Technological Innovation in Mortgage Underwriting*

Christopher L. Foote[†]   Lara Loewenstein[‡]   Paul S. Willen[§]

October 5, 2018

> **Preliminary Draft: Please do not quote or cite without permission.**

Abstract

The 1990s stand out as a period of rapid technical change in mortgage underwriting. Advances in computerization and information technology allowed lenders to rely on default predictions from increasingly sophisticated empirical models. Debt-to-income ratios at origination add little to the predictive power of these models. As a result, the automated underwriting systems allow higher debt-to-income ratios than previous human underwriters would typically accept. As economic theory would predict, the relaxed standards for income in mortgage lending helped raise the homeownership rate modestly during the late 1990s, but had a small effect on housing prices. However, the relaxation of income constraints engendered by automated underwriting may have allowed the 2000s housing boom to grow, because these systems allowed borrowers throughout the income distribution to act on distorted beliefs about future house-price growth by taking out increasingly larger loans.

At first glance, Louise Beyler of Gainsville, Georgia might appear as an unlikely candidate to buy a $105,000 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. 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

An ironic feature of the catastrophic housing cycle of the 2000s was that it followed a decade of rapid technical change, when mortgage lending was widely thought to have improved. During the 1990s mortgage lenders leveraged rapid gains in data-storage and information-processing capabilities to originate mortgages faster and at less cost than before. These technical advances were a central focus of end-of-decade retrospectives on mortgage lending, which not only touted the 1990s gains but also looked forward to additional progress during the 2000s. Of course, taken as a whole the 2000s did not turn out to be a particularly good decade for the mortgage industry. In the wake of the housing crisis, a number of new government regulations have sought to remedy problems in the mortgage-origination process, for example by requiring that lenders ensure that borrowers have the ability to repay their loans. In this paper, we study how the ostensibly positive changes in mortgage origination during the 1990s can be reconciled with the significant problems in mortgage finance during the decade that followed. We have four main findings.

First, we use new data and new empirical approaches to confirm that the people who wrote about technical change in the mortgage industry during the 1990s knew what they were talking about. During that decade computerization did allow originators to deliver mortgages more efficiently. Advances in information technology also enhanced the ability of mortgage lenders to compete geographically, with the share of mortgages originated by out-of-state lenders rising from around 10–20% in 1990 to around 60% by the end of the decade. Some of these improved metrics have reversed course recently, as lenders adapt to the new lending regulations intended to prevent another housing crisis. For example, the time required to close a standard refinance application fell from around two months in the early 1990s to less then 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 today. A qualitatively similar pattern is present for the financial costs of mortgage origination that are borne by borrowers; large declines during the 1990s have been partially reversed after

1

the housing crisis.

The second finding of the paper concerns a change that has not been reversed: technology has fundamentally altered the way that mortgage lenders evaluate risk. Before 1990, loan officers and originators evaluated mortgage applications by applying so-called knockout rules, which specified maximal cutoffs for LTV (loan-to-value) ratios, DTI (debt-to-income) ratios, and other variables. One would think that a rules-based system such as this would be tailor-made for computers, which have transformed large sections of the U.S. economy because they can follow sequential instructions and perform routine tasks efficiently. In fact, as the 1990s began many lenders tried to exploit computerization in precisely this way, by coding up their lending rules into formal algorithms. The idea was that the artificial intelligence (AI) systems would then evaluate loans just as humans always had, but at much lower cost.

The coders soon discovered that intuitive human judgment was in fact central to mortgage lending. Decisions were easy for humans when a loan application cleared all of the relevant hurdles, for example by having low DTI and LTV ratios. But for marginal cases the underwriter had to use his own experience and judgment to decide whether, say, a high DTI that a requested could be offset by a low LTV ratio or 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 under-writers 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).

A big problem with basing lending decisions on codified human experience is that mortgage defaults are rare. Individual underwriters are therefore unlikely to accumulate enough personal experience with defaults to properly quantify the trade-offs among a loan's risk characteristics.

Others in the mortgage industry chose to use computers differently. They pulled together large datasets of loan-level records to estimate empirical models of default, which could then be used to supplant human decision-making in lending decisions. These modellers realized that the predictions from empirical default models could be significantly enhanced by the use of credit scores, including the FICO score, that were then being developed to predict defaults on consumer debt. Thus, the same technologies that revolutionized the credit-card

industry and other aspects of consumer lenders during the 1990s became integrated into mortgage lending as well.

In addition to finding that credit scores helped predict mortgage default, the 1990s-era mortgage modellers found that DTI ratios did not. As we explain below, modern theories of mortgage default—which are based on income *shocks* and not initial income *levels*—are quite consistent with this result. The relative unimportance of origination DTI for mortgage default had also been foreshadowed in previous, small-scale studies of mortgage default, including one sponsored by the National Bureau of Economic Research in the late 1960s (Herzog and Earley, 1970). In the 1990s, however, the unimportance of DTI at origination began to influence lending decisions. Informed by the findings of empirical default models, during the 1990s mortgage lenders weakened the relationship between a borrower's income and the amount of mortgage debt that he was allowed to assume.

A third contribution of the paper exploits the fact that the 1990s changes in lending standards were plausibly exogenous with respect to the state of the housing market. The empirically based AU systems were developed in the 1990s not because the housing market was particularly hot or cold at that time, but rather because computerization had advanced to the point that estimation of large-scale predictive models became feasible. By and large, we find that effects of these exogenous changes are essentially those that are predicted by theory. For example, relaxing DTI limits should have the largest effects on individuals whose future income is high relative to their current incomes. Young college graduates have steep age-earnings profiles, and we find that the increase in homeownership rates after 1994 were particularly large for that group. In this respect, the credit expansion of the 1990s played out as a miniature version of the much larger expansion in mortgage lending between 1940 and 1960. As shown by Fetter (2013), the mid-century credit expansion driven by the Veterans' Administration (VA) and Federal Housing Administration (FHA) also had its largest effect on the young. Fetter (2013) argues that this credit expansion allowed individuals who would have purchased homes eventually to buy them at an earlier stage in the life cycle. We find the same for the 1990s.

