 1990s essentially allowed younger and better-educated households to purchase houses sooner than otherwise would have otherwise been the case. These households had expectations of relatively high permanent incomes (because of their educational attainment) but low current incomes (because of their ages).[34]

Although the 1994–2004 increase in homeownership rates figures prominently in many narratives of the US housing crisis, this change was small relative to the ownership increase that occurred after World War II, when the introduction of zero-down-payment Veterans Administration (VA) loans and relaxed lending standards for mortgages insured by the Federal Housing Administration (FHA) helped increase homeownership by nearly 20 percentage points.[35] Yet like the recent homeownership increase, the more significant changes in housing finance during the middle of the twentieth century also allowed persons to buy homes sooner than they otherwise would have, as Fetter (2013) has argued. Figure 13 uses data from the decennial census and the American Community Survey (ACS) to provide a more detailed look at homeownership changes during various periods. The use of census and ACS data rather than the CPS/HVS allows us to disaggregate these changes in homeownership rates by four educational groups, rather than two, and also permits analysis by single-year-of-age rather than five-year age groups.

The top panel of Figure 13 depicts ownership changes by single-year-of-age and education over the 1940–1960 period. As Fetter (2013) points out, the main effect of underwriting changes immediately after World War II resulted in increased homeownership rates among younger adults who would have probably purchased homes later in life. Eligibility for the mid-century lending programs was broadly distributed across the population with respect to education, so it is not surprising that young persons in all educational classes, except high

[33]For household heads under 25 years of age, homeownership rates rise for both higher- and lower-educated households. But ownership rates for both groups remain small throughout the sample period.

[34]The better-educated groups would also be less likely to experience income disruptions, because the probability that a worker transitions from employment to unemployment is an inverse function of his education level (Mincer 1991; **?**).

[35]The homeownership rate rose from 43.9 percent in 1940 to 61.9 percent in 1960. See https://www.census.gov/hhes/www/housing/census/historic/owner.html for ownership rates based on decennial census data.

school dropouts, saw their homeownership rates increase. The lower panels depict changes in ownership between 1990 and 2000 (bottom left panel) and 2000 and 2005 (lower left panel) using the 2000 census and the 2005 ACS. Although the 1990–2000 panel shows little within-group increases, those in the 2000–2005 panel show that younger college graduates were most affected by the changes in mortgage-lending requirements. Homeownership rates for persons with no college education did not change during this period.

6.2 Effects on Housing Prices: Gross Flows of Residences

We now turn to the effect of developments during the 1990s on housing prices. The main idea behind much recent research on credit standards and house prices is that relaxed lending standards should raise the price of owner-occupied homes by raising the set of eligible home buyers. But as pointed out by Kaplan, Mitman, and Violante (2017) and others, the ultimate effect of lending standards on house prices depends on the ease with which residences can flow between rented and owner-occupied status. Ultimately, the price of a house should depend on the present discounted value of future rents. If changes in lending standards do not change the expected stream of rental income, then the price of owner-occupied homes should be unaffected as well. Paraphrasing Kaplan, Mitman, and Violante (2017), an implication of the link between house prices and future rents is that relaxed lending standards should merely encourage current renters to purchase their existing residences from their landlords.[36] As long as homes can flow between the rental and owner-occupied segments of the market, then higher effective demand for owner-occupied homes as opposed to rented ones can be satisfied through expanded supply, not by requiring significantly higher owner-occupied prices.

Housing experts have long recognized that the conversion of owner-occupied homes into rentals, known as "filtering," is a critical source of supply at the lower end of the housing market (Rosenthal 2014). The importance of filtering is one reason that the government and private housing groups expend considerable effort to measure flows of homes between various tenure arrangements. To get a sense of these flows, Table 3 depicts some results from various Components of Inventory Change (CINCH) reports, which are commissioned by the Department of Housing and Urban Development (HUD) to measure flows among homes that are owner-occupied, rented, or vacant. The table depicts CINCH data relating to losses of owner-occupied homes to either rented or vacant status, as well as losses resulting from destruction, conversions to nonresidential uses, and so on.[37] The underlying data source for CINCH is the AHS, which returns to the same houses every two years and asks respondents

[36] Because homes can flow in the other direction—that is, owner–occupied homes can be rented out—an increase in demand for owner-occupied homes can also be partially met by a *decline* in this gross flow.

[37] Other tables could be produced for gains of homes from one year to another, as well as for gains and losses of rental and vacant units.

whether they own or rent their residences. When no residents occupy a home, the AHS designates it as vacant.

The first row of Table 3 shows that, according to the AHS, about 56.8 million homes were owner-occupied in 1985. Of those units, about 52.2 million remained in existence and were still owner-occupied in 1987, so that total owner-occupied losses from 1985 to 1987 comprised 4.6 million units, or 8.1 percent, of the 1985 stock. The top row also shows that about 6.7 percent of this stock was either converted to rentals or became vacant by 1987, with the remainder lost due to conversion, destruction, etc. In 2009, during the housing bust, more than one in 10 owner-occupied homes were either rented or vacant two years later. But even in non-crisis periods, these flows are in the neighborhood of 8 percent.

Unfortunately, the published CINCH data do not allow us to determine how many owner-occupied homes become rental properties as opposed to becoming vacant from one survey to the next. Yet because the CINCH reports are based on AHS data, we can calculate this decomposition ourselves. Figure 14 depicts gross outflows of housing units that start out as either owner-occupied (top panel), rented (lower left panel) or vacant (lower right panel). The top panel shows that, among the owner-occupied properties that flow to another status, about half flow to rentals and half become vacant. The lower left panel shows the outflow data for rented units, where about two-thirds of rental outflows become vacant while one-third becomes owner-occupied. Yet the rented-to-owned flow is still relatively large. Specifically, about $1/35 \approx 2.9$ percent of renter-occupied homes flow to owner-occupied status every year.[38] More importantly, the upward trend in the red dashed line in the lower left panel indicates that the flow of rented-to-owned homes rose significantly in the 1990s, as we would expect with a modest credit expansion.

Table 4 depicts the annualized gross flows between rented and owner-occupied status, as well as the net flow of properties from rented to owner-occupied. From the 1985–1987 period through the 1991–1993 period the net flow in the last column of the table is negative. The negative sign is consistent with the filtering hypothesis. Because low-cost rental housing in the United States comes in large part from the steady depreciation of the owner-occupied stock, the flow of owner-occupied to rented homes tends to be larger than the flow in the opposite direction. In the late 1990s, however, the gross flow of rented to owner-occupied units rose significantly.[39] Not much happens to the gross flow in the opposite direction, so the net flow between rented and owner-occupied homes becomes positive, reaching 247,000 units per year in the 1997–1999 period.

[38]Because the AHS is conducted every two years, the survey-to-survey flow of about 2 million units in the table corresponds to about 1 million rental units becoming owner-occupied in any given year. Census data indicate that there were about 35 million renter-occupied units in the 1990s (a number that was relatively stable over the course of the decade).

[39]This increase is simply the annualized version of the red dashed line in the lower left panel of Figure 14.

Figure 15 studies all of the relevant margins for growth in the owner-occupied stock of homes. The left panel of this figure shows the growth of owner-occupied homes using data from the CPS/HVS.[40] Growth is calculated over two-year periods, so the chart implies yearly increases of about 1 million units. The right panel decomposes the sources of change in owner-occupied homes, with the contribution from rental properties depicted by the light blue bars. As we saw in Table 4, early in the sample period (and consistent with the filtering hypothesis) this contribution is negative, so the light blue bars lie below the horizontal axis through the mid-1990s. In the late 1990s, however, the rented-to-owned contribution becomes a significant positive contributor to growth in the owner-occupied stock, so the light blue bars lie above the horizontal axis. The contribution from vacant properties (dark blue bars) is generally negative throughout the sample, although it is a small negative contributor during the late 1990s and a large negative contributor during the 2000s. The construction contribution (red bars) is always the largest positive contributor to the growth of owner-occupied homes, proving especially large in the early 2000s.[41]

Two patterns apparent in the right panel are directly relevant for assessing the impact of relaxed lending standards over time. First, the contribution from rentals depicted by the light blue bars is never positive after 2001, as it was in the late 1990s. Second, the contribution from vacancies depicted is more negative in the 2000s than it was during the late 1990s. It would be difficult to replicate these two patterns with a model in which, first, lending standards were exogenously relaxed in the 2000s, and second, both rentals and vacancies were potential margins of adjustment. An exogenous loosening of lending standards that raised effective demand of owner-occupied homes would encourage the conversion of rentals into owner-occupied properties, and it would also draw more owner-occupied homes out of the vacancy pool. These patterns are apparent in the late 1990s, as the contribution from rentals becomes positive then, while the vacancy contribution becomes less negative. Yet these patterns reverse in the 2000s, the period when growth in US housing prices was strongest. Consequently, the two patterns cast doubt on the claim that price increases during the 2000s were driven in large part by an exogenous change in lending standards.

[40] The panel shows the growth in the number of owner-occupied homes after adjusting for a level break in the series in 2000, caused by the introduction of weights from the 2000 decennial census in that year.

[41] Conceptually, this contribution is the gross gains from newly constructed properties less the losses from destruction and non-residential conversions of owner-occupied homes that appear in the right-most columns of Table 3. Rather than build up this contribution using AHS microdata (as was done for rentals, vacancies, and seasonal properties) we calculate the net construction contribution by simply subtracting the sum of the latter three contributions from the overall change in the owner-occupied stock, which is depicted in the upper panel.

6.3 Effects on Housing Prices: House Prices at the Zip-Code Level

Another way to assess the effect of mortgage-lending standards on house prices is to study how the cross-section of house prices evolved at the zip-code level. As suggested by Mian and Sufi (2009), an exogenous relaxation of credit should bid up the prices of homes in low-income areas, where formerly credit-constrained individuals would be most likely to buy homes. Accordingly, we use zip-code level house price data from housing-data provider Zillow to estimate the following regression:

$$P_{zct} = \alpha_{ct} + \beta_t I_{zct}^{2000} + \epsilon_{zct}, \tag{3}$$

where z indexes a zip code in CBSA c in year t, P_{zct} is the log of Zillow's estimate of the average house price in the zip code, I_{zct}^{2000} is the log of median household income of that zip code from the 2000 decennial census, and α_{ct} are CBSA by year fixed effects. The structure of this regression is similar to the canonical debt-income regression, but here the goal is to assess the relationship between house prices and income at the zip-code level, not mortgage debt and income at the individual level. Like the previous regressions, we would expect the average value of β in the price-income regression to be positive: within a city, wealthier areas should have more expensive homes. If the βs decline over time, then the prices of low-income zip codes are rising relative to those in high-income zip codes.

The monthly Zillow data are available for about 14,500 zip codes starting in April 1996. By August 2018, the end of the Zillow sample, house-price data for about 15,600 zip codes are available. Because the regressions are at the zip-year level, we create two samples that together generate annual data from 1996 through 2018. The first sample consists of yearly averages of monthly data from January through August, and therefore include the years 1997 to 2018. The second sample uses the monthly data from April through December and can extend from 1996 to 2017.[42]

The estimated βs from the (unweighted) zip-code level regressions are plotted in the top panel of Figure 16.[43] Until 2002, the coefficients are flat, indicating that the prices of homes in low-income zip codes were quite stable relative to home prices in higher-income areas of the same city. This stability argues against significant and immediate effects resulting from the modest and exogenous credit expansion that we have documented for the 1990s. After 2002, however, the relative prices of houses in low-income areas began to rise (that

[42] We also created both balanced and unbalanced samples consisting of 12-month averages from 1996 to 2017. The results were essentially identical to those using the two main samples described above.

[43] Standard errors are excluded for clarity. When the data are clustered by zip code, the errors for any individual year range from about 0.013 to 0.017. When the data are clustered by CBSA, the standard errors range from 0.038 to 0.068. Also, there are seven zip codes for which median household income is topcoded at $200,000 in the 2000 census and five for which it is bottomcoded at –$2,500. These zip codes are excluded from all regressions.

is, the estimated income coefficients β began to decline). This pattern is consistent with previous empirical findings, and has often been interpreted as evidence that an exogenous credit expansion during the 2000s did raise housing prices.

The behavior of income coefficients after the boom ended can help reconcile these competing interpretations. During the housing bust, relative prices in low-income zip codes decline (that is, the income coefficients rise) until the early 2010s. After that, prices in low-income areas start rising again, even though few people would argue that the 2010s were a period of loose mortgage credit. One quantitative measure of credit availability, constructed by Li and Goodman (2015), is based on the idea that a higher ex ante probability of default for mortgages originated in a particular period indicates looser lending standards and greater credit availability at that time. The authors use a nonparametric model to estimate default probabilities for different mortgage cohorts based on their DTI ratios, LTV ratios, and FICO scores, as well as the types of mortgage products available (for example, whether the mortgage cohort includes large numbers of interest-only or 40-year loans). The model can also hold constant macroeconomic conditions such as house prices and interest rates across cohorts. Goodman (2017) writes that based on DTI ratios alone, credit conditions were seemingly "about average" in 2016 (p. 236). But the comprehensive Li-Goodman index, which incorporates many other mortgage characteristics, indicates that credit availability fell dramatically after the peak of the housing boom in 2006 and remained low thereafter. A separate credit-availability index published by the Mortgage Bankers Association displays a similar pattern.[44]

Yet in spite of this credit tightening, the regression coefficients in the upper panel of Figure 16 indicate that relative house prices in low-income areas began rising in the early 2010s. The lower panel suggests one reason why: the state of the aggregate housing market. This panel graphs Zillow's national house price estimate divided by a measure of expected annual earnings that is derived from the CPS. Specifically, the denominator in this ratio is the current-dollar value of usual median weekly earnings for full-time wage and salary workers from the CPS, annualized by multiplying it by 52.[45] The ratio is therefore an inverse measure

[44] According to the October 2018 update of the Li-Goodman index, credit availability declined by 61 percent from 2006 to 2012. By 2017 availability had fallen by 66 percent relative to the 2006 peak. The Li-Goodman index, which reflects mortgage-credit conditions only for purchases of owner-occupied properties, is available at https://www.urban.org/policy-centers/housing-finance-policy-center/projects/housing-credit-availability-index. The index published by the Mortgage Bankers Association is not based on a default model, but rather aggregates the published lending guidelines of several lenders and loan securitizers. It is available at https://www.mba.org/news-research-and-resources/research-and-economics/single-family-research/mortgage-credit-availability-index.

[45] The CPS earnings data are published quarterly, but these are also seasonally adjusted. Thus, even though there are only two quarters of median weekly earnings data available for 2018 as of the time of this writing, we can use the average of those two quarters to construct an average estimate of housing affordability for 2018 as a whole.

of housing affordability at the national level. Since 2012, this ratio has been rising as house prices have been increasing more quickly than earnings.

A comparison of the two panels in Figure 16 suggests a negative correlation between the two series, which is easily confirmed with regressions. These regressions show that the lagged value of the price-income ratio enters the regression more strongly than the contemporaneous value of this ratio.[46] A visual sense of this negative correlation is provided by Figure 17, which plots the income coefficients from the regression that uses the January-August averages against the lagged price-income ratio. From 2006 to 2008, US houses were very expensive relative to incomes, and the prices of houses in low-income zip codes were relatively high (that is, the income coefficients were small). At the trough of the recent cycle (2012–2014), housing in the United States had become more affordable and the relative price of houses in low-income areas had declined. By 2018, US house prices had risen relative to income for several years, and the relative prices of houses in low-income areas had risen once more. One story consistent with this pattern is that the relative price of houses in low-income areas fluctuates with the *general affordability of US homes*, not with the tightness of mortgage-credit constraints. When housing prices are high relative to incomes, affordability concerns increase the relative demand for lower-priced homes. This theory can explain the data during the house-price increase of the early-to-mid 2000s (when credit constraints were relaxed) as well as during the early 2010s (when credit conditions were tight). It also explains why the relative price of houses was stable in the late 1990s, even though credit had eased somewhat.

6.4 Characterizing the Consequences of Technical Change

The most important takeaway from this section is that the consequences of the technical changes of the 1990s were close to what economic theory would predict. Relaxing current-income constraints should have the largest effect on individuals with steep age-earnings profiles, and homeownership rates of young persons with at least some college education did rise the most during the modest expansion of homeownership after 1994. Also consistent with theory is that increases in effective demand for owner-occupied housing were met in part by the conversion of rental properties to owner-occupied ones. This finding is important because the effect of lending standards on housing prices depends critically on the quantity margins through which higher effective demand for homeownership can be satisfied. Indeed, the zip-code level price regressions show no relative appreciation of housing prices in low-income areas during the late 1990s, as these conversions were taking place. In the 2000s,

[46] In a regression of the January-August regression coefficients on both the contemporaneous and lagged values of the price-income ratio, the contemporaneous ratio enters with a coefficient of –0.00535 and an uncorrected standard error of 0.0149, while the lagged value enters with a coefficient of –0.0742 and a standard error of 0.0147. The R-squared of this regression is 0.88. The lagged coefficient is essentially unchanged with the contemporaneous value is dropped from the regression, as is the R-squared term.

the net conversion of rented-to-owned properties reversed, suggesting that relaxed credit constraints were not the main driver of the housing cycle after 2000. As has been noted previously, the relative prices of lower-valued homes rose during the early 2000s boom, but the behavior of the US housing market after 2012 suggests that this relative-price pattern had more to do with aggregate house-price increases in general, as opposed to credit-led booms in particular. Although we do not sketch out the analysis here, it is probably not too difficult to generate a model in which an overall increase in the prices of houses, relative to income, tends to raise the relative demand for lower-priced homes due to affordability concerns.

7 Conclusion

During the last three decades, improvements in information technology have transformed numerous aspects of American life. Mortgage lending is no different. During the 1990s, both mortgage and consumer lending were enhanced by technologies that not only processed data more rapidly, but could also evaluate risk more efficiently and accurately than humans could. For the mortgage industry, the new empirical models downplayed the role of current income in future mortgage default, so the 1990s saw mortgage credit expand exogenously with respect to income, a fact that makes the 1990s a particularly valuable period to consider when studying how changes in lending standards affect the housing market. By and large, the modest effects of this exogenous change on the overall market are in line with theory, but the effects run counter to the claim that an exogenous change in standards during the 2000s was responsible for the period's aggregate housing boom.

A question left unaddressed by this paper is what *did* cause the housing boom. Several papers have noted that numerous stylized facts about the housing boom make sense if expectations about future house prices were overly optimistic. Yet this simply pushes the question back one step: why, as the 2000s began, did expectations for housing prices rise so much? One possibility, explored by Kaplan, Mitman, and Violante (2017), is that some discrete event caused expected future rents to increase. House prices rose immediately in anticipation of the higher future rents, and then crashed when the rent increase did not materialize. This model can be solved with standard rational-expectations methods and is potentially consistent with the cross-section of house-price increases across US states. Figure 18 graphs state-level house price growth from 2003 to 2007 against US Census projections made in 1996 for state-level population growth from 2000 to 2025.[47] There is a strong upward

[47] The census population projections were taken from Campbell (1997) and the methodology for constructing them is described in Campbell (1996). The Census Bureau constructed updated state-level population estimates in 2004, and these projections are also correlated positively with house-price growth from 2003 to 2007. However, in unreported regressions, we found that house-price increases are more closely correlated

relationship between these variables, both in the raw data (top panel) and in a binned and population-weighted scatter plot (bottom panel). These plots are consistent with expected housing shortages and rent increases that were projected to be more acute in Sun Belt states such as Arizona, California, and Florida than they were in Rust Best states such as Indiana, Michigan, and Ohio.[48] Another story might go farther and drop the rational-expectations assumption: shocks to owner-occupied housing demand could have generated price increases that were irrationally extrapolated by agents to continue into the future. No matter how expectations are formed, the resulting boom is likely to be stronger if agents are able to trade on their expectations with few frictions. By removing such frictions, the technological improvements to mortgage lending in the 1990s could turn out to be an important reason why the housing cycle of the 2000s was so pronounced.

with the 1996 projections than they were with the 2004 projections and with actual state-level 1980–1995 population growth when all three variables are entered in the same regression. The 2004 projections are available from the Centers for Disease Control at https://wonder.cdc.gov/population-projections.html and the methodology for constructing them is available at https://wonder.cdc.gov/WONDER/help/populations/population-projections/methodology.html.

[48]See also Nathanson and Zwick (2018) for a model that links house-price increases to expected future development constraints ("arrested development") in particular localities.

References

Acemoglu, Daron, and David Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, Orley Ashenfelter and David Card, eds., vol. 4, Part B, chap. 12, 1043–1171. Elsevier. doi:10.1016/s0169-7218(11)02410-5.

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): 1635–1670. doi:10.1093/rfs/hhw018.

Albanesi, Stefania, Giacomo De Giorgi, and Jaromir Nosal. 2017. "Credit Growth and the Financial Crisis: A New Narrative." Technical Report. National Bureau of Economic Research. doi:10.3386/w23740.

Avery, Robert B., Neil Bhutta, Kenneth P. Brevoort, and Glenn B. Canner. 2012. "The Mortgage Market in 2011: Highlights from the Data Reported Under the Home Mortgage Disclosure Act." *Federal Reserve Bulletin* 98(6): 1–46.

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: 621–648.

Avery, Robert B., and Kenneth P. Brevoort. 2015. "The Subprime Crisis: Is Government Housing Policy to Blame?" *Review of Economics and Statistics* 97(2): 352–363. doi:10.1162/rest_a_00491.

Avery, Robert B., Kenneth P. Brevoort, and Glenn Canner. 2007. "Opportunities and Issues in Using HMDA data." *Journal of Real Estate Research* 29(4): 351–380.

Bhutta, Neil. 2010. "GSE Activity and Mortgage Supply in Lower-Income and Minority Neighborhoods: The Effect of the Affordable Housing Goals." *The Journal of Real Estate Finance and Economics* 45(1): 238–261. doi:10.1007/s11146-010-9258-z.

Bhutta, Neil. 2011. "The Community Reinvestment Act and Mortgage Lending to Lower Income Borrowers and Neighborhoods." *Journal of Law and Economics* 54(4): 953–983. doi:10.1086/661938.

Bhutta, Neil, Jane Dokko, and Hui Shan. 2017. "Consumer Ruthlessness and Mortgage Default during the 2007 to 2009 Housing Bust." *Journal of Finance* 72(6): 2433–2466.

Bogdon, Amy S. 2000. “Editor’s Introduction: Expansion of Mortgage Markets and Increasing Use of Technology.” *Journal of Housing Research* 11(2): 161–172. Available at http://www.jstor.org/stable/24833778.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru. 2018. “Fintech, regulatory arbitrage, and the rise of shadow banks.” *Journal of Financial Economics* 130(3): 453–483. doi:10.1016/j.jfineco.2018.03.011.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo. 2016. “Understanding Booms and Busts in Housing Markets.” *Journal of Political Economy* 124(4): 1088–1147. doi:10.1086/686732.

Campbell, John Y., and João F. Cocco. 2015. “A Model of Mortgage Default.” *The Journal of Finance* 70(4): 1495–1554. doi:10.1111/jofi.12252.

Campbell, Paul. 1996. ”Population Projections for States by Age, Sex, Race, and Hispanic Origin: 1995 to 2025,” Issued by the U.S. Census Bureau as PPL-47. October. Available at https://www.census.gov/content/dam/Census/programs-surveys/population-projections/methodology/ppl-47.pdf.

Campbell, Paul. 1997. ”Population Projections, States: 1995–2025,” Issued by the U.S. Census Bureau as *Current Population Reports: Population Projections, P25–1131*. May. Available at https://www.census.gov/content/dam/Census/library/publications/1997/demo/p25-1131.pdf.

Colton, Kent W. 2002. “Housing Finance in the United States: The Transformation of the U.S. Housing Finance System.” Harvard Joint Center for Housing Studies Working Paper. Available at http://www.jchs.harvard.edu/sites/default/files/w02-5_colton.pdf.

Corradin, Stefano. 2014. “Household Leverage.” *Journal of Money, Credit and Banking* 46(4): 567–613. doi:10.1111/jmcb.12118.

Dallas, William D. 1996. “A Time of Innovation.” *Mortgage Banking* 57(1): 20–26.

David, Paul A. 1990. “The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox.” *American Economic Review* 80(2): 355–361.

Eggers, Frederic J. 2007. “The American Housing Survey and Non-Traditional Mortgage Products.” Report prepared for the U.S. Department of Housing and Urban Development, Office of Policy Development & Research. Available at https://www.huduser.gov/publications/pdf/mortgage_prod.pdf.

Eggers, Frederick J. 2008. "Data on Home Mortgage Finance from the 1997, 1999, 2001, 2003 and 2005 American Housing Surveys." Report prepared for the U.S. Department of Housing and Urban Development, Office of Policy Development & Research. Available at https://www.huduser.gov/portal/publications/ahsrep/data_onhome.html.

Einav, Liran, Mark Jenkins, and Jonathan Levin. 2013. "The impact of credit scoring on consumer lending." *The RAND Journal of Economics* 44(2): 249–274. doi: 10.1111/1756-2171.12019.

Evans, David S., and Richard Schmalensee. 2005. *Paying with Plastic: The Digital Revolution in Buying and Borrowing.* Second edition. Cambridge, MA: MIT Press.

Federal Reserve and Conference of State Bank Supervisors. 2016. "Community Banking in the 21st Century: 2016." Available at https://www.communitybanking.org/~/media/files/communitybanking/2016/cb21townhallpub_2016_web.pdf.

Fetter, Daniel K. 2013. "How Do Mortgage Subsidies Affect Home Ownership? Evidence from the Mid-Century GI Bills." *American Economic Journal: Economic Policy* 5(2): 111–147. doi:10.1257/pol.5.2.111.

Foote, Christopher, Kristopher Gerardi, Lorenz Goette, and Paul Willen. 2010. "Reducing Foreclosures: No Easy Answers." *NBER Macroeconomics Annual* 24(1): 89–138. doi: 10.1086/648289.

Foote, Christopher L., Lara Loewenstein, and Paul S. Willen. 2016. "Cross-Sectional Patterns of Mortgage Debt During the Housing Boom: Evidence and Implications." Federal Reserve Bank of Boston Working Paper No. 16-12. Available at https://www.bostonfed.org/-/media/Documents/Workingpapers/PDF/wp1612.pdf.

Foote, Christopher L., and Paul S. Willen. 2018. "Mortgage-Default Research and the Recent Foreclosure Crisis." *Annual Review of Financial Economics* 10(1): 59–100. doi:10.1146/annurev-financial-110217-022541.

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther. 2017. "Predictably Unequal? The Effects of Machine Learning on Credit Markets." doi:10.2139/ssrn.3072038.

Fuster, Andreas, Laurie S. Goodman, David Lucca, Laurel Medar, Linsey Molloy, and Paul S. Willen. 2013. "The Rising Gap between Primary and Secondary Mortgage Rates." Federal Reserve Bank of New York *Economic Policy Review* 19(2): 17–40.

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, James Vickery, Andreas Fuster, Matthew Plosser, Philipp Schnabl, and James Vickery. 2018. "The Role of Technology in Mortgage Lending." Working Paper 24500. National Bureau of Economic Research. doi:10.3386/w24500.

Gates, Susan Wharton, Vanessa Gail Perry, and Peter M. Zorn. 2002. "Automated underwriting in mortgage lending: Good news for the underserved?" *Housing Policy Debate* 13(2): 369–391. doi:10.1080/10511482.2002.9521447.

Gennaioli, Nicola, and Andrei Shleifer. 2018. *A Crisis of Beliefs: Investor Psychology and Financial Fragility.* Princeton, NJ: Princeton University Press.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny. 2012. "Neglected risks, financial innovation, and financial fragility." *Journal of Financial Economics* 104(3): 452–468. doi:10.1016/j.jfineco.2011.05.005.

Goodman, Laurie S. 2017. "Quantifying the Tightness of Mortgage Credit and Assessing Policy Actions." *Boston College Journal of Law and Social Justice* 37(2): 235–265.

Goodman, Laurie S., Jim Parrott, and Jun Zhu. 2015. "The Impact of Early Efforts to Clarify Mortgage Repurchases: Evidence from Freddie Mac and Fannie Mae's Newest Data." Urban Institute Housing Finance Policy Brief, April. Available at https://www.urban.org/sites/default/files/publication/43676/2000142-The-Impact-of-Early-Efforts-to-Clarify-Mortgage-Repurchases.pdf.

Herzog, John P., and James S. Earley. 1970. *Home Mortgage Delinquency and Foreclosure.* National Bureau of Economic Research General Series No. 91. New York, NY: Columbia University Press. Available at http://papers.nber.org/books/herz70-1.

Hurst, Erik, Benjamin J. Keys, Amit Seru, and Joseph Vavra. 2016. "Regional Redistribution through the US Mortgage Market." *American Economic Review* 106(10): 2982–3028. doi:10.1257/aer.20151052.

Kaplan, Greg, Kurt Mitman, and Giovanni Violante. 2017. "The Housing Boom and Bust: Model Meets Evidence." Working Paper 23694. National Bureau of Economic Research. doi:10.3386/w23694.

Kau, James B., Donald C. Keenan, and Taewon Kim. 1994. "Default Probabilities for Mortgages." *Journal of Urban Economics* 35(3): 278–296. doi:10.1006/juec.1994.1017.

LaCour-Little, Michael. 2000. "The Evolving Role of Technology in Mortgage Finance." *Journal of Housing Research* 11(2): 173–205.
Available at http://www.jstor.org/stable/24833779.

Lam, Ken, and Bulbul Kaul. 2003. "Analysis of Housing Finance Issues using the American Housing Survey." Report prepared for the U.S. Department of Housing and Urban Development, Office of Policy Development & Research. Available at https://www.census.gov/content/dam/Census/programs-surveys/ahs/publications/AhsAnalysis.pdf.

Laufer, Steven. 2018. "Equity extraction and mortgage default." *Review of Economic Dynamics* 28: 1–33. doi:10.1016/j.red.2017.07.006.

Levy, Frank, and Richard J. Murnane. 2004. *The New Division of Labor: How Computers Are Creating the Next Job Market.* Princeton, NJ: Princeton University Press.

Li, Wei, and Laurie S. Goodman. 2015. "A New Mortgage Credit Availability Index." *The Journal of Structured Finance* 20(4): 67–83. doi:10.3905/jsf.2015.20.4.067.

Livshits, Igor, James C. Mac Gee, and Michèle Tertilt. 2016. "The Democratization of Credit and the Rise in Consumer Bankruptcies." *The Review of Economic Studies* 83(4): 1673–1710. doi:10.1093/restud/rdw011.

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.

Makuch, William M. 1998. "Scoring Applications." In *Credit Risk Modeling: Design and Application*, Elizabeth Mays, ed., chap. 1, 1–12. Chicago, IL: Glenlake Publishing Company, Ltd.

McDonald, David W., Charles O. Pepe, Henry M. Bowers, and Edward J. Dombroski. 1997. "Desktop Underwriter: Fannie Mae's Automated Mortgage Underwriting Expert System." In *Innovative Applications of Artificial Intelligence–Conference Proceedings*, 875–882. American Association for Artificial Intelligence. Proceedings of the 1997 9th Conference on Innovative Applications of Artificial Intelligence, Providence, RI, 27–31 July.

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): 1449–1496. doi:10.1162/qjec.2009.124.4.1449.

Mian, Atif, and Amir Sufi. 2017. "Fraudulent Income Overstatement on Mortgage Applications During the Credit Expansion of 2002 to 2005." *The Review of Financial Studies* 30(6): 1832–1864. doi:10.1093/rfs/hhw104.

Mincer, Jacob. 1991. "Education and Unemployment." Working Paper 3838. National Bureau of Economic Research. doi:10.3386/w3838.

Morgenson, Gretchen, and Joshua Rosner. 2011. *Reckless Endangerment: How Outsized Ambition, Greed, and Corruption Led to Economic Armageddon.* New York, NY: Times Books.

Munnell, Alicia H., Geoffrey M. B. Tootell, Lynn E. Browne, and James McEneaney. 1996. "Mortgage Lending in Boston: Interpreting HMDA Data." *The American Economic Review* 86(1): 25–53.
Available at http://www.jstor.org/stable/2118254.

Nathanson, Charles G., and Eric Zwick. 2018. "Arrested Development: Theory and Evidence of Supply-Side Speculation in the Housing Market." *The Journal of Finance* doi:10.1111/jofi.12719.

Philippon, Thomas. 2018. "The Fintech Opportunity." New York University Working Paper. Available at http://pages.stern.nyu.edu/~tphilipp/papers/FinTech.pdf.