We then investigate the link between exogenous changes in lending standards and housing prices. As pointed out by Kaplan, Mitman, and Violante (2017) and others, an exogenous change in lending standards raises the potential demand for owner-occupied homes, as opposed to rented homes. The increase in owner-occupied demand can be satisfied by new construction or by the conversion of rental units into owner-occupied ones. To the extent that the higher owner-occupied demand is satisfied by these margins, then relaxed lending standards will not raise owner-occupied prices very much. To investigate these potential supply margins, we use the American Housing Survey (AHS), which is also used in our study of mortgage debt and income. Because the AHS follows a panel of the same housing units

over time, we can track the evolution of tenure status for a given unit over time. When we construct "gross flows" of properties in and out owner-occupied status, we find that the net flow of rented to owned properties, which is usually negative, becomes positive in the 1990s, consistent with theory. There is now such switch in the 2000s, nor do we find that the homeowner vacancy rate declines as house price appreciation picks up. We view these findings as evidence that gross flows can offset the effect of exogenous lending standards on prices. These flows provide no support for an exogenous change in standards during the 2000s boom.

Finally, our fourth point concerns the central 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 reducing the importance of human-decision making in lending led to more efficient predictions of default. The success of model-based lending explains the continued development of "fintech," which includes the current development of sophisticated machine-learning algorithms to predict default. But improving lending decisions at the individual level can lead to bad macro outcomes if beliefs about macro variables are distorted. The central purpose of any financial system is to match people who want to lend to 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 to them. New technologies facilitated these transactions, and the downgraded importance of current income in the lending decision weakened any depressing effect on lending that human underwriters, using old DTI ratios, would have exerted. In this way, the more-efficient mortgage lending system developed in the 1990s could have played an important part in the housing boom by allowing individuals to act on their overly optimistic believes about further house-price gains more easily.

The remainder of this paper is organized as follows. Section 2 discusses the two main data sources we use to quantify productivity improvements in mortgage lending and to show how these changes influenced the relationship between income and mortgage debt. Section 3 quantifies the significant improvements in mortgage lending during the 1990s, and section 4 studies the evolution of the debt-income relationship. Section 5 provides a historical review outlining the ways that credit scores and automated underwriting models enhanced mortgage lending by improving the evaluation of default risk. Section 6 traces out the effects of such changes on the homeownership rate and on housing prices, with particular attention to how gross flows among rented and owned status affects price effects. In the conclusion, we discuss how and why a more-efficient system of mortgage lending can be reconciled with a damaging housing cycle, as long as this cycle is driven by distorted beliefs about housing prices.

2 Data

2.1 The Home Mortgage Disclosure Act (HMDA)

Our first dataset comes from the Home Mortgage Disclosure Act (HMDA), a 1975 law requiring 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 relevant information is not available in HMDA on a consistent basis due to changes to reporting guidelines, including a relatively major change in 2004.[1] Information on the lien status of mortgages—that is, whether the mortgage is a first or second lien— becomes available starting in 2004. Because it was common during the 2000s boom for first-lien mortgages to be supplemented by second liens at purchase ("piggybacks"), it is important to account for second liens consistently when making comparisons over many years. We therefore created an algorithm to identify second liens before 2004, and then validated this rule using the reported liens available beginning in that year. 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.[2]

The HMDA data also include a field that can separate owner-occupiers from investors.[3] 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. Additionally, investors complicate any analysis of income and debt at the zip-code level, because investors may not live in the zip code where the investment property is located.

We then clean the HMDA data by following advice from Avery, Brevoort, and Canner (2007). We drop loans that lack information on race and gender, as these are probably business loans. Whenever applicable we combine home improvement loans with refinances. We also remove outliers by calculating the implied monthly payment of each loan, assuming that it is a 30-year fixed-rate mortgage at the current interest rate. We then calculate the implied DTI ratio (that is, the ratio of the implied monthly payment to the borrower's implied

---

[1]For details of the changes that went into effect on January 1, 2004, see the article "HMDA changes are on the way; new rules take effect in 2004" in Community Dividend from the Federal Reserve Bank of Minneapolis.

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

[3]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.

monthly income), and drop loans with ratios in the bottom and top 1% of the distribution.

Figure 1 depicts some summary data from purchase-mortgage originations in HMDA. The top panel shows the total number of purchase loans originated and the fraction of those loans with associated second liens. The number of purchase loans originated increases steadily from 1990 until 2005, while the use of piggyback loans grows dramatically near the end of the boom. After 2007, both series drop sharply. The bottom panel displays the median total loan amount (reflecting the sum of first liens and associated piggybacks) for owner-occupied purchase mortgages. The panel also shows the share of purchase mortgages made to owner-occupiers. The owner-occupier share declines by almost 10 percentage points during the boom. There is also hump shape in the median loan amount at the height of the boom, when the use of piggyback loans was most prominent and the investor share or purchase mortgages was highest.