Poon, Martha. 2009. "From new deal institutions to capital markets: Commercial consumer risk scores and the making of subprime mortgage finance." *Accounting, Organizations and Society* 34(5): 654–674. doi:10.1016/j.aos.2009.02.003.

Rajan, Raghuram. 2010. *Fault Lines: How Hidden Fractures Still Threaten the World Economy.* Princeton, NJ: Princeton University Press.

Rosenthal, Stuart S. 2014. "Are Private Markets and Filtering a Viable Source of Low-Income Housing? Estimates from a "Repeat Income" Model." *American Economic Review* 104(2): 687–706. doi:10.1257/aer.104.2.687.

Rudolph, Patricia M., and John Griffith. 1997. "Integration of the Mortgage Market into the National Capital Markets: 1963–1993." *Journal of Housing Economics* 6(2): 164–183. doi:10.1006/jhec.1997.0207.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2018. "IPUMS USA: Version 8.0 [American Community Survey]." doi:10.18128/D010.V8.0.

Schelkle, Thomas. 2018. "Mortgage Default during the U.S. Mortgage Crisis." *Journal of Money, Credit and Banking* 50(6): 1101–1137. doi:10.1111/jmcb.12546.

Stamper, Michael K. 1995. "The Predictive Power of Selected Credit Scores." Available as Attachment 2 of "Statement of Edward J. Pinto before the Committee on Oversight and Government Reform, United States House of Representatives, December 9, 2008." https://democrats-oversight.house.gov/sites/democrats.oversight.house.gov/files/documents/FannieFreddieTestimonyofEdwardPinto12.9.08writtensubmissionFull.pdf.

Straka, John W. 2000. "A Shift in the Mortgage Landscape: The 1990s Move to Automated Credit Evaluations." *Journal of Housing Research* 11(2): 207–232.
Available at http://www.jstor.org/stable/24833780.

Talebzadeh, Houman, Sanda Mandutianu, and Christian F. Winner. 1995. "Countrywide Loan-Underwriting Expert System." *AI Magazine* 16(1): 51–64.

U.S. Department of Housing and Urban Development. 1996. "Privatization of Fannie Mae and Freddie Mac: Desirability and Feasibility." Report prepared by the Office of Policy Development & Research. Available at https://www.huduser.gov/portal/Publications/pdf/private.pdf.

U.S. Department of Housing and Urban Development. 2017. "Introducing the 2015 American Housing Survey (Slide Presentation)." Office of Policy Development & Research. Available at https://www.huduser.gov/portal/sites/default/files/pdf/Introducing-the-2015-American-Housing-Survey-Slides-and-Notes.pdf.

Wallison, Peter J., and Edward J. Pinto. 2012. "Free Fall: How Government Policies Brought Down the Housing Market." Financial Services Outlook, American Enterprise Institute. April 26. Available at http://www.aei.org/publication/free-fall-how-government-policies-brought-down-the-housing-market/.

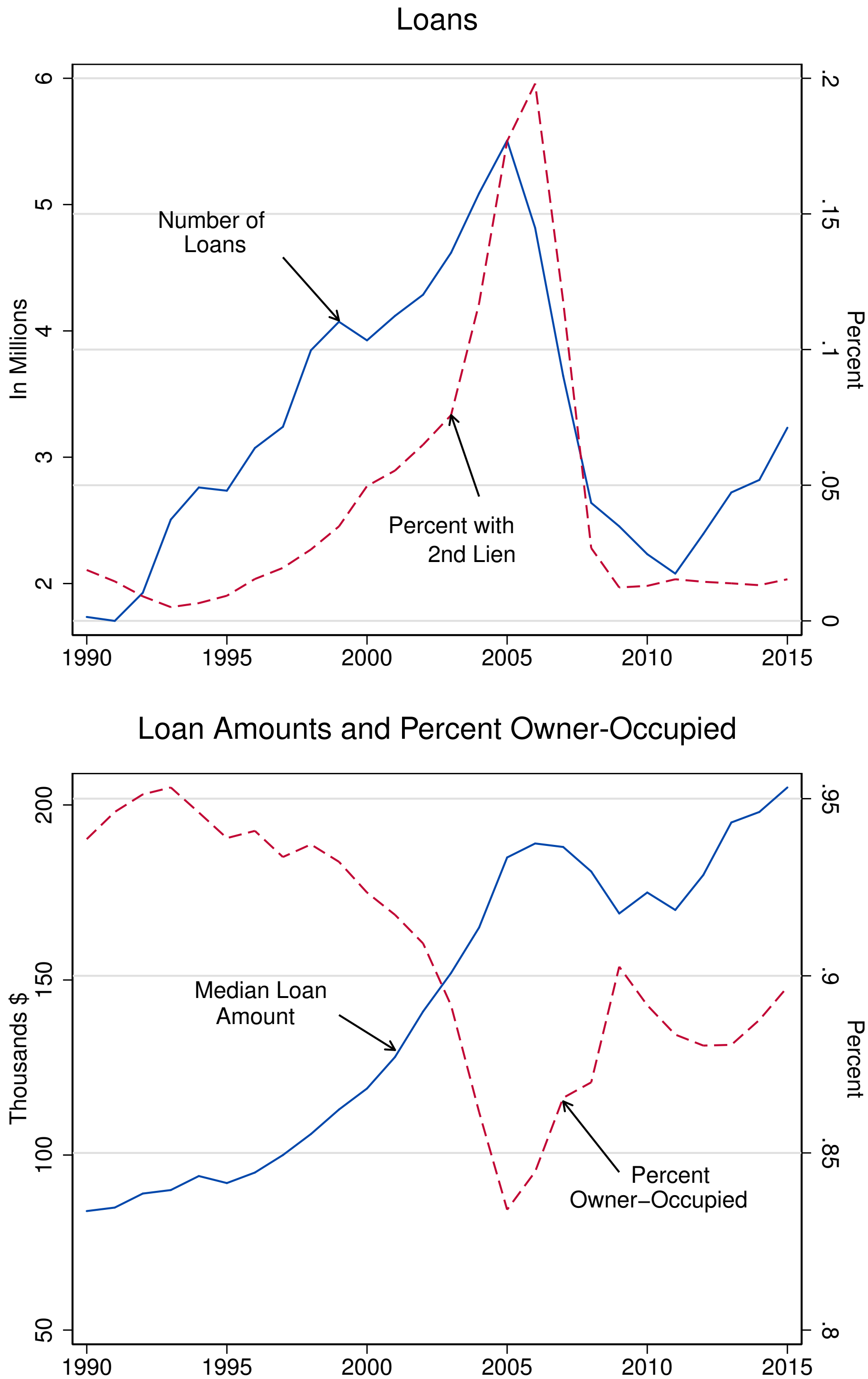

Figure 1. Summary Statistics from HMDA

Note: The total median loan amount refers to the median value of the sum of first liens and associated piggyback loans on owner-occupied properties. The percent with second lien is the share of all loans that have a piggyback loan.

Source: Microdata from the Home Mortgage Disclosure Act.

AHS Survey Year	(1) Total AHS Observations	(2) Occupied Interviews w/Nonzero Weights	(3) Homeowners	(4) Recently Moving Owners...	(5) ...with Nonzero Income & Mortgage Debt	(6) Truncate Top & Bottom 5% of Debt & Income (Baseline Sample)
1985	53,558	37,470	24,312	3,232	2,880	2,338
1987	54,052	43,436	28,857	3,184	2,740	2,259
1989	58,942	39,399	25,557	2,435	2,093	1,717
1991	59,491	44,764	29,608	2,851	2,440	2,004
1993	64,998	40,931	26,460	2,433	2,087	1,701
1995	63,143	45,675	29,384	3,436	3,424	2,742
1997	58,287	39,981	26,309	2,664	2,635	2,136
1999	67,177	46,589	30,799	3,228	3,198	2,587
2001	62,314	42,487	28,703	2,892	2,853	2,322
2003	71,170	48,197	32,586	3,209	3,174	2,589
2005	69,020	43,360	29,603	3,395	3,370	2,735
2007	65,419	39,107	26,671	2,616	2,598	2,094
2009	73,222	45,057	30,228	2,177	2,155	1,747
2011	186,448	134,918	82,418	5,372	5,313	4,252
2013	84,355	60,097	35,852	2,075	2,042	1,621
Totals					43,002	34,844

Table 1. Unweighted Sample Counts in the American Housing Survey
Source: Microdata from the American Housing Survey.

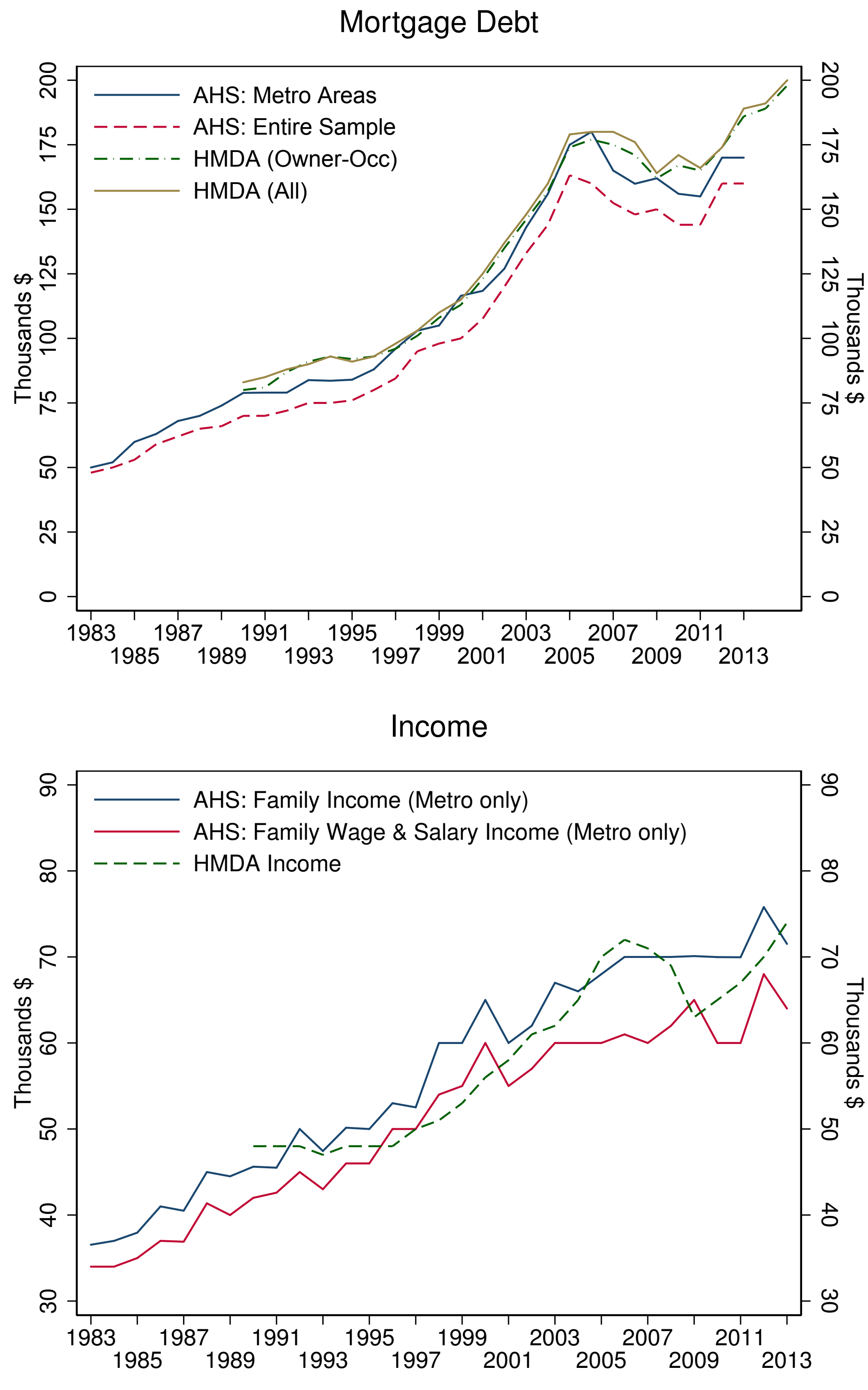

Figure 2. Comparison of New Mortgage Debt and Income from the Home Mortgage Disclosure Act and the American Housing Survey
Source: Microdata from the Home Mortgage Disclosure Act and the American Housing Survey.

Figure 3. Out-of-State Lending in the HMDA Records
Note: Lenders are matched to their branch locations, parents and bank holding companies (if any) using their Federal Reserve Board Entity number. Mortgage companies and credit unions are excluded from the sample.
Source: Home Mortgage Disclosure Act, the National Information Center, and the Federal Deposit Insurance Corporation Summary of Deposits.

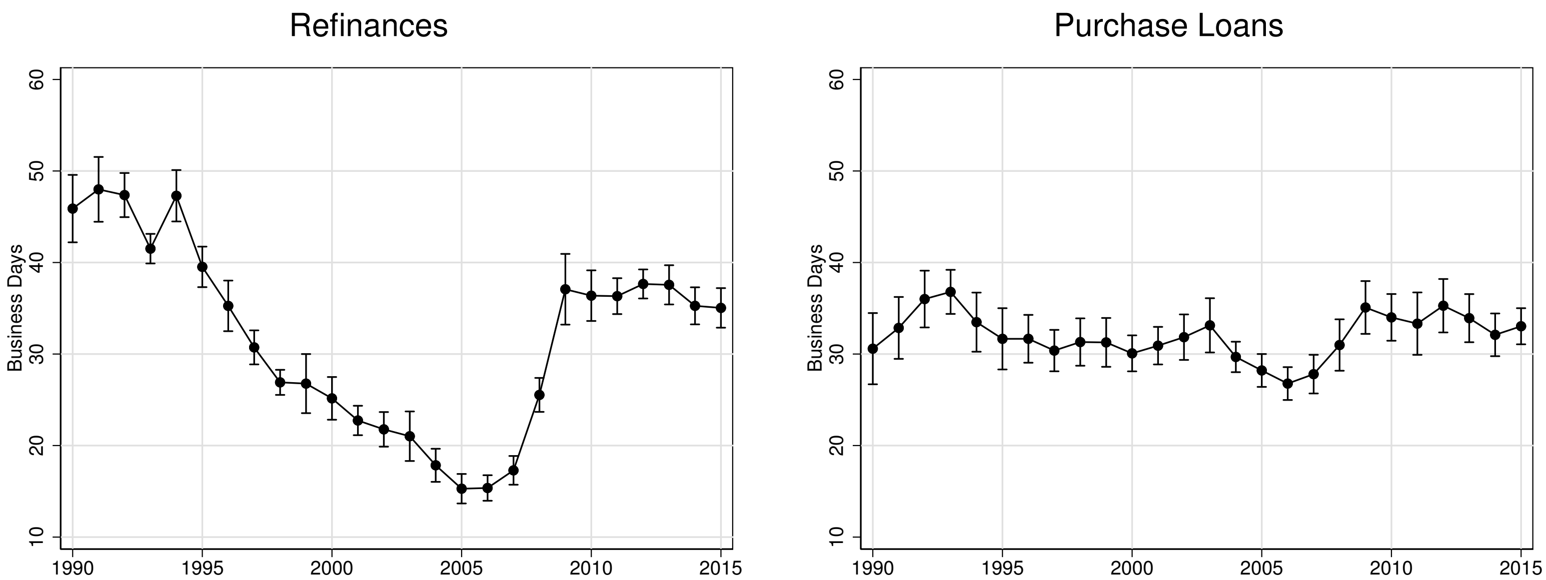

Figure 4. Time-to-Close Regressions

Note: Each panel shows a plot of the average processing time by year after stripping out any variation explained by the size of the lender, the borrower's race and gender, whether the borrower has a coapplicant, and the concurrent monthly application volume. The processing times are calculated as of the year of application and include both closed loans and denials.

Source: Microdata from the Home Mortgage Disclosure Act.