A potential problem with the income information in HMDA is that it reflects what the mortgage lender verified for the purpose of the loan, and is not necessarily total household 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 he and his mortgage broker may not go to the trouble of documenting the stability of income that is not needed for the loan.[4] Consequently, when house prices rise, incomes reported on HMDA records may also rise, as borrowers require additional income sources to qualify for the larger loans. Individual-level income as reported on HMDA records may therefore be 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.[5] In light of discrepancies in HMDA incomes, in our main regression specification we instrument for HMDA income using the median household income by census tract from the decennial census and ACS.

---

[4]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."

[5]Some evidence on the potential endogeneity of HMDA incomes comes from a comparison of HMDA incomes to Census income data. The decennial 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 HMDA incomes were up to 30 percent higher than those reported in the ACS in 2005 and 2006 in certain states: specifically, 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 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.

2.2 The American Housing Survey

Another source of both income and mortgage-debt data is the American Housing Survey (AHS), which began in 1973 and has been conducted in odd-numbered years since 1981. We use data after a significant redesign in 1985.[6] The AHS is a joint project of the HUD and Census designed to measure the size, quality, and composition of the U.S. housing stock. The survey also measures monthly housing costs for U.S. residents and collects demographic information and income on 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 mortgages currently on the house *at the time those mortgages were originated*.[7] We use the origination amounts of current mortgages for recent movers as a measure of new mortgage debt over time. For income, the AHS includes data on both salary and wage income as well as total income. Because these income measures come from surveys, not from mortgage applications, they are much less likely to be influenced by housing prices.

Table 1 displays unweighted sample counts from the AHS. While our main regressions focus only on households that moved since the previous survey, we still have a sizable sample of households. Unless otherwise noted in our empirical analysis we use the same of households that moved since the previous survey. This sample ranges in size from just over 3,000 households to over 7,000 households per survey year.

Figure 2 compares median levels of new mortgage debt and income from HMDA and the AHS. 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 HMDA incomes rise relative to AHS income during the height of the boom. This pattern confirms the lessons from Avery et al. (2012) that HMDA income may be overstate true income in areas or periods where house prices are rising rapidly.

3   Quantifying Technical Progress in Mortgage Lending

In this section, we quantify three areas in which technological improvements in mortgage lending are reflected.

---

[6]For a general description of the AHS see U.S. Department of Housing & 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).

[7]Although there are no questions asking whether the current mortgage on the home is actual purchase mortgage, for very recent movers this is likely to be the case. When it isn't, then the origination amount of the current mortgage will be still close to the size of the origination amount of the purchase mortgage, absent large amounts of cash-out refinancing soon after the house was purchased. As a further check, we can use a field in the AHS to verify that we only consider current mortgages that were taken out in the same year that the house was purchased.

3.1 Cross-State Lending

Since 1990 lenders have been increasingly able to handle loan applications from areas outside of their normal markets. We calculate what percentage of loans were originated by lenders on properties outside of states where they have a physical presence. We use HMDA data matched with information on each lender's parent company and/or bank holding company from the National Information Center[8] and bank branches from the FDIC's Summary of Deposits (SOD). Mortgage companies or credit unions are excluded from this analysis. [9] This analysis also does not include locations of loan production offices (LPs), which are office locations that process loans. However, the loans are approved and funds are dispersed by other branches or offices in the bank. For this reason, a certain level of technology is necessary for LP offices to exist any significant distance from a branch of office that is able to approve applications, such as fax machines or internet to transfer documents between locations with sufficient speed.

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) or 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 location from the perspective of the topholder. Both panels show an increase in out-of-state lending during the 1990s.[10] In 1990 fewer than 20 percent of loans originated in HMDA were made on properties in a different state. By 2000 this had increased to over 60 percent at the lender-level, and over 55 percent from the perspective of the topholder. The percentage of out-of-state loans originated has remained high 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 were denied or originated.

Factors that determine loan-closing timelines include the volume of applications processed

---

[8]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 RSSD IDs back to 1990. This dataset also provides information for the lender's parent company or bank holding company.

[9]Credit unions are not included in the SOD. Credit-union branch locations are available from the National Credit Union Association, but only beginning in 2010. Mortgage companies do not have branch locations, and it is not always feasible given our data to see connections between mortgage companies and commercial banks. For example, a mortgage company may be owned by a bank holding company, but that would not be taken into account when considering the locations where the bank holding company has a physical presence.

[10]This pattern holds whether looking specifically at purchase loans, refinances, or all applications (not just closed loans).

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 model 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, etc.) the race and sex of the borrower, whether the borrower has a co-applicant, and a concurrent measure of mortgage application volume.[11].

Figure 4 depicts the results of these regressions. There is a dramatic decline in processing times for refinances between 1995 and 1998, and then continue to drift lower until 2005. Processing times increase again after 2007, but on average about 10 business days faster than they were prior to 1995.

In contrast, we do not see a similar pattern for purchase loans. This is not surprising. Closing dates for purchase loans are often chosen to accommodate the needs of the borrower and seller. Consequently, the time between a purchase-loan application and the closing date can be lengthy, even if the borrower has been pre-approved and provided much of the necessary documentation before making an offer on a 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 average points and fees paid by borrowers over the course of the decade. As seen in the top panel of Figure 5, this series, calculated by Freddie Mac, declined from about 3.0 percent in the early 1980s to about $1/2$ percent in the early 2000s. Although the decline in the series reflects to some extent the reduced costs of mortgage intermediation, it is also influenced by other factors, such as the fraction of borrowers who choose to "buy down" their interest rate by paying higher points on their loans.

The lower panels of Figure 5 plots a measure of mortgage-originator profits and un-measured costs (OPUCs) that accounts for many of these other factors. The details of its calculation is described in Fuster et al. (2013). It measures the portion of the revenues from originating a loan that is kept by the originator.[12] 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,

---

[11]Our measure of volume is the average number of mortgage applications per business day in each month using HMDA. 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

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

while an increase could reflect an increase in profits.[13] Despite that caveat, they provide a clearer measure of costs of intermediation borne by borrowers than the series from Freddie Mac described above.