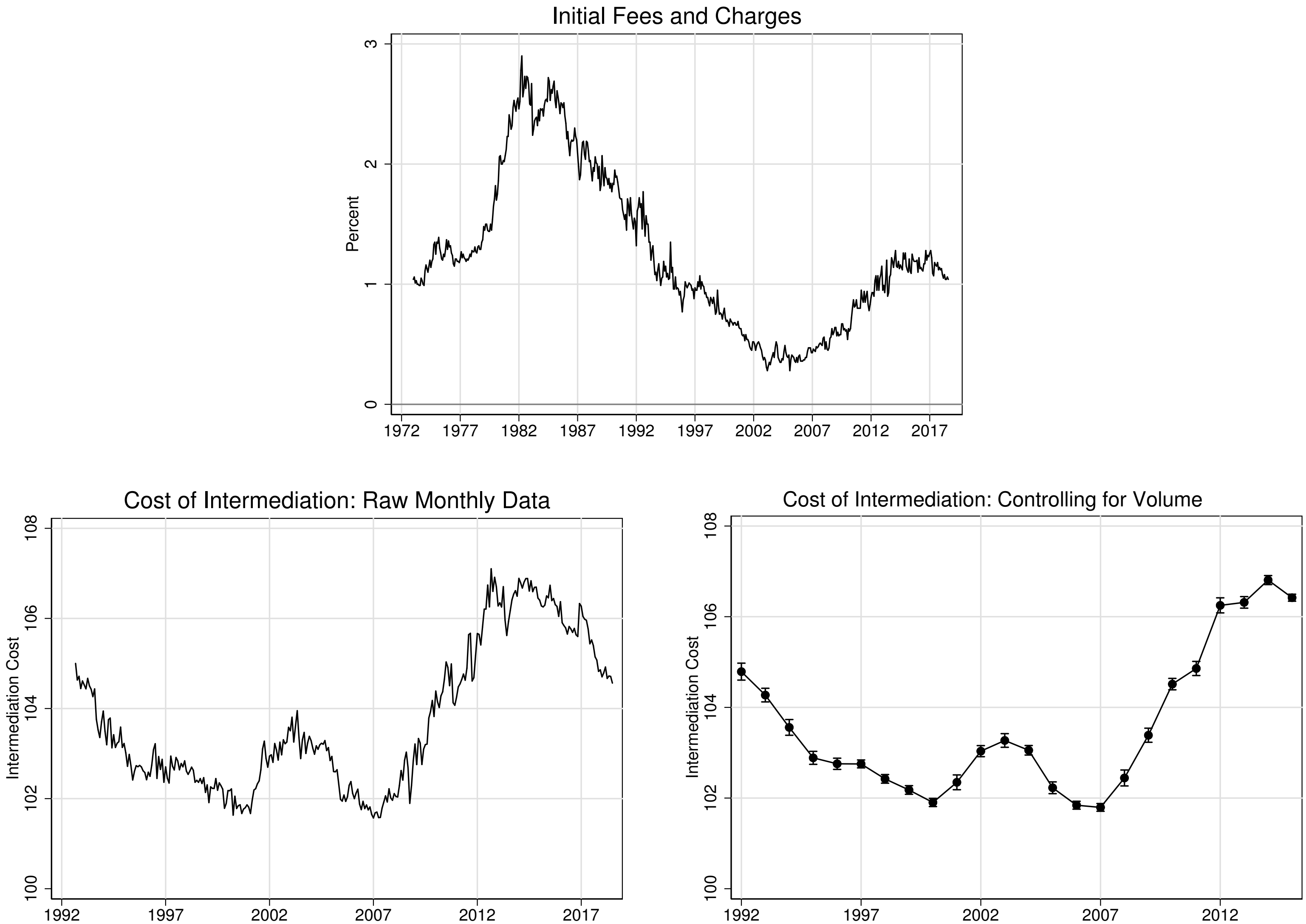

Figure 5. Costs of Intermediation

Note: The intermediation cost is defined the amount that a borrower pays for $100 worth of mortgage credit. The measure of volume used to construct the lower right panel is the number of loan applications per business day in each month in HMDA.

Source: Federal Housing Finance Agency, Home Mortgage Disclosure Act, and publicly available data used for Fuster et al. (2013).

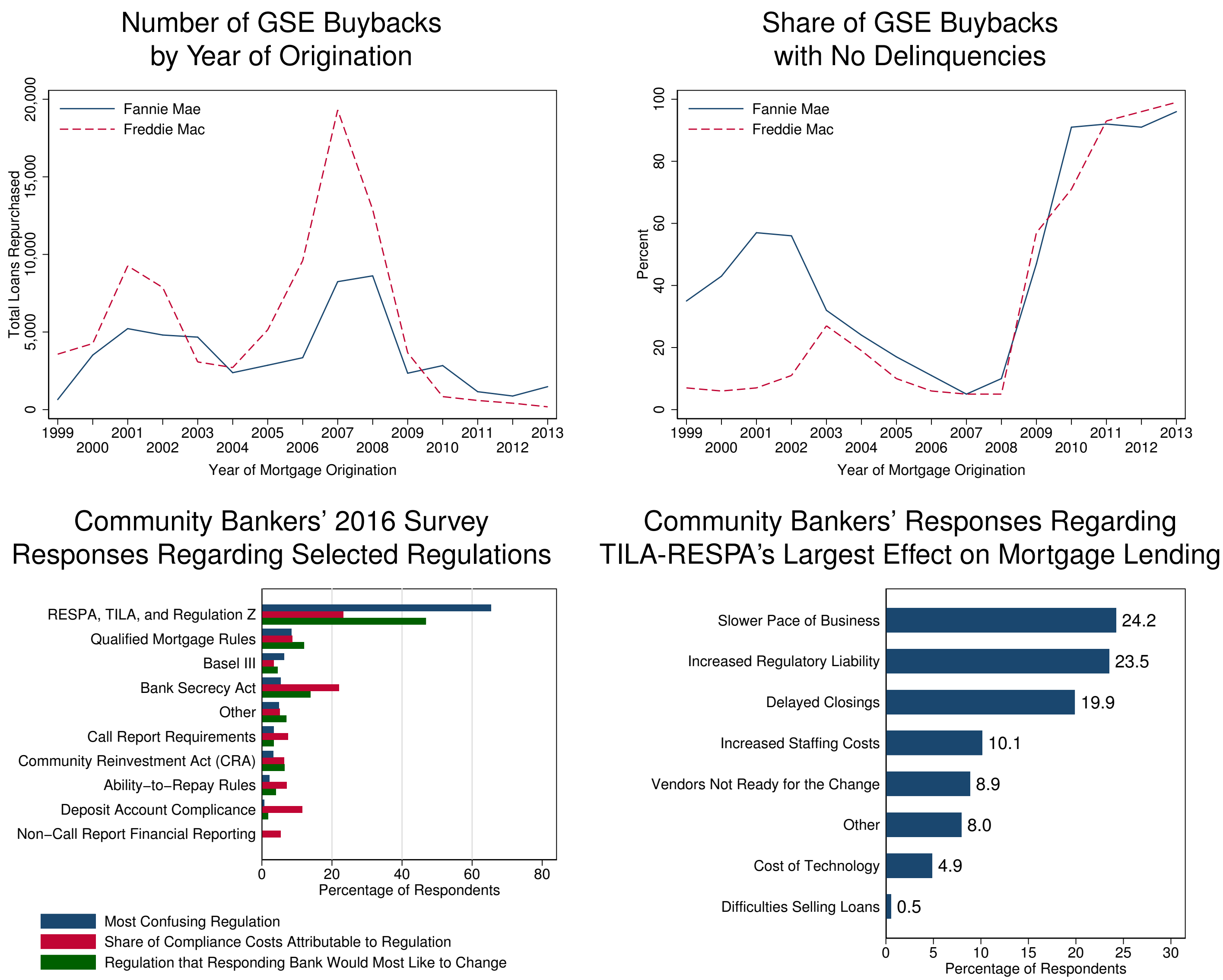

Figure 6. Explaining the Decline in Mortgage-Lending Efficiency after the Housing Boom
Source: Top panels: Table 1 of Goodman, Parrott, and Zhu (2015); bottom panels: Figures 12–15 of Federal Reserve and Conference of State Bank Supervisors (2016).

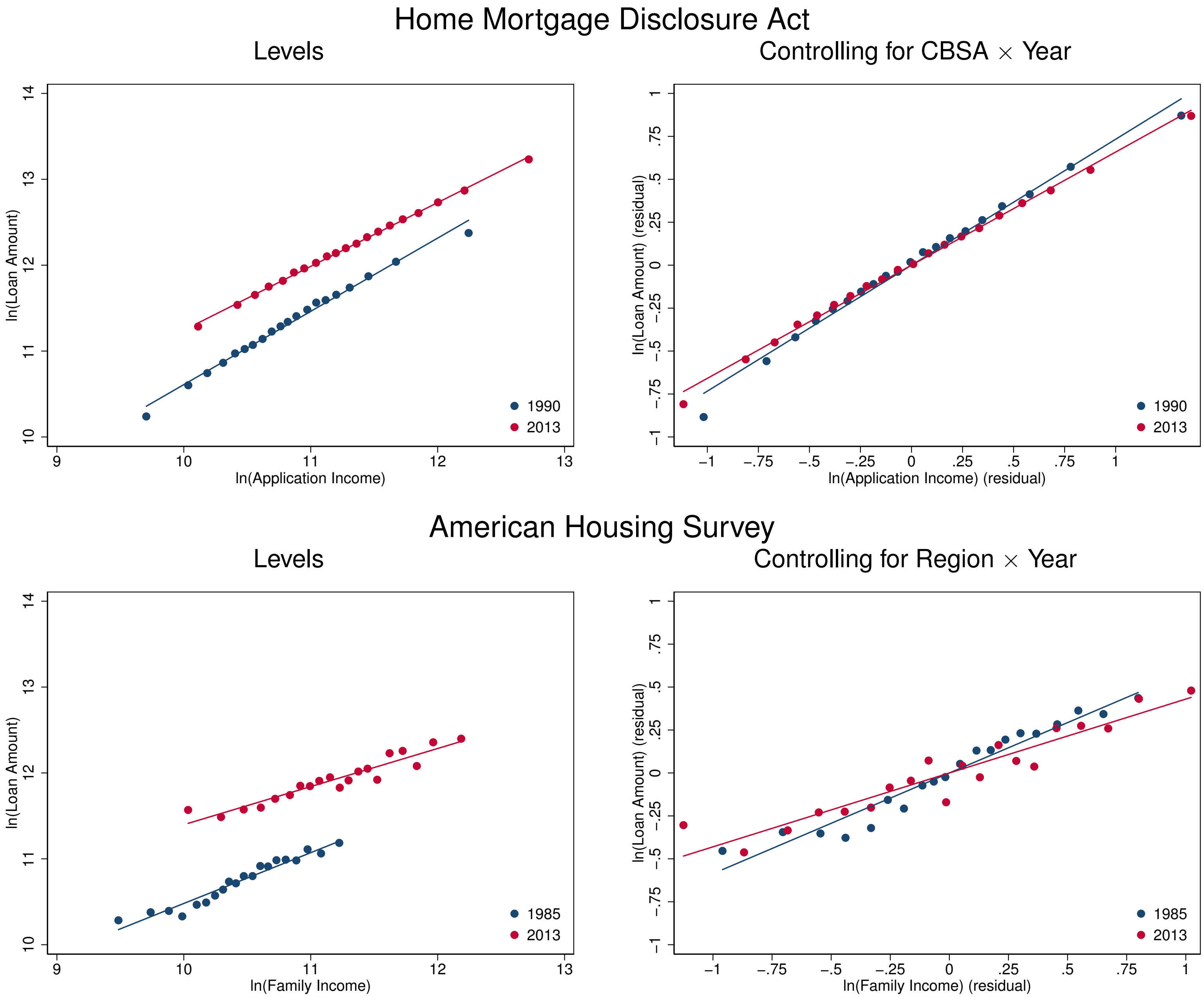

Figure 7. Binned Scatter Plots of New Mortgage Debt and Income at the Individual Level
Note: Each dot plots the average loan value for a given income quantile. The two left panels are binned scatter plots of residuals from regressions of the natural log of loan amounts and income on geographic area by year fixed effects.
Source: Microdata from the American Housing Survey and the Home Mortgage Disclosure Act.

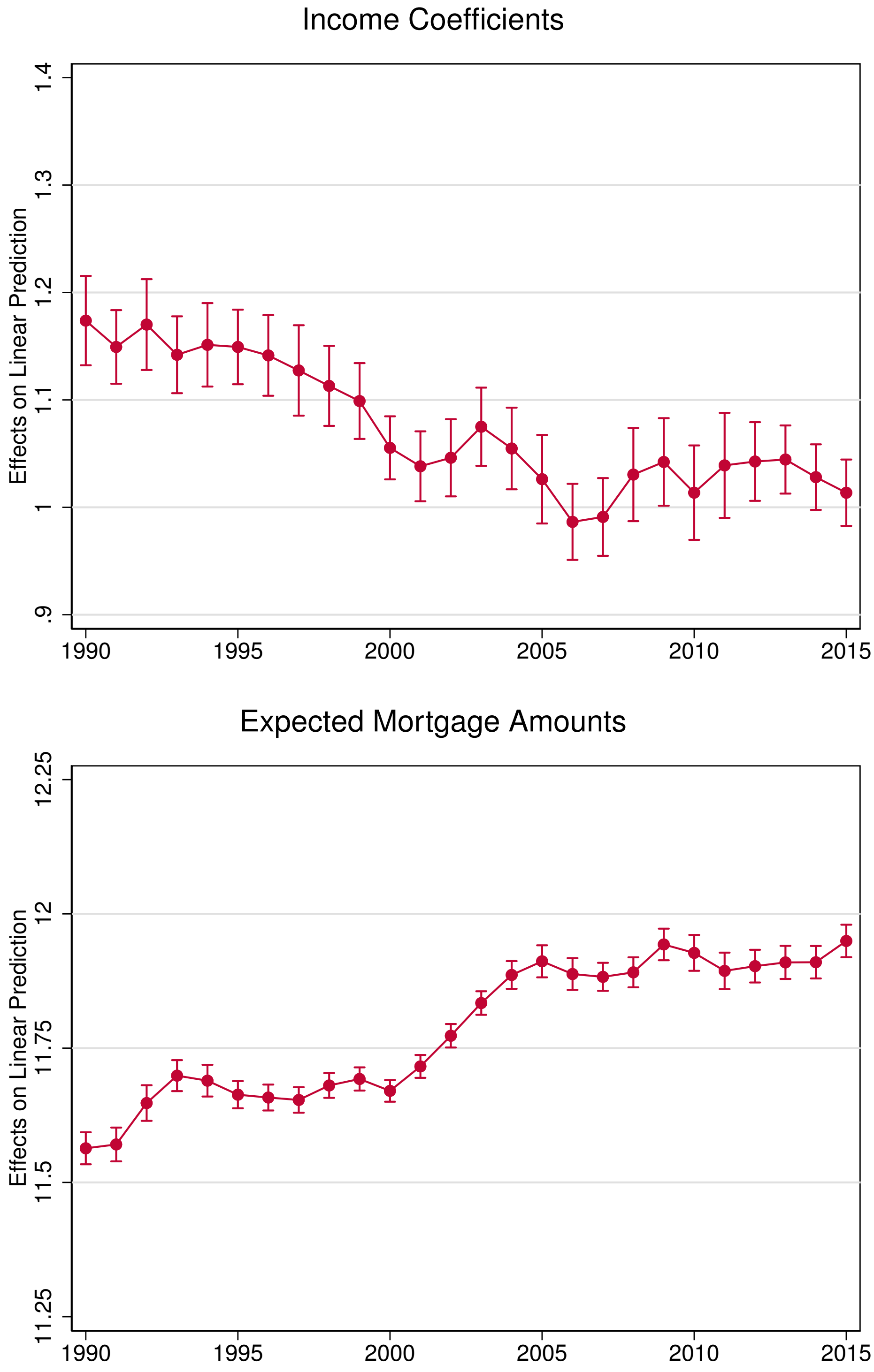

Figure 8. Regression using Individual-Level HMDA Mortgage Balances and Income Levels
Note: The top panel graphs income coefficients (and 95 percent confidence intervals) from regressions of individual purchase mortgage origination amounts from HMDA on measures of income from HMDA. The HMDA income measure is instrumented with the most recently available Census tract income from the Census and ACS. The regressions include CBSA $\times$ year fixed effects, and also control for the borrowers race and gender interacted with year. Expected mortgage amounts are predictions from an identical regression without CBSA fixed effects, holding income constant at its average value across all years.
Source: Home Mortgage Disclosure Act, US Decennial Census, and the American Community Survey.