These panels show that consistent with the changes in cross-state lending and processing timelines, OPUCs declined sharply during the 1990s. They then increased slightly during the housing boom,[14] and then increased dramatically during the Great Recession. Because intermediate costs are influenced by capacity constraints, the bottom right panel controls for volume of applications using the same measure as in the previous subsection on HMDA timelines (specifically, the number of applications per business day per month). The results are little changed.

3.4   Reversal of Trends after the Great Recession

An interesting facet of both the timeline declines in Figure 4 and the cost declines in Figure 5 is their reversal after the 2000s boom. Today, mortgage loans today take about as long to close as they did in 1995, and intermediation costs are even higher than they were two decades ago. In this subsection, we discuss the policy and regulatory changes that significantly increased the burden on individual loan originators after the housing boom. These regulatory changes lengthened timelines and raised costs even though the underlying technology had improved. Then, in the following sections, we discuss how the relationship between mortgage debt and income, which reflects lending technology, did in fact change in the 1990s, yet has undergone no subsequent reversal.

One of the most significant factors to affect mortgage originators after the 2000s was a change in the repurchase policies of 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 originated near the peak of the housing boom, which had very high default rates. The top right panel, also derived from data in Goodman, Parrott, and Zhu (2015), depicts the fraction of repurchased loans that had always been current at the time of repurchase. Virtually none of the problematic boom-era repurchases are in that category, but virtually all of the post-crisis repurchases are. The implication is that after the housing boom lenders selling to the GSEs had to be much more

---

[13]For example, during the refinancing boom of 2000–2003, intermediate costs went up, but loan processing timelines stayed the same. This is consistent with a story in which people had to pay more to keep processing times low and were willing to do so.

[14]This could reflect higher originator profits due to higher demand for mortgage debt.

careful on every loan that they sold. Failing to follow the letter of the agencies' rules could result in a costly repurchase, and a stellar repayment history on the loan no longer protected the originators in this regard.

The lower panels of Figure 6 illustrates the consequences of another change in the post-crisis lending landscape: the effect of 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 the 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. The goal of this regulatory change was to give mortgage borrowers more information about potential loans at the start of the origination process, so they could better shop around for the best mortgage deal. Additionally, disclosures near the end of the process about the mortgage the borrower ultimately chose were designed to make sure that their were no surprises for the borrower about the loan at closing.

As illustrated by the lower left panel of Figure 6, lenders generally found the TILA-RESPA rule to be costly and burdensome. The panel shows the results of a survey conducted by Federal Reserve and Conference of State Bank Supervisors that asked community banks about the burdens of various banking regulations, including those unrelated to mortgage lending. The community bankers reported that the new TILA-RESPA rule was the most confusing regulation, accounted for the highest share of compliance costs attributed to regulation, and was the regulation they would most like to change. The lower right panel provides direct evidence of the consequences of TILA-RESPA changes on various aspects of mortgage origination. Among the most important consequences of the new rule was a slowdown in the pace of orignation, delayed closings, and increased staffing costs.

Taken together, the new buyback policy of the GSEs and the increased regulatory burden after the crisis explain the longer timelines and the increased costs noted above. That is, even though lending technology has improved, lenders struggled to adapt their software to the new rules. Yet the relationship between income and debt embedded in this software did not change, as we will show in the next section.

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

4.1 The Canonical Debt-Income Regression

The relationship between mortgage debt and income is the focus of a growing body of empirical research, most of which uses some variant of what we call the canonical specification

for the debt-income relationship:

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

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 richer 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. If lending standards relax and higher DTI ratios are allowed, then low-income households will be able to borrow more, increasing the amount of debt at the bottom of the income distribution and causing the positive relationship between debt and income β to decline.

Figure 7 motivates our regression analysis with binned scatters of log debt and income using our two main data sources, HMDA and the AHS. The plots show that both both series, the relationship between income and debt is well approximated by a log-linear specification. The plots also show that for both datasets, the positive relationship between debt and income became less steep over the course of the sample periods. By running yearly regressions, we can trace out exactly when this decline in the importance of income and debt took place.

4.2   Regressions Using HMDA Data

We first describe the results of the canonical specification using individual-level debt and income data from HMDA. Some limited demographic variables are available in the HMDA data and using HMDA data allows us to include CBSA fixed effects to control for local housing market characteristics including local demand effects. The CBSA fixed effects also correct for any over reporting of income that is consistent across an MSA and control for any local demand effects. We modify Equation 1 to

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

where the vector X_{it} includes dummy variables denoting the borrower's race and gender and c indexes the CBSA. We also cluster the standard errors by CBSA.[15] Due to the measurement issues with individual HMDA income described in Section 2, we instrument for HMDA income using the median household income in each census tract from the decennial census and ACS.[16]

---

[15]The results are robust to clustering by state.
[16]We

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 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, calculated as the estimates of α_t plus $\beta_t \times \overline{I}_{.t}$, where $\overline{I}_{.t}$ denotes the mean of income in the entire sample. These expected amounts increase sharply after 2000, when, as noted in the introduction, the ratio of aggregate U.S. mortgage debt to aggregate income rises.[17]

4.3 Regressions Using AHS Data

Figure 9 presents analogous results using data from the AHS. Recall that in this dataset, there is no need to IV for income, because it comes from survey data and not loan applications. Unfortunately, small sample sizes in the AHS mean that we cannot run regressions with CBSA fixed effects.[18] The top panel mimics the HMDA results in that debt-income relationship summarized in β_t declines throughout the 1990s and flattens out in the 2000s. The bottom panel shows estimates of expected mortgage amounts. As with the HMDA data, these amounts increase rapidly during the boom.

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

What factors underlie the declining relationship between income and mortgage debt in the 1990s? Figure 10 shows that nominal interest rates are unlikely to explain this pattern. The top left panel of Figure 10 shows the mean and median of nominal interest rates on all AHS movers, along with the conventional 30-year rate from Freddie Mac's mortgage survey. Interest rates declined significantly in the late 1980s and early 1990s and continued to trend downward through the mid-2000s. The top right panel shows that dispersion in interest rates declines from 1985 on, consistent with the increased geographic competition among lenders discussed earlier. But the lower left panel shows the the β_ts from our canonical regression display the same pattern even after controlling for individual-level interest rates in the regression canonical; the interest-rate coefficients themselves appear in the lower right panel.

---

[17]Figure A.1 in the appendix contains results from parallel regressions not including CBSA fixed effects and without using census tract level median household income as an instrument.

[18]We do, however, include yearly fixed effects for the four Census regions of the country: Northeast, Midwest, South and West. These results along with additional robustness checks are included in Figure A.2 in the appendix.

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). An oft-cited event in this literature occurred in 1995, when the Clinton Administration brought together housing-market participants from the private and public sectors to form the National Homeownership Strategy, which was designed to raise the number of homeowners by eight million by 2000. The varied nature of the 100 action items 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, 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 GSEs efforts to expand lending to underserved borrowers were conducted through affordability programs, such as Freddie Mac's Affordable Gold program, which was begun 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 downpayments.[19] Fannie Mae had a similar affordable-loan program, the Community Homebuyer Program, which began in 1990. Government policies that rewarded lenders form 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 GSE acts had much of an effect (Bhutta, 2011, 2012; Avery and Brevoort, 2015).

5.1 Credit Scoring and Mortgage Lending in Theory and Practice

In contrast to government policy and interest rates, 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 human decisions about mortgage risk. But the intuitive decision rules were hard to code into formal algorithms, and the development of AI systems in the early 1990s soon turned out to be a dead end. As one industry professional wrote, the initial AI 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, 1998b, p. 4).

---

[19]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 homeowner-counseling course.

At the other end of the spectrum from human judgment 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' past credit histories and the risk characteristics of loans. A more complicated algorithm could be estimated via machine learning to predict default based on non-linear and/or hierarchical relationships among these variables. In either case, the resulting parameter estimates can be used to construct predicted probabilities of default that would inform a lender's decision on whether to make a loan and on what interest rate to charge. 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.

Yet even as credit scoring transformed consumer lending in the 1990s, many lenders doubted that credit scores would be much help in predicting mortgage default. Unlike credit cards and other forms of consumer debt, mortgages are secured loans, and the borrower's equity stake in the underlying collateral plays a critical role in his default decision. In fact, the central academic project for 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, this decision is completely characterized by the level of negative equity at which a borrower "ruthlessly" or "strategically" defaults. This threshold equity level is a complex and time-varying function of borrower equity, house prices, and interest rates. Adverse life events such as job loss and income shocks do not lead to default in the FOM, because borrowers can ride out adverse events with unsecured borrowing at the risk-free rate.

In reality many borrowers are liquidity constrained, so job losses and other adverse shocks often prompt default on underwater mortgages. The mortgage-default literature is now attempting to blend insights from the FOM with those of the "double-trigger" model of default, which links default to the simultaneous occurrence of negative equity and an adverse life event such as unemployment.[20] Such borrowers lack the liquid funds to make their monthly payments, and those that also have negative equity are unable to sell their homes for enough to pay off their loans. Because of liquidity constraints, the unemployed borrowers are also prevented from taking out unsecured bridge loans during their periods of financial distress, as the FOM assumes. Default is then the only possible outcome.[21]

Another relevant finding from the mortgage-default literature concerns negative-equity

---

[20]Examples of this work include Corradin (2014), Campbell and Cocco (2015), Laufer (2018), and Schelkle (2018).

[21]For a more complete exposition of the FOM and double-trigger model, see the survey in Foote and Willen (2018).

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% negative equity (Kau, Keenan, and Kim, 1994; Bhutta, Dokko, and Shan, 2017). In the data, however, negative equity typically exceeds this level without the borrower defaulting. This result is relevant for mortgage-default modelling because it suggests that empirical models do not have predict the relevant negative-equity default thresholds an an FOM-consistent way.

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 mortgage-default model. Borrowers that start their loan histories with low LTV ratios (that is, with high down payments) are less likely to experience any level of negative equity, which is a necessary condition for default. Low-LTV borrowers are also less likely to experience the very deep levels of negative equity at which ruthless defaults sometimes occur. Borrowers with high credit scores should also have low mortgage-default risk 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.[22] Although income is critical in the double-trigger model, 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 this 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 an borrower will experience a shock large enough to engender 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. 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.

5.2 Integrating Credit Scores into Mortgage Lending

The default models run by mortgage researchers throughout the 1990s confirmed these theoretical predictions. Two decades earlier, a small-scale NBER study on individual-level default Herzog and Earley (1970) found that initial LTV ratios were good default predictors, and the 1990s results confirmed these results. The new credit scores being developed at the time also entered mortgage-default regressions significantly. Mahoney and Zorn (1997) discuss Freddie Mac's modelling work during 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 the authors also noted researchers typically found that initial DTI ratios were weaker default predictors. "While capacity is an

---

[22]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.

important underwriting component, debt-to-income ratios generally are less powerful predictors of loan performance than other factors. This sample points to both down payments and credit-bureau scores as better indicators of mortgage risk" (Mahoney and Zorn, 1997, p. 16). Results in an influential Federal Reserve study also supported the use of credit scores for mortgage-lending decisions (Avery et al., 1996).