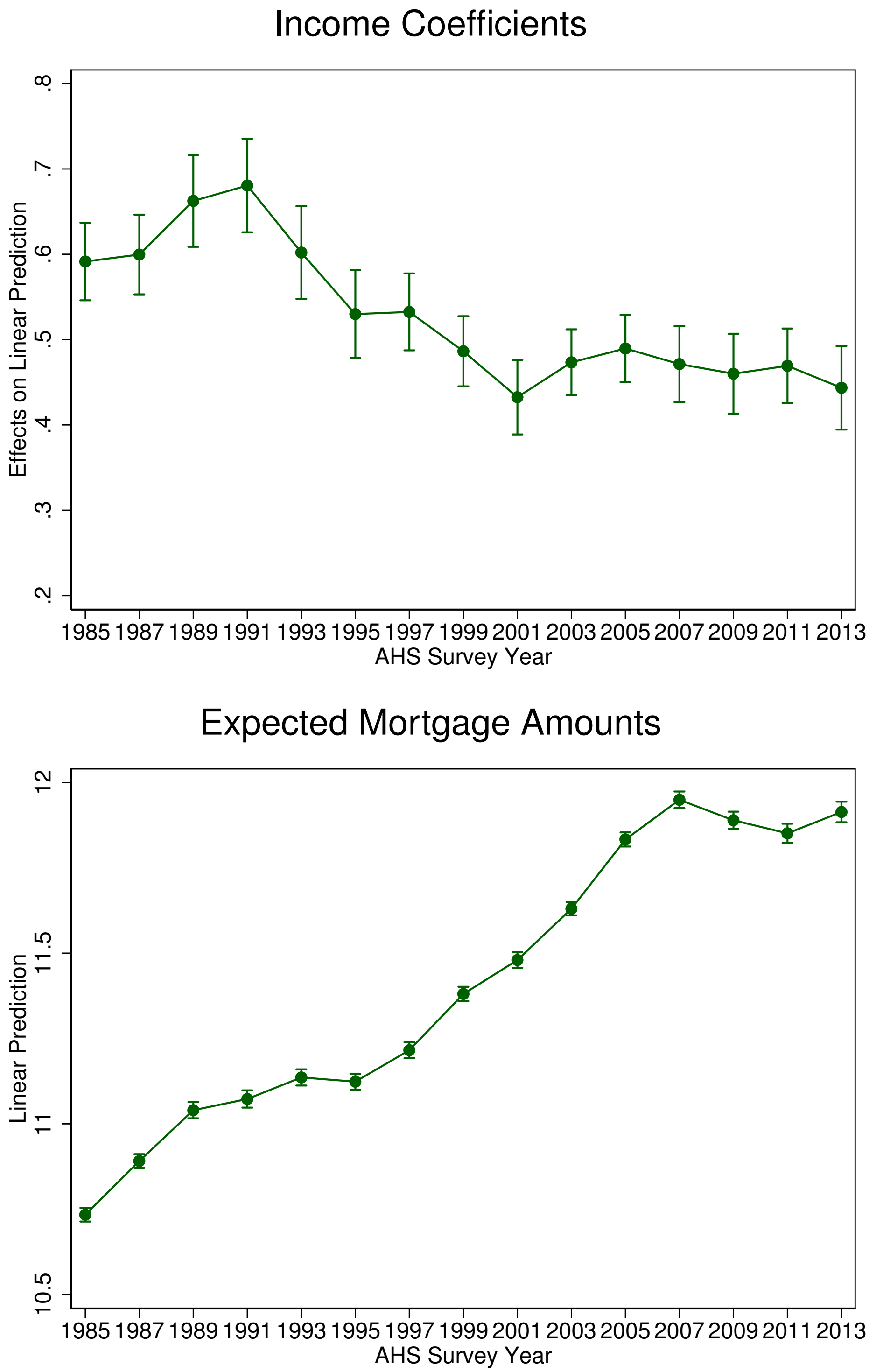

Figure 9. Regression using Individual-Level AHS Mortgage Balances and Income Levels
Note: The top panel graphs income coefficients (and 95 percent confidence intervals) from regressions of AHS mortgage amounts on AHS income at the individual level. The sample for the regressions is the baseline sample, described by the last column of Table 1. The years in the panel refer to the wave of the AHS, and the data are therefore generated by owners who move in the previous two years. Interactions between Census region and year are also included in the regression that generates the income coefficients in the top panel. The expected mortgage amounts in the lower panel are calculate as in Figure 8 from a regression that does not include region $\times$ year interactions.
Source: Microdata from the American Housing Survey.

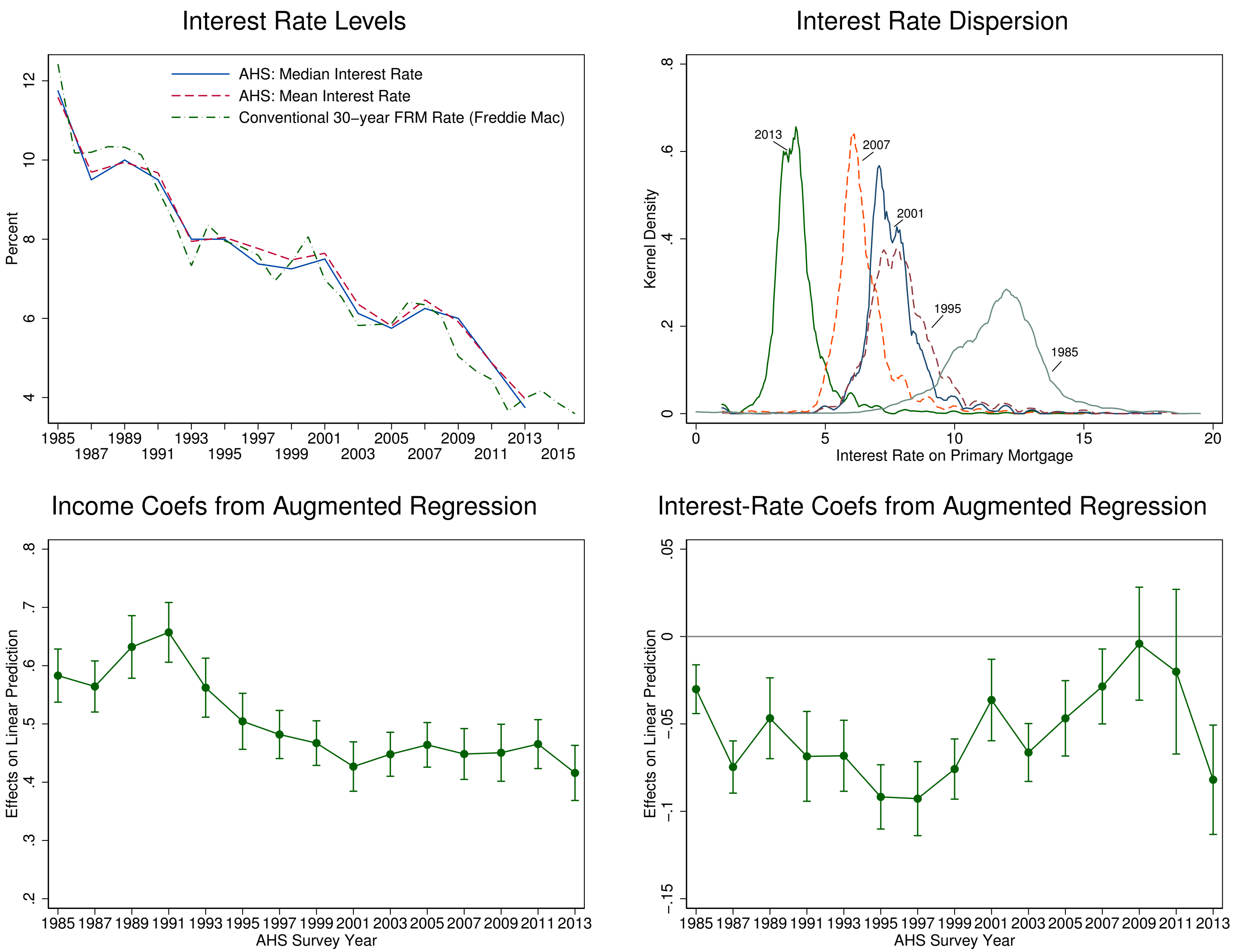

Figure 10. Interest Rates and Mortgage Lending
Source: American Housing Survey and Freddie Mac.

Freddie Mac

GOLD MEASURE WORKSHEET—Version 2.0

Borrower/Co-borrower Name(s) ______________________

City, State ______________________
Lender Loan Number ______________________
Origination Date ______________________
Completion Date ______________________

Seller Name ______________________
Freddie Mac Seller Number ______________________
Freddie Mac Loan Number (if available) ______________________
Branch Office/Channel ______________________
Underwriting Center ______________________
TPO Name ______________________
Underwriter ______________________

Loan Decision Approved __ Denied __ Withdrawn __ File Closed __

Directions:
- Circle the appropriate "Risk Units" (RUs) for each category. Total the RUs in each section and enter on the Subtotal line. Then combine Subtotals for each section and enter the Grand Total on the Total RUs line. Note that negative numbers such as "–2" are risk offsets.
- It is important to read the accompanying *Gold Measure Worksheet and Instructions—Version 2.0* booklet and to refer to it for additional information on completing this worksheet.
- This worksheet is an aid, not a substitute for the underwriting decision.
- Complete **either** Credit File A **or** Credit File B, but not both. Use Credit File A if 3 credit scores are requested. Use Credit File B if fewer than 3 credit scores are requested. See the *Gold Measure Worksheet and Instructions—Version 2.0* booklet for easy instructions on how to order bureau and bankruptcy scores for use with Credit File A.

I. Credit File A

Directions: When using Credit File A, complete **either** the Bureau Score **or** the Bankruptcy Score, but not both.

Bureau Score Equifax Beacon Score, Trans Union Empirica Score and TRW-FICO Score (See instructions)

Bankruptcy Score Equifax DAS Score, Trans Union Delphi Score and TRW-MDS Score (See instructions)

Bureau Score	RUs	Bankruptcy Score	RUs
Over 790	–16	150 or less	–12
771 – 790	–14	151 – 200	–10
761 – 770	–11	201 – 240	–4
731 – 760	–7	241 – 300	–3
721 – 730	–5	301 – 320	–1
701 – 720	0	321 – 360	0
681 – 700	6	361 – 420	4
661 – 680	8	421 – 480	8
641 – 660	12	481 – 540	11
621 – 640	17	541 – 620	15
601 – 620	20	621 – 700	18
581 – 600	23	701 – 740	21
541 – 580	25	741 – 840	23
540 or less	32	841 – 960	25
		Over 960	29
No reported Score available	20	No reported Score available	20

I. Credit File A. **Subtotal** of circled **RUs**: ______

I. Credit File B

Directions: Use Credit File B if fewer than 3 credit scores are requested.

No delinquencies or other derogatory tradeline or derogatory public record information and number of tradelines (open or closed) is:

	RUs
11 or more	–4
6 – 10	–3
1 – 5	0

One or more revolving tradelines and **total** revolving balances are under $500: **RUs** –4

Fewer than 3 tradelines (open or closed): **RUs** 2

Percent of all tradelines (open or closed) ever delinquent or worse (30-90 days or more, collection, charge-off, etc.):

	RUs
0 – 10%	–3
11 – 15%	0
16 – 40%	4
41 – 60%	8
Over 60%	11

Worst ever derogatory credit file entry is either:
- 30–180 days delinquent: **RUs** 6

or
- Public record (bankruptcy, foreclosure, judgment, lien, garnishment, suit, certain collections) or tradeline reported as over 180 days delinquent, charge-off, repossession or collection: **RUs** 10

Number of derogatory Public records:

	RUs
0 – 1	0
2 – 3	4
Over 3	9

Number of inquiries in the past 3 months:

	RUs
0	–2
1	0
2 – 3	5
4	8
5	11
More than 5	14

Age of **oldest tradeline** (in months):

	RUs
0 (no tradelines) – 6	18
7 – 12	13
13 – 24	7
25 – 48	3
49 – 72	2
73 – 120	0
121 – 168	–1
169 or more	–2

If age of oldest tradeline is 13–48 months and **any one or more** of the following:
- More than 3 inquiries within the past 3 months
- More than 3 tradelines opened in the past year
- Total open balances exceed $10,000

RUs 6

I. Credit File B. **Subtotal** of circled **RUs**: ______

II. Income

	RUs
Self-employed and above area median income:	5
Majority of income from commissions:	5
Employed second earner on application:	–2
Borrower's time on job is 5 years or more:	–2
Co-borrower's time on job is 2 years or more:	–1

II. Income. **Subtotal** of circled **RUs**: ______

III. Loan, Collateral, Assets

LTV/TLTV (including secondary financing*) is:

	RUs
60.5% or lower	–27
60.6 – 70.5%	–16
70.6 – 80.0%	–5
80.1 – 85.5%	–1
85.6 – 90.5%	0
90.6 – 93.5%	2
93.6 – 94.5%	5
94.6 – 95.5%	8
95.6 – 96.5%	10
96.6 – 98.5%	11
98.6 – 99.5%	13
99.6 – 99.9%	15

Property seller contributions exceed 3% of value: **RUs** 5

Reserves are:

	RUs
Less than 1 month	8
At least 1, but less than 2 months	5
At least 2, but less than 4 months	0
At least 4, but less than 5 months	–3
5 or more months	–6

Less than 5% down from borrower funds with 95% LTV (e.g. Affordable Gold with 3/2 Option): **RUs** 8

III. Loan, Collateral, Assets. **Subtotal** of circled **RUs**: ______

*When secondary financing is included, if the secondary financing provides for any amortization (payments) before maturity of the Freddie Mac loan, then add 1% to LTV for every rounded percentage point of secondary financing. Likewise, add 0.5% to LTV for every rounded percentage point of secondary financing, if there is no amortization (no payments due) before maturity of the Freddie Mac loan. Unsecured grants or gifts require no adjustments to LTV.

IV. Debt-Payment Burden

Debt-to-income ratio is:

	RUs
Less than 32.6%	0
32.6 – 38.5%	2
38.6 – 40.5%	4
40.6 – 42.5%	7
42.6 – 44.5%	10
44.6 – 46.5%	13
46.6 – 48.5%	15
48.6 – 50.5%	18
Over 50.5%	30

Spread between total debt and housing ratios (i.e. nonhousing debt ratio) is:

	RUs
10 to 15%	2
More than 15%	5

Proposed housing expense is less than 120% of previous housing expense: **RUs** –1

IV. Debt-Payment Burden. **Subtotal** of circled **RUs**: ______

V. Loan/Property Type

Loan type is:

		RUs
Fixed-Rate:	15-Year	–6
	20-Year	–4
	25-Year	–1
ARM:	Rate-Capped	6
	Payment-Capped	8

Property type is:

	RUs
2 Unit	5
3–4 Unit	11
Condominium	5

V. Loan/Property Type. **Subtotal** of circled **RUs**: ______

Total of sections I A or B, II, III, IV and V. **TOTAL RUs:** ______

Freddie Mac Risk Unit Guideline:	15 RUs
If pre-purchase counseling:	16 RUs
If post-purchase counseling:	17 RUs
If pre- and post-purchase counseling:	18 RUs

Refer to *Gold Measure Worksheet and Instructions—Version 2.0* booklet for more information. This worksheet is an aid, not a substitute for the underwriting decision. Call your Account Representative for additional information.

Figure 11. Freddie Mac's Gold Measure Worksheet
Source: Avery et al. (1996).

			Automated Underwriting			
			1995 Model		2000 Model	
			Accept	Caution	Accept	Caution
Manual Underwriting	Accept	51.6	24.0	27.5	44.7	6.9
		[1.04]	[0.21]	[1.75]	[0.66]	[3.52]
	Caution	48.4	20.8	27.7	42.7	5.8
		[0.96]	[0.21]	[1.52]	[0.75]	[2.55]
	Total	100	44.8	55.2	87.3	12.7
		[1]	[0.21]	[1.63]	[0.70]	[3.08]

Table 2. Comparing Manual and Automated Underwriting on a Sample of Freddie Mac Loans
Note: The non-bracketed numbers in this table refer to group shares (as percentages of the entire evaluation sample). The bracketed numbers refer to the relative default rates (relative to the sample-wide rate).
Source: Gates, Perry, and Zorn (2002).