Armed with these empirical results, 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, the letter said, 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

> perform 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)

In addition to supporting the general use of credit scores in the underwriting process, the empirical default models developed by the GSEs were the basis of scorecards that could weigh all numeric 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 GSE's Affordable Gold program, and which we reprint as Figure 11. The worksheet allocates risk units to a loan files based on the application's LTV ratio, DTI ratio, credit scores, and other credit information.

The weights in the worksheet reflect the high importance of equity and credit scores—and the low importance of origination DTI—in empirical default models. The table below illustrates this fact by using the worksheet to evaluate three hypothetical loans. Loan A has an LTV of 90%, a FICO score of more than 90%, and a DTI ratio exceeding 50.5%.[23] Although the high DTI ratio costs the potential borrower 30 risk units, the high FICO

---

[23]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.

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 borrower with a low credit score and a carries modest DTI ratio. This loan scores one unit above Freddie's cutoff. Finally, Loan C has a very high LTV and DTI ratios (99.5% and 50.5%, respectively), as well as 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.

	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.

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 risk. The superior speed and accuracy of scorecard-based evaluations were illustrated powerfully by a head-to-head comparison 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, quality-control investigators there indicated that the loans' risk characteristics were too high, so a sample of 700 was scored using the Gold Medal 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. But while there was some overlap, the humans and the worksheet differed substantially on the set of investment-quality mortgages.

By following the mortgages over time, Freddie Mac could conduct a horse race between the worksheet and human underwriters regarding their 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 non-investment 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: "Review 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 default models and their substantial financial resources by developing software to predict default, and then distribute this software to loan originators.[24] By 1995, both GSEs had developed automated underwriting (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 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 sample as a whole. 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 designated about half of the loans as acceptable for purchase by Freddie Mac, although this half ultimately performed about the same as

---

[24]Initially, Fannie Mae was a leader in trying to mimic human decisions with computerized AI algorithms. But the company eventually realized that modelling default directly was also needed. See McDonald et al. (1997) for details.

non-investment-quality half.[25] 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 humans, 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 the humans rejected (a group comprising 20.8 percent of the sample), while it rejected many mortgages the humans 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).

How much of this credit-box expansion involved an increase in permissible DTIs? Once lending policies have been encoded into a proprietary AU system, we can no longer evaluate them with comparisons of hypothetical loans, as we did for the Gold Measure Worksheet. But it is very likely that relaxed DTI standards were an important part of the credit expansion. As the use 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 models 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 level of the DTI ratio, but rather the *difference* between the new monthly mortgage payment and the amount that the homeowner was already paying for housing.[26]

More recently, the GSEs evaluation of DTI has taken on additional importance after the passage of the Dodd-Frank Act, which imposes ability-to-repay rules for mortgages. In general, these requirements are automatically met for loans acceptable for purchase by the GSEs. As of late 2017, Fannie Mae reported that for manually underwritten loans, the

---

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

[26]Yet even this weaker income test "is not a major-tripper upper," the journalist quoted an industry participant as saying. See "Building Blocks of Automated Underwriting," by Lew Sichelman, United Feature Syndicate. (Available at `http://articles.orlandosentinel.com/2000-06-04/business/0006010404_1_automated-underwriting-payment-shock-fixed-loans`.)

maximum permissible back-end DTI ratio (which includes non-housing debt) in most cases is 36 percent of the borrower's "stable monthly income." Manual underwriters could permit DTIs up to 45 percent, however, as long as the borrower met certain requirements related to her credit score and available liquid reserves. Yet for loans underwritten through Desktop Underwriter, the maximum allowable DTI ratio is 50 percent.[27]

All told, we find ample evidence that origination DTIs became less important in lending decisions during the 1990s, consistent with modern theories of mortgage default and the debt-income regressions that were presented in the previous section. It is important to note, however, that the technological improvements to mortgage lending during the 1990s were exactly that—improvements. Empirical models informed by credit scores generated simple scorecards that were much more accurate in predicting default risk that human judgment, particularly for marginal loan applications like the Affordable Gold sample. To the extent that the information used for these models is racially neutral, then the models also deliver racially unbiased lending decisions, something that 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. They came to dominate the mortgage-origination market by the start of the 2000s, after being introduced only a few years before. And they remain in place today.

A deeper point about these technological improvements is that the computerization of mortgage lending has close parallels with technological change in many other industries. When the 1990s began, it was not obvious how computers would or should affect mortgage lending. Some lenders believed that the best way to use the new technology was to develop computerized AI systems that would largely replicate their existing way of doing business, which was based on human decision-making. Over time, however, lenders realized that information technology allowed them to largely replace human judgment with predictions from empirical models, which were cheaper and more accurate. There is a close parallel with initial use of electricity by manufacturers in the 20th century, as described in the seminal paper of David (1990). After electric power became available, U.S. manufacturers took some time to realize that electricity could do more than simply replace the central steam-power source in their predominantly multistory factories. Because electricity could be distributed more easily throughout factories than steam power could, electricity essentially allowed manufacturers to essentially turn their factories on their sides, by constructing single-story buildings through which materials could move more easily. David and others have argued that like the advent of electricity, computerization in the modern era has also led to delayed productivity gains, because today's firms also need time to fundamentally change business practices in ways that

---

[27]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 November 10, 2017.

computerization allows. The story of how computers transformed 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

Research on the 2000s housing boom often links the rapid increase in house prices then to a credit expansion. Results above show that to the extent there was a change in the debt-income relationship, it was driven by technological developments in the 1990s, not by securitization or some other development in the 2000s. Still, a valid question is whether the earlier change in underwriting methods might have contributed to the 2000s housing boom. In this section, we discuss the potential impacts that the 1990s developments might have had on the overall homeownership rate and on housing prices.