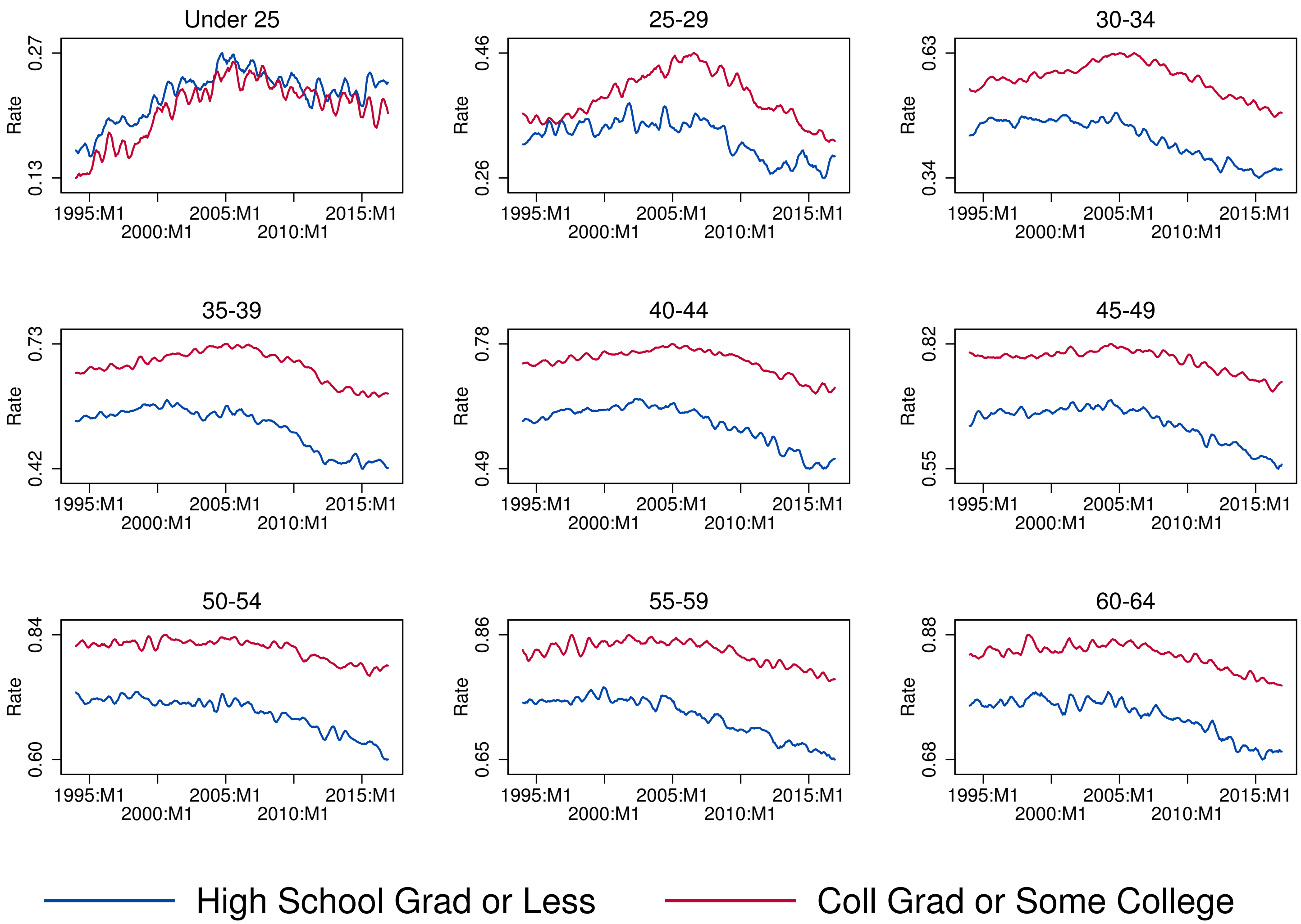

Figure 12. Homeownership Rates by Age of Householder and Education in the Current Population Survey/Housing Vacancy Survey

Note: Data are six-month moving averages of monthly rates.

Source: Microdata from the Current Population Survey/Housing Vacancy Survey.

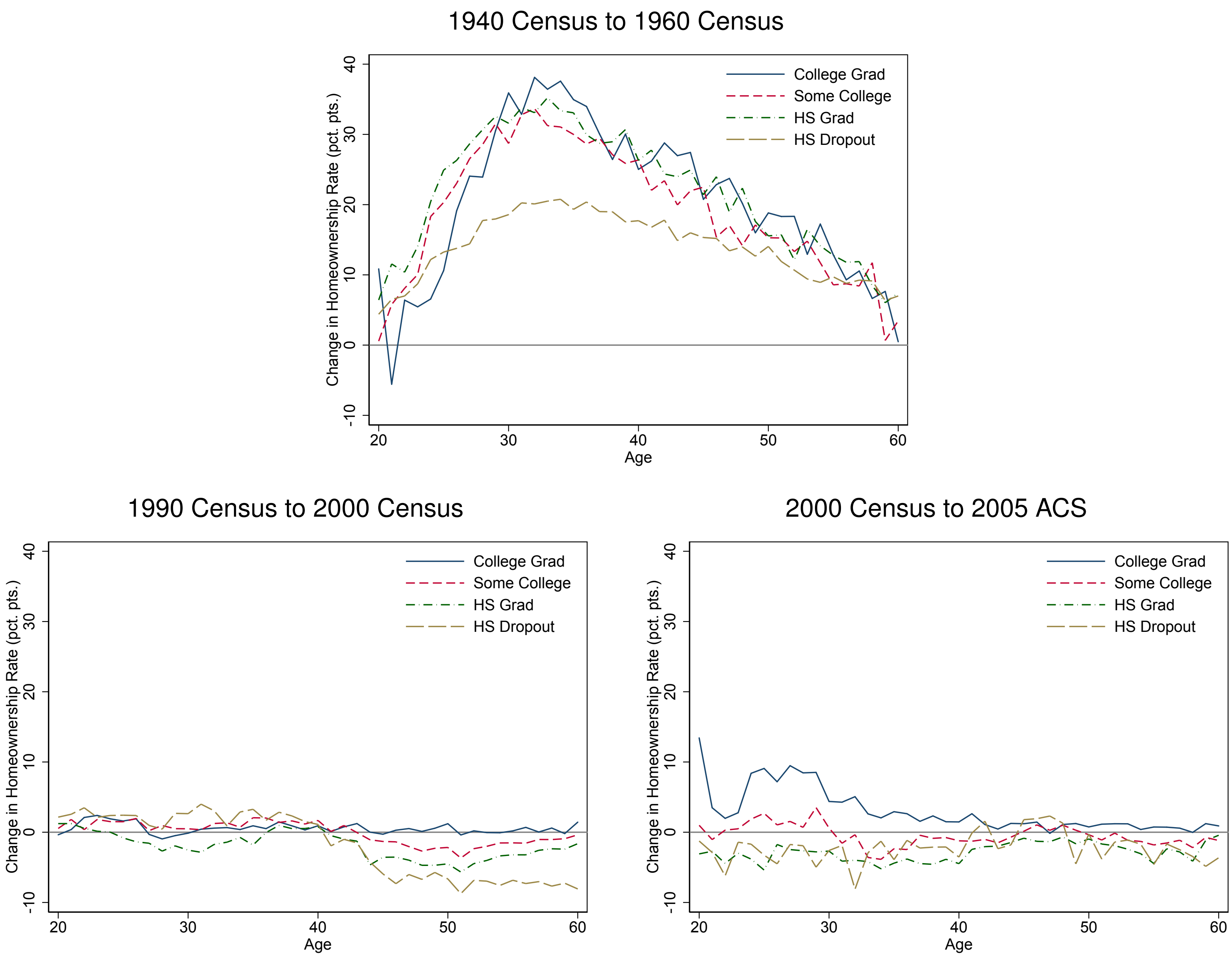

Figure 13. Changes in U.S. Homeownership in Three Historical Periods
Note: The top panel graphs the change in homeownership rates for households with heads of various ages and educational attainment from 1940 to 1960, using individual-level Census data (which are not available for 1950). The lower panels construct analogous graphs displaying homeownership changes between the 1990 and 2000 decennial Censuses (lower left panel) and the 2000 census and the 2005 ACS (lower right panel).
Source: 1940, 1960, 1990 and 2000 Decennial Censuses and 2005 ACS (from IPUMS: Ruggles et al. (2018)).

AHS Surveys	Initial Units	Remaining Units	Total Losses: Number	Total Losses: As % of Init. Units	Conversions to Rented or Vacant: Number	Conversions to Rented or Vacant: As % of Init. Units	Losses Other than Conversions to Rented or Vacant: Mobile Homes Moved Out	Demolitions & Disasters	Badly Damaged & Condemned	Conversions to Nonresidential	Units Lost in Other Ways	Total Number	As % of Init. Units
1985-87	56,766	52,162	4,604	8.1%	3,781	6.7%	528	100	45	37	113	823	1.4%
1987-89	58,746	54,232	4,514	7.7%	3,801	6.5%	454	114	57	35	53	713	1.2%
1989-91	59,764	55,211	4,553	7.6%	3,924	6.6%	385	95	37	38	74	629	1.1%
1991-93	59,580	55,198	4,382	7.4%	3,661	6.1%	467	68	38	23	125	721	1.2%
1993-95	60,999	55,907	5,092	8.3%	4,392	7.2%	472	77	40	33	78	700	1.1%
1995-97	63,314	58,016	5,298	8.4%	4,544	7.2%	405	80	48	34	187	754	1.2%
1997-99	65,396	60,292	5,104	7.8%	4,434	6.8%	361	135	52	35	87	670	1.0%
1999-01	68,712	62,787	5,925	8.6%	5,092	7.4%	412	86	50	38	247	833	1.2%
2001-03	71,708	65,558	6,150	8.6%	5,625	7.8%	116	97	55	23	234	525	0.7%
2003-05	72,238	66,061	6,177	8.6%	5,716	7.9%	80	96	44	28	213	461	0.6%
2005-07	74,931	67,620	7,311	9.8%	6,759	9.0%	174	181	68	25	104	552	0.7%
2007-09	75,647	68,551	7,096	9.4%	6,642	8.8%	124	96	40	57	137	454	0.6%
2009-11	76,428	68,281	8,147	10.7%	7,722	10.1%	109	145	29	26	116	425	0.6%
2011-13	76,092	69,324	6,768	8.9%	6,418	8.4%	83	116	26	14	111	350	0.5%

Table 3. Losses of Owner-Occupied Housing Units between Adjacent American Housing Surveys

Note: Numbers are in thousands. This table displays statistics related to losses of owner-occupied housing units that are derived from the U.S. Department and Urban Development's *Components of Inventory Change* reports, which are in turn based on comparisons of data from adjacent American Housing Surveys. The column labelled "Initial Units" lists the total number of owner-occupied units in the AHS in the first year of a two-year period (for example, in the first row, the initial year is 1985). The column titled "Remaining Units" lists the number of those initial units that remain the next AHS survey (for the first row, this year is 1987). The column titled "Units Lost in Other Ways" includes the net losses from mergers or conversions of owner-occupied units into units of different sizes, as well as a statistical discrepancy that totals no more more than 4,000 units in any two-year period.

Source: Authors' calculations from various issues of the *Components of Inventory Change* reports.

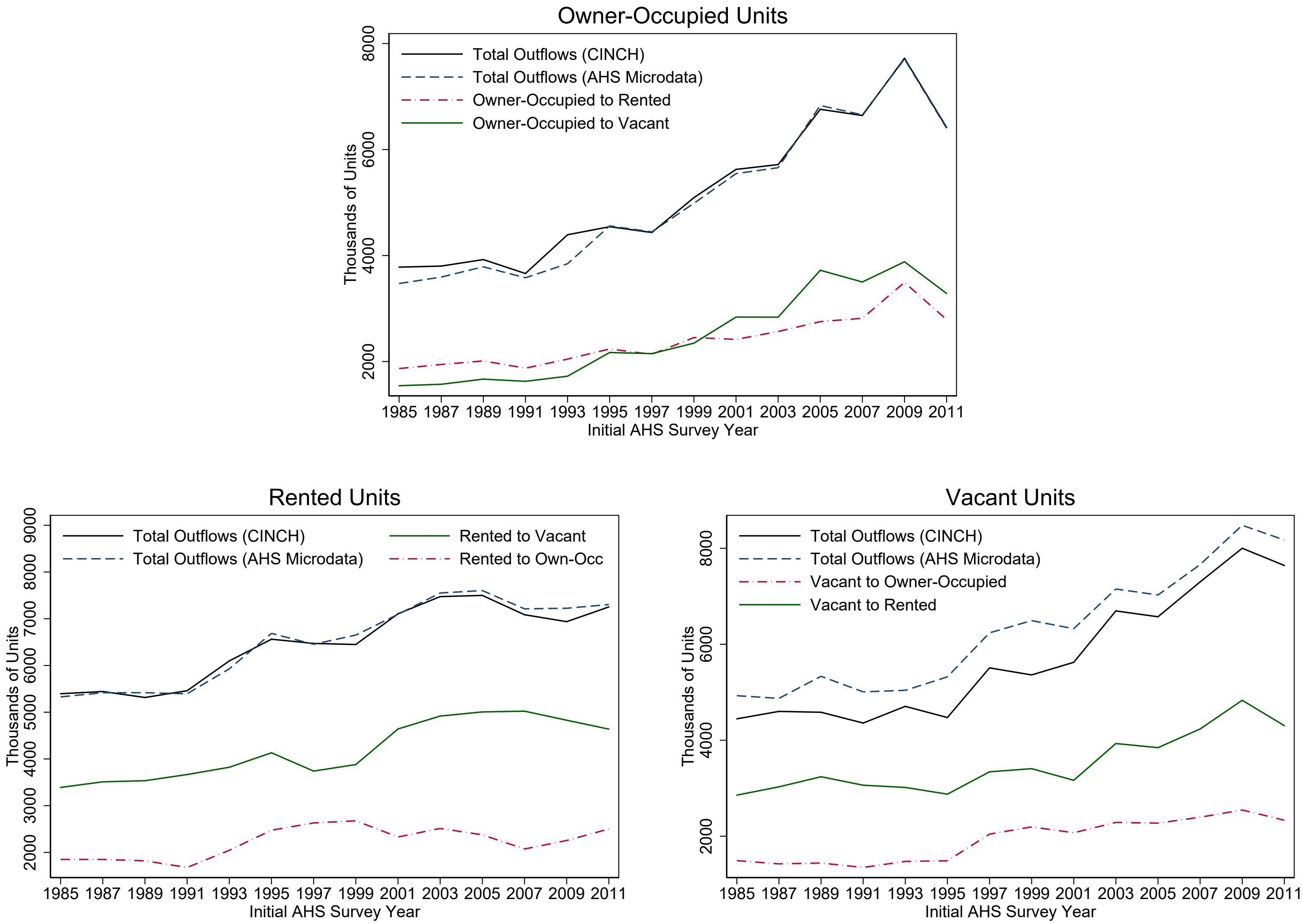

Figure 14. Gross Flows of Housing Units out of Owner-Occupied, Rented, and Vacant Status
Note: The black solid lines in each panel display total outflows according to the Components of Inventory Change (CINCH) reports, while the dashed blue lines depicts our estimates of these series. The estimates are not in perfect agreement, due in part to our use of beginning-year weights only when constructing flows; CINCH combines a sample properties weights from both the beginning and ending years.
Source: American Housing Survey and various issues of the *Components of Inventory Change* reports.

Period	Gross Flow: Rented to Owner-Occupied	Gross Flow: Owner-Occupied to Rented	Difference: Net Flow from Rented to Owner-Occupied
1985-87	925	934	–9
1987-89	925	973	–48
1989-91	911	1,006	–95
1991-93	839	938	–98
1993-95	1,023	1,023	0
1995-97	1,238	1,119	119
1997-99	1,316	1,068	247
1999-01	1,339	1,227	112
2001-03	1,165	1,209	–43
2003-05	1,257	1,283	-26
2005-07	1,189	1,377	–188
2007-09	1,037	1,408	–371
2009-11	1,128	1,747	–619
2011-13	1,252	1,396	–144

Table 4. Gross Flows of Occupied Housing Units between Rented and Owned Status
Note: Numbers are in thousands. Differences are not always exact because of rounding.
Source: American Housing Survey and Bureau of the Census.