According to the Census Bureau, the U.S. homeownership rate rose from 64.0% in 1994 to 69.0% 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 the increase in homeownership to be especially large for households with high 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 CPS/HVS microdata by age and education. Young college graduates are known to have steeper age-earnings profiles than non-college persons of the same age, 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 but no degrees). The middle panel of the top row shows that among households headed by a 25–29 year old, ownership 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.[28] For all but the youngest age groups, ownership rates for college-educated households are substantially higher than for the less-educated. Putting the pieces together, we see that credit expansion of the 1990s essentially allowed younger, better-educated households to purchase households sooner than the otherwise would have. These households had relatively high permanent incomes (because of their education) but low current incomes (because of their ages).[29]

---

[28]For households under 25, ownership rates rise for both higher- and lower-educated households. But ownership rates for both group remain small throughout the sample period.

[29]These groups would also be less likely to experience income disruptions. As is well known, the probability that a worker transitions for employment to unemployment is an inverse function of his education level (Mincer, 1991; Mukoyama and Sahin, 2006).

Although the 1994–2004 increase in homeownership figures prominently in many narratives of the housing crisis, this change was small relative to the ownership increase immediately after World War II, when the introduction of zero-down-payment Veterans Administration (VA) loans and other changes in housing finance increased homeownership by yearly 20 percentage points.[30] Yet like the recent homeownership increase, the more-significant changes in housing finance during the middle of the 20th century also appear to have allowed persons to buy homes sooner than they ordinarily would, as Fetter (2013) has argued. Figure 13 uses data from decennial censuses and the American Community Survey (ACS) to provide a more detailed look at homeownership changes during various periods. The use of Census and ACS rather than CPS/HVS data allows us to disaggregate by four educational groups, rather than two, and also permits analysis by single-year-of-age rather than 5-year age groups. The top panel depicts ownership changes by single-year-of-age and education for the 1940–1960 changes. As Fetter (2013) points out, the main effect of underwriting changes immediately after World War II was to increase homeownership rates among young persons 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 young persons in all educational classes except high school dropouts experienced rising ownership during this period. The lower two panels depict changes in ownership between 1990 and 2000 (bottom left panel) and 2000 and 2005 (lower left panel, which uses 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 underwriting changes. Homeownership for persons with no years of college education did not change during this period.

6.2   Effects on Housing Prices

We now turn to the effect of 1990s developments on housing prices. The main idea behind much research on 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 housing 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), and implication of the link between house prices

---

[30]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.

and future rents is that relaxed lending standards should merely encourage current renters to purchase their existing residences from their landlords. 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. The fundamental point is that as long as homes can flow between the rental and owner-occupied markets, an increase in demand for owner-occupied homes resulting from changes in lending standards can be satisfied without much of an increase in housing prices.

Housing experts have long recognized that the conversion of owned 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 in measuring flows of homes between various ownership 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 from destruction, conversions to nonresidential uses, and so on.[31] The underlying data source for CINCH is the AHS, which returns to the same houses every two years and queries the residents as to 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 owner-occupied in 1987, so that total owner-occupied losses from 1985 to 1987 comprised 8.1% of the 1985 stock. About 6.7% of this stock was converted to rentals or became vacant by 1987s, with the remainder lost to conversions, destruction, or in other ways. The key takeaway from this table is that individual properties very often flow from one tenure arrangement to another. 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%.

We use the AHS microdata to calculate how many owner-occupied homes become rentals as opposed to becoming vacant. 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 the share of owned properties that about 50 percent of owned properties flow to rentals, while the remaining 50 percent become vacant. The lower left panel shows the outflow data for rented units, where the split is no longer 50:50. Instead,

---

[31]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.

about two-thirds of rental outflows become vacant while one-third becomes owner-occupied. Yet the rented-to-owned flow is still relatively large. These estimates suggest that about $1/35 \approx 2.9\%$ of renter-occupied homes flow to owner-occupied status every year.[32] 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 after 1991. This consistent with the claim that higher effective demand from relaxed lending standards can be satisfied in part by gross flows among rented to owned homes.

Table 4 depicts the annualized gross flows between rented and owned status, as well as the net flow of properties from rented to owned. For most years immediately following 1985, the net flow is negative, consistent with the filtering hypothesis. Much of the low-cost rental housing in the United States comes from the steady depreciation of the owner-occupied stock, so we would expect the flow from owned to rented status to be larger than the flow in the other direction. From 1991-93 to 1997-99, however, the gross flow of rented to owned units rises significantly.[33] Not much happens to the gross flow in the other direction, so the net flow between rented and homes swings positive in the late 1990s, reaching 241,000 units per year in the 1997-99 period. This is consistent with a change in lending standards in the 1990s that expanded relative demand in the owner-occupied market.

Figure 15, studies all of the relevant margins for growth in the owner-occupied stock of homes. The top panel of this figure shows the growth of owner-occupied homes according to published data from the CPS/HVS. Growth is calculated over two year periods, so the chart implies yearly increases of about 1 million units. The lower panel decomposes the sources of change in owner-occupied homes. As we saw in the previous table, early in the sample period this flow was negative, consistent with the general pattern of owned homes filtering down to the rental market. During the mid-1990s, however, the rental-to-owned flow becomes positive, as these conversions contribute significantly to the growth of the owner-occupied stock. The contributions from seasonal and vacant properties also fluctuate during the 1990s, with the vacancy flow tending to reduce the owner-occupied stock and the seasonal flow tending to increase it. However, the net contribution from construction is by far the largest positive contributor to the growth of owner-occupied homes in all years of the sample, and is especially large in the early 2000s.[34]

---

[32]Because the AHS is conducted every two years, the survey-to-survey flow of about two 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).