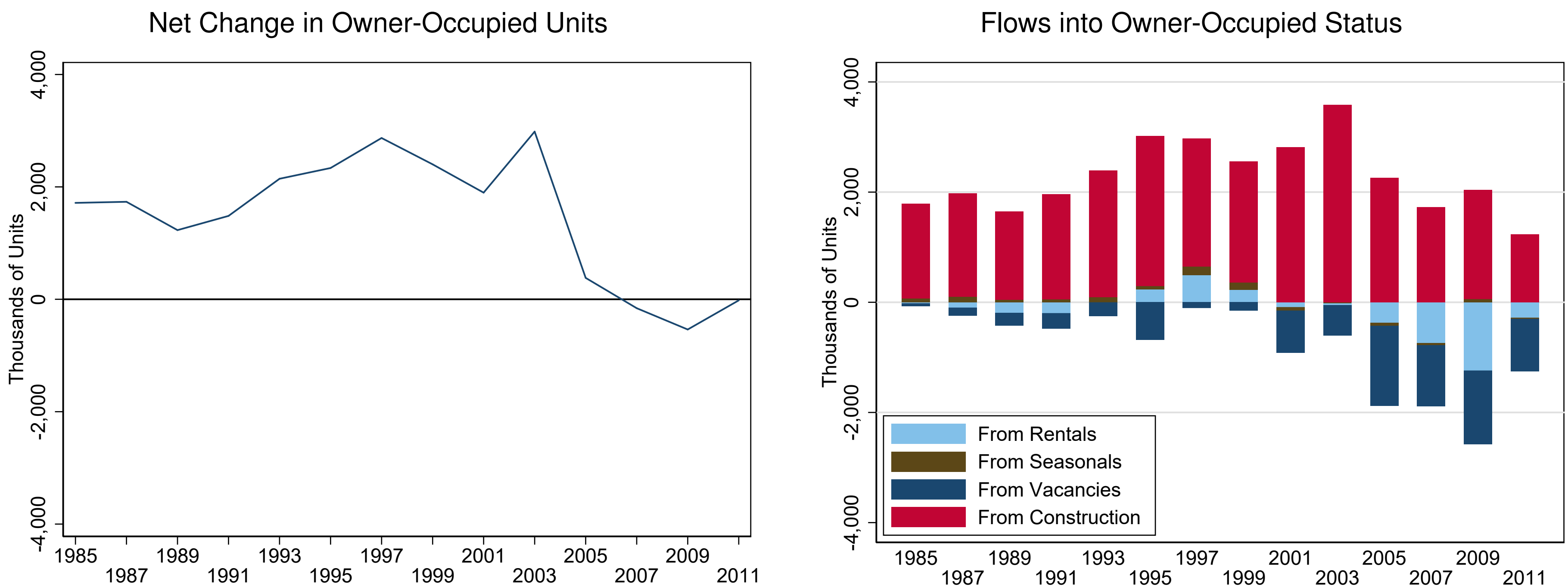

Figure 15. Net and Gross Flows of Owner-Occupied Homes
Note: The left panel depicts the 2-year change in the number of owner-occupied homes according to the Census Bureau; for example, the change from 1985 to 1987 appears as the 1985 data point. The Census series has been adjusted for a break in 2000 arising from the use of weights from the 2000 decennial census beginning in that year. In the right panel, the height of the stacked bar above the y-axis, less the height of the bar below that axis, equals the net change in owner-occupied units shown in the left panel. The light blue bars in the right panel are based on AHS microdata, and depict the net flow of existing homes from rented to owner-occupied status—that is, the gross flow of rented to owner-occupied homes less the gross flow of owner-occupied to rented homes. Contributions arising from seasonal and vacant properties (brown and dark blue bars) are also based on AHS microdata and are calculated analogously. The construction contribution in the right panel (red bars) is calculated as the change in owner-occupied units from the left panel less the combined contributions from rented properties, seasonal properties, and vacancies shown in the right panel. The construction contribution therefore reflects the gross flow of newly constructed owner-occupied properties less the losses of such properties from destruction and conversions.
Source: American Housing Survey microdata, US Bureau of the Census, and Haver Analytics.

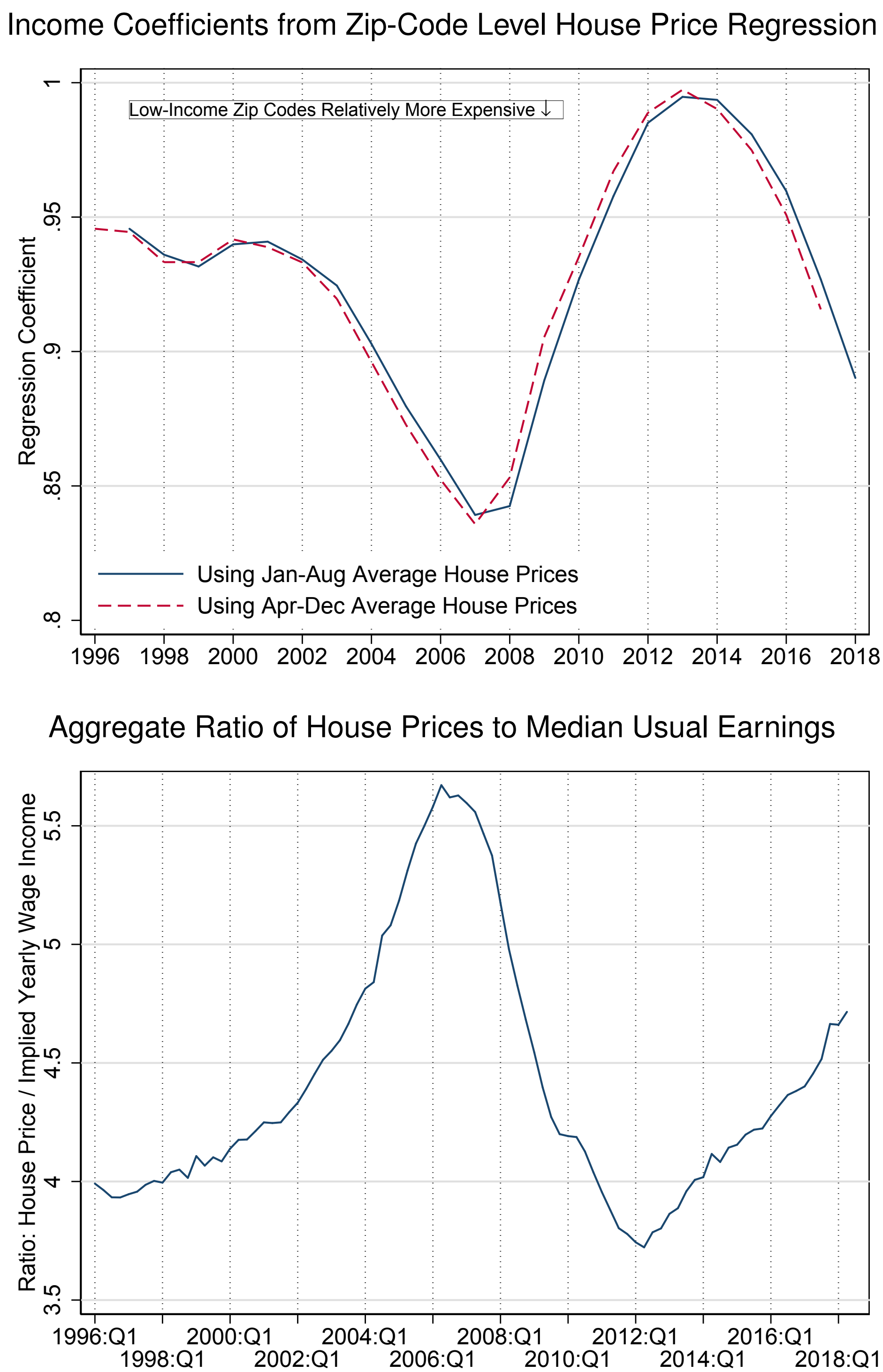

Figure 16. Effects of Income on House Prices at the Zip-Code Level and the Aggregate Ratio of House Prices to Income

Note: The upper panel graphs yearly income coefficients from a regression of the log of zip-code level house prices on the log of median income for zip codes from the 2000 Census, interacted with year. The lower panel graphs the national Zillow house-price estimate divided by 52 times the median weekly earnings for full-time wage and salary workers in the CPS.

Source: Zillow, US Bureau of the Census, and US Bureau of Labor Statistics.

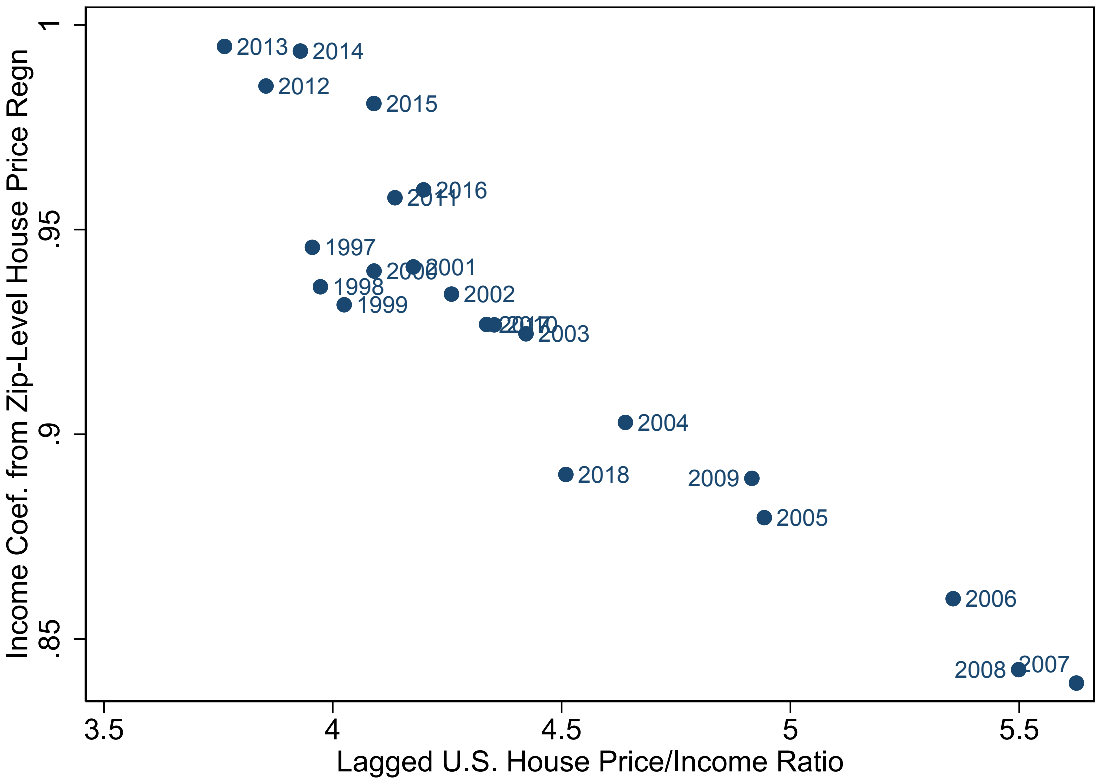

Figure 17. The Relationship between Income Coefficients from the Zip-Code Level House Price Regression and the Lagged Aggregate House Price–Income Ratio
Note: The vertical axis of this graph corresponds to the income coefficients in the upper panel of Figure 16, and the horizontal axis corresponds to lagged values of the price-income ratio in the lower panel of this figure.
Source: Zillow, US Bureau of the Census, and US Bureau of Labor Statistics.

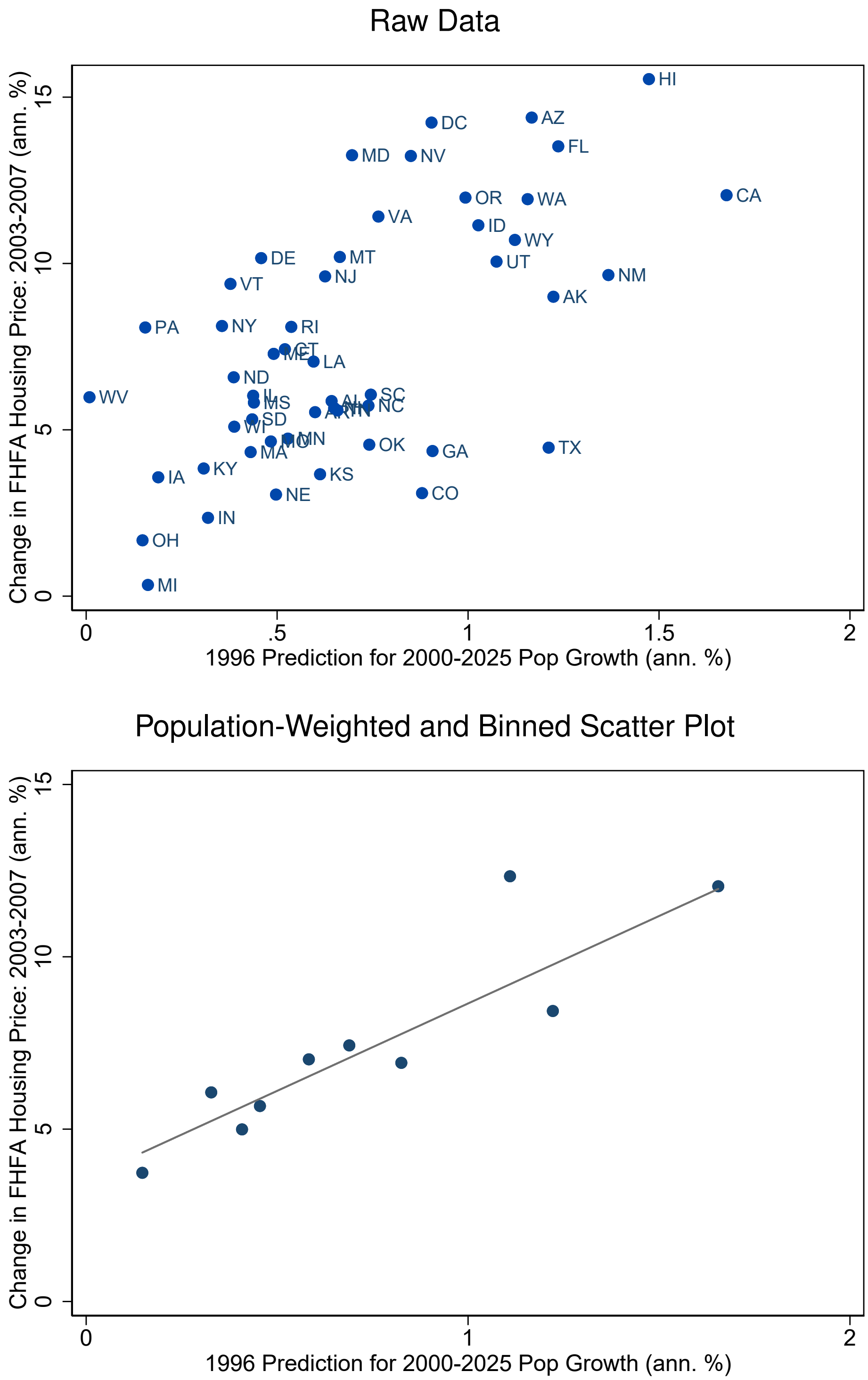

Figure 18. Expected Population Growth and House-Price Increases at the State Level
Note: In the top panel, the horizontal axis depicts the annualized growth rate of state-level population from 2000 to 2025 that was expected by the US Census Bureau as of 1996. The vertical axis of this panel plots the annualized growth rate of house prices from 2003–2007 according to the Federal Housing Finance Agency. The lower panel is a binned scatter plot of the same data that is also weighted by actual 2004 population levels.
Source: Campbell (1997) and Federal Housing Finance Agency.

A Internet Appendix

A.1 Identifying Piggyback Loans in HMDA

The procedure for identifying piggyback loans in HMDA involves multiple stages. Each stage involves the identification of some piggyback loans and after each stage, those piggyback loans and their associated first liens are removed from consideration.

The first stage uses loan-level characteristics available in HMDA to identify duplicate observations. Of the two observations that comprise a duplicate, we assume that the larger one in terms of loan value is the first lien and that the smaller is the piggyback loan. The loan-level characteristics that we used to identify these duplicates are as follows:

1. The banking institution originating the loan.
2. The week the loan was originated.
3. The month the loan was originated.
4. The census tract in which the property is located.
5. The loan type (whether it is conventional, or guaranteed by the Federal Housing Authoris (FHA), Veterans Administration (VA), Farm Service Agency (FSA) or the Rural Housing Service (RHS)).
6. Whether or not the borrower will be occupying the property.
7. The income, race, and sex of the borrower.

The second identification stage is similar to the first, but in place of some of the loan-level characteristics, this stage assumes a specific ratio between the origination amount of the first and second lien. First, we assume that the first lien is four times the value of the second lien (an 80–20) ratio, the most common ratio seen of first liens to their associated piggybacks. Then we assume that the ratio is 80–10, 80–5, and finally 80–15, which are also common ratios seen in the data. The loan level characteristics used are the month of origination, the census tract of the property, and the income of the borrower.

The third stage is more ad hoc and identifies piggyback loans by using duplicates in a variety of loan-level variables.

1. Origination date, application date, census tract of the property, income of the borrower, and whether the borrower indicates that they plan to occupy the property.

2. Origination date, application date, the census tract of the property, the income of the borrower rounded to the nearest $10,000, whether the borrower plans to occupy the property, the borrower's sex and race, the co-applicant's sex and race, and whether the loan was conventional or guaranteed by one of the federal authorities listed above.

3. Origination date, census tract of the property, whether the borrower is an owner-occupant, the banking institution that made the loans, and the income, race, and sex of the borrower.

4. Origination date, the banking institution that made the loans, the census tract of the property, whether the borrower intends to occupy the proerty, whether the loan was conventional or guaranteed by one of the federal authorities listed above, the race, sex, and income of the borrower, and the race and sex of the co-applicant.

5. Origination date, the banking institution that made the loan, the week the loan was applied for, the census tract of the property, whether the borrower intends to occupy the property, whether the loan was conventional or guaranteed by one of the federal authorities listed above, the income of the borrower rounded to the nearest $5,000, the race and sex of the borrower and co-borrower.

6. Origination date, the week of application, the census tract of the property, whether the borrower intends to occupy the property, whether the loan was conventional or guaranteed by one of the federal authorities listed above, the income, race, and sex of the borrower, and the race of the co-borrower.

As a last step, we use county-level median house prices from Zillow, and code any loan that is less than 0.01 percent of the value of the median house price in the county as a piggyback loan. We are not able to identify the associated first liens to these loans, so they are removed from any analysis of individual mortgages.

Starting in 2004, HMDA identifies the lien-type of all reported loans, but does not link associated first and second liens. To assess the quality of our identification of piggyback loans, we compare our results to the HMDA identified second liens in 2004. Of loans that HMDA says are first liens, we identify over 99 percent of them correctly as first liens, and of loans HMDA codes as second liens, we correctly identify 84.8 percent as second liens. Therefore, we falsely identify 15.2 percent of HMDA-coded second liens as first liens, and 0.9 percent of HMDA-coded first liens as second liens. Since there are many more first liens than piggyback loans, the number of falsely identified second liens is a little under half the number of falsely identified second liens. Unfortunately, these errors do not cancel each other out, but it does mean that overall we correctly identify 97 percent of the loans. this rate holds for all the years following 2004 as well. We also checked to see if our procedure was more accurate in higher or lower income areas. Using 2000 census tract income, in 2004, we correctly identify 97.8 percent of the loans in the lowest quartile of census tracts by median income, and 96.9 percent of the loans in the highest quartile of census tracts.

Figure A.1 compares the results of regressions of individual loan amounts on income using only first liens identified in HMDA starting in 2005, our correction using combined loan amounts from matched first and second liens, to regressions that include first and second liens as separate observations. It is clear that identifying the second liens has a significant effect. It is less important whether or not the piggyback loans are included in total loan amount.

A.2 Robustness Checks for Debt-Income Regressions

Figures A.2 and A.3 display results from alternative specifications of the debt-income regressions using either HMDA or AHS data. Recall that our baseline HMDA specification instruments individual-level HMDA income with tract-level income from the US Census or ACS, and also includes CBSA-year fixed effects. The top left panel of Figure A.2 displays the income coefficients from a HMDA regression in which individual-level income is not instrumented and no CBSA-year fixed effects are included. The middle panel of the top row shows the expected mortgage amounts from that regression. The top right panel of Figure A.2 shows income coefficients from a non-instrumented regression that does include the CBSA-year fixed effects.

Each of the panels in the lower row comes from a regression for which tract-level income serves as an IV for individual-level income. The first two panels of the lower row display the income coefficients and the expected mortgage amounts from an instrumented regression without CBSA-year fixed effects. The bottom right panel in the lower row depicts income coefficients from an IV regression that includes CBSA-year fixed effects. The last two panels of the lower row are therefore identical to the panels in Figure 8.

Figure A.3 presents alternative versions of the AHS regressions. The top left panel displays the coefficients using the baseline sample and fixed effects for census region $\times$ year. The expected mortgage amounts from this regression are graphed in the bottom left panel. The top right panel displays the income coefficients without using the 5-percent truncation rule for income and mortgage debt and with the region-year FEs. The middle left panel displays results from a quantile regression using the baseline sample that includes region-year FEs, and the middle right panel omits the region-year FEs. The lower right panel shows the expected mortgage amounts from this regression. The panels that appear in Figure 9 are thus the middle right and lower right panels.

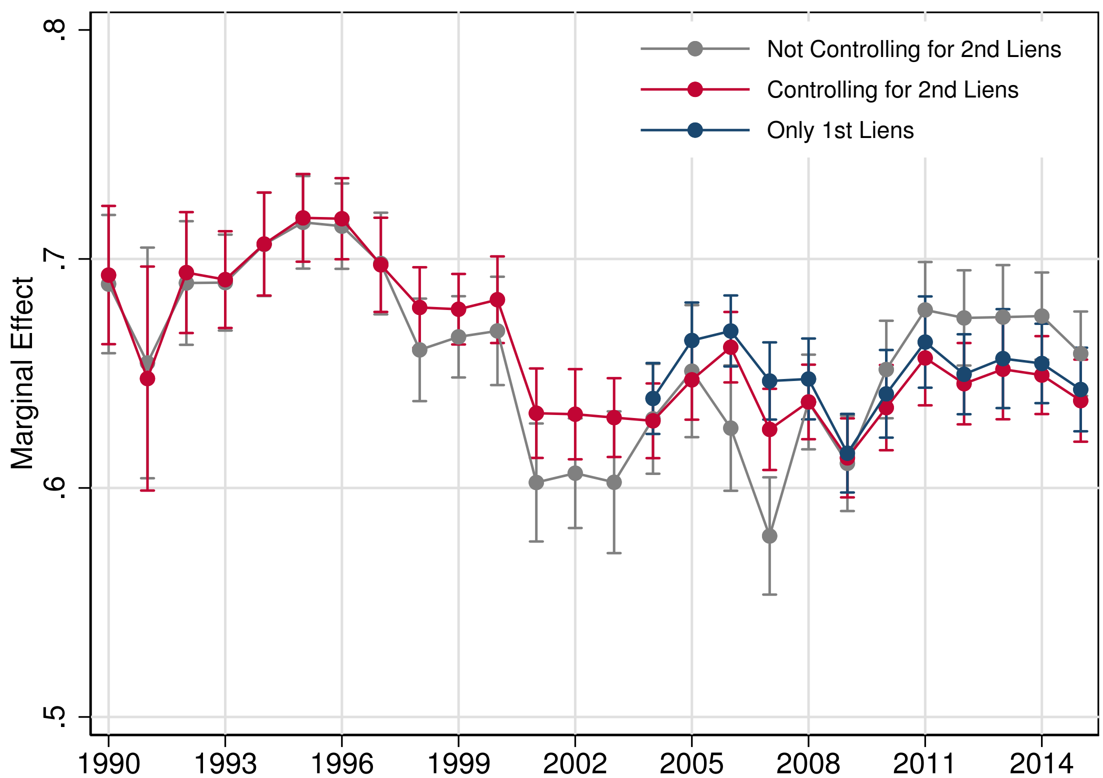

Figure A.1. Comparison of Author-Identified First Liens and those Identified by HMDA
Note: This graph plots coefficients from regressions of individual loan amounts on income and covariates from HMDA data. The blue line plots the coefficients only using first liens as identified by HMDA. The grey line does not make any correction for second liens. The red line uses the authors' algorithm to identify second liens back to 1990, and corrects the loan amount to account for both mortgages.
Source: HMDA.

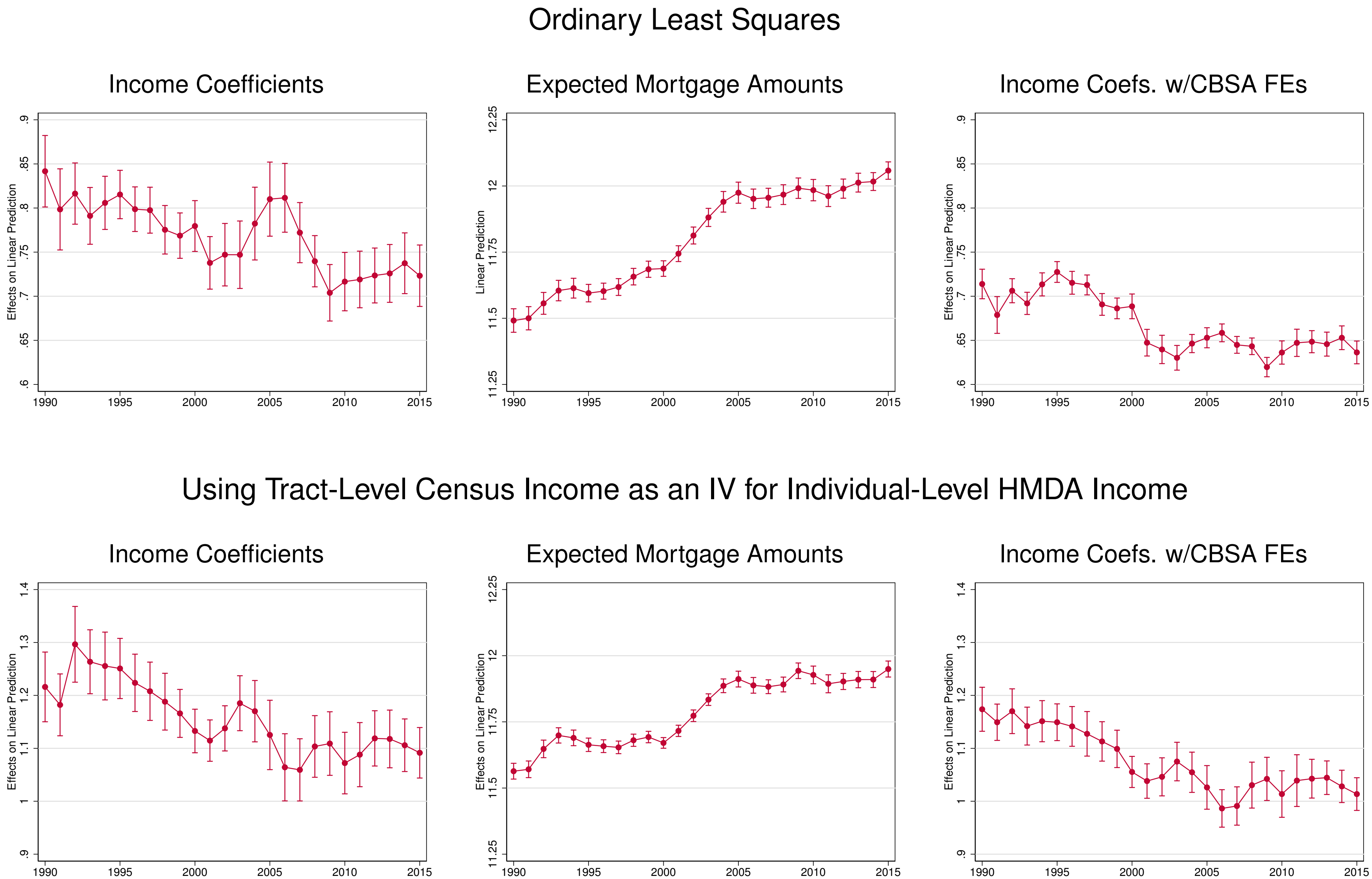

Figure A.2. Canonical Regression using Individual-Level HMDA Mortgage Balances and Income Levels
Note: These panels graph income coefficients (and 95 percent confidence intervals) from regressions of individual purchase mortgage origination amounts from HMDA on measures of income. The top panel uses individual income as reported in HMDA. The bottom panel instruments for HMDA income using the most recently available Census tract income from the Census and ACS. Both specifications include CBSA by year fixed effects and control for the borrowers race and gender interacted with year. Expected mortgage amounts are predictions, holding income constant at its average value across all years.
Source: HMDA, decennial census, and the American Community Survey.

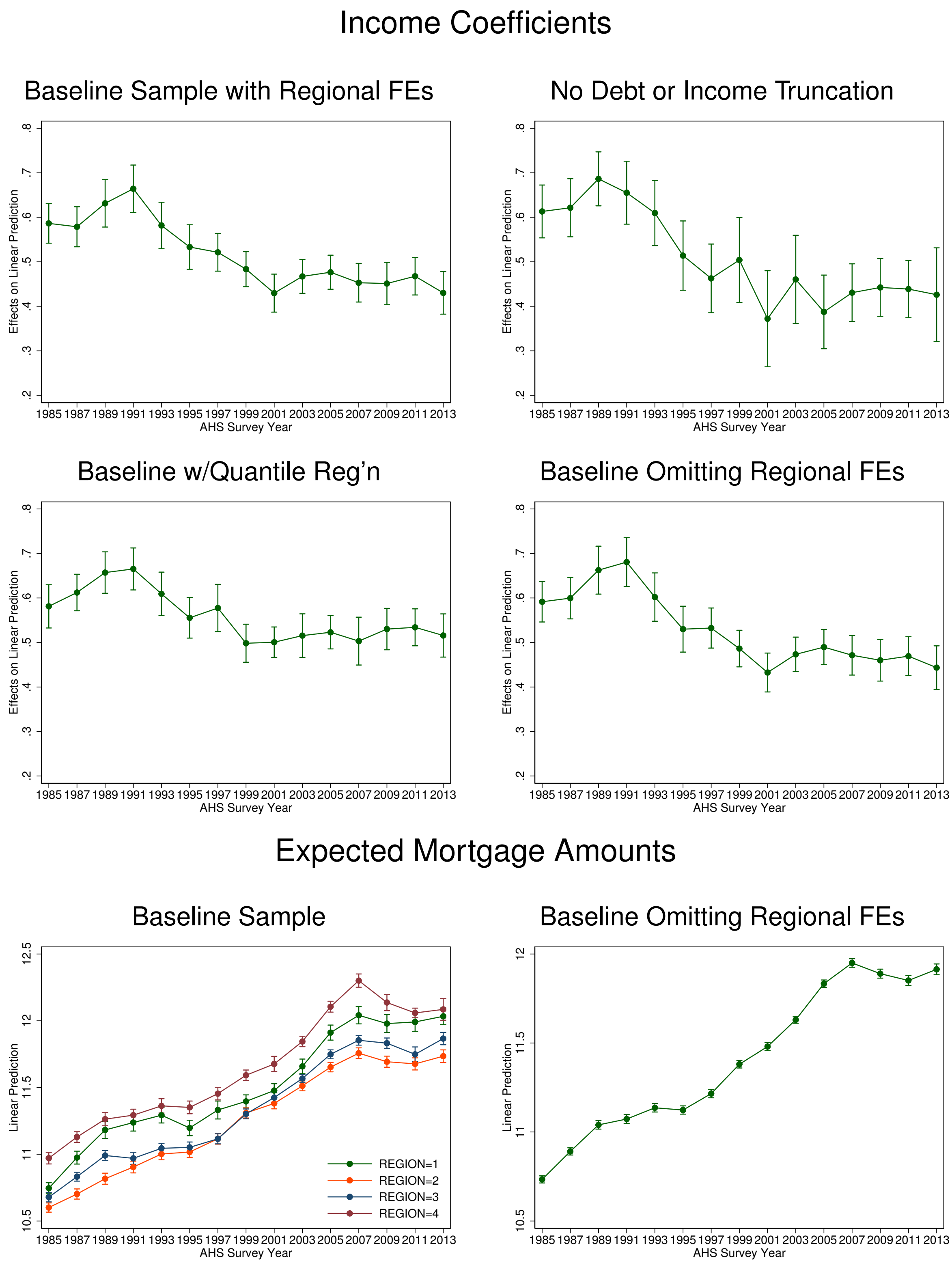

Figure A.3. Canonical Regression using Individual-Level AHS Mortgage Balances and Income Levels
Note: See the notes to Table 9 for details of these regressions.
Source: AHS.