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

[34] Conceptually, this contribution is the gross gains from newly constructed properties less the losses from destruction and non-residential conversions of owned 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, depicted in the upper panel.

Two patterns apparent in the lower panel are directly relevant for assessing the impact of relaxed lending standards over time. First, the contribution from rentals depicted by the green bars is never positive after 2001, as it was in the late 1990s. Second, the contribution from vacancies shows no general trend toward raising the stock of owner-occupied homes during the 2000s. It would be difficult to replicate these patterns with a model in which lending standards were relaxed and both rentals and vacancies were potential margins of adjustment. A loosening of lending standards that raised effective demand would encourage the conversion of rentals into owner-occupied properties (as was the case in the 1990s), and it would shrink the pool of vacant homes available homes for sale. Yet none of these factors appear to have occurred during the 2000s, when growth in housing prices was strongest. Consequently, these two patterns cast some doubt on the view that price increases during the 2000s were driven in large part by looser lending standards.

6.3 Characterizing the Consequences of Technical Change

There are three main takeaways from our analysis of how technical changes during the 1990s affected the housing market. First, the change in the relationship between mortgage debt and current income during the 1990s had particularly large effects on young college graduates. This pattern is consistent with economic theory, because young college graduates tend to have low incomes today relative to their permanent incomes. It is also consistent with broader changes in lending during the middle of the 20th century. A second finding, also consistent with theory, is that increases in effective demand brought about by 1990s development were met in part by the conversion of rental properties to owner-occupied properties. This finding supports a central claim of Kaplan, Mitman, and Violante (2017): the effect of lending standards on housing prices depends critically on the quantity margins through which higher effective demand for owned homes can be satisfied. Third, with respect to these margins, the behavior of the housing market during the 2000s provides little evidence that the acceleration of house prices during that decade was driven by an exogenous relaxation of lending standards to individuals with low current incomes. The net conversion of rental-to-owned properties reverses course in that decade, and vacancy rates for homeowners continued to drift higher. These findings argue against the claim that an exogenous credit expansion drove price appreciation higher during the 2000s.

7 Conclusion

[To be added.]

References

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.

Adelson, Mark. 2003. "A Journey to the Alt-A Zone: A Brief Primer on Alt-A Mortgage Loans." Nomura Fixed Income Research, June 3. Available at `http://www.markadelson.com/pubs/Journey_to_the_Alt-A_Zone.pdf`.

Amel, Dean, Elliot Anenberg, and Rebecca Jorgensen. 2018. "On the Geographic Scope of Retail Mortgage Markets." Board of Governors of the Federal Reserve System, *FEDS Notes*, June 15. Available at `https://www.federalreserve.gov/econres/notes/feds-notes/on-the-geographic-scope-of-retail-mortgage-markets-20180615.htm`.

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

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(July): 621–648.

Avery, Robert B., Raphael W. Bostic, Paul S. Calem, and Glenn B. Canner. 2000. "Credit Scoring: Statistical Issues and Evidence from Credit-Bureau Files." *Real Estate Economics* 28(3): 523–547.

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.

Bennett, Paul, Richard Peach, and Stavros Peristiani. 2001. "Structural Change in the Mortgage Market and the Propensity to Refinance." *Journal of Money, Credit and Banking* 33(4): 955–975.

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.

Bhutta, Neil. 2012. "GSE Activity and Mortgage Supply in Lower-Income and Minority Neighborhoods: The Effect of the Affordable Housing Goals." *Journal of Real Estate Finance and Economics* 45(1): 238–261.

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. "Editor's Introduction: Expansion of Mortgage Markets and Increasing Use of Technology." *Journal of Housing Research* 11(2): 161–172.

Campbell, John Y. 2006. "Household Finance." *Journal of Finance* 61(4): 1553–1604.

Campbell, John Y., and João F. Cocco. 2015. "A Model of Mortgage Default." *Journal of Finance* 70(4): 1495–1554.

Corradin, Stefano. 2014. "Household Leverage." *Journal of Money, Credit and Banking* 46(4): 567–613.

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 & 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 & Urban Development, Office of Policy Development & Research. Available at `https://www.huduser.gov/portal/publications/ahsrep/data_onhome.html`.

Faucett, James. 2001. "Freddie Mac Says Loan Prospector Has Yielded $6.5 Billion of Savings." *National Mortgage News* 25(21): 13.

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.

Fetter, Daniel K. 2016. "The Home Front: Rent Control and the Rapid Wartime Increase in Home Ownership." *Journal of Economic History* 76(4): 1001–1043.

Foote, Christopher L., and Paul S. Willen. 2018. "Mortgage-Default Research and the Recent Foreclosure Crisis." *Annual Review of Financial Economics* 35(1): 278–296.

Fuster, Andreas, Laurie 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* (December): 17–40.

Fuster, Andreas, Stephanie H. Lo, and Paul S. Willen. 2017. "The Time-Varying Price of Financial Intermediation in the Mortgage Market." Federal Reserve Bank of New York Staff Report No. 805, August. Available at `https://www.newyorkfed.org/research/staff_reports/sr805.html`.

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.

Goodman, Laurie, 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`.

Kaplan, Greg, Kurt Mitman, and Giovanni L. Violante. 2017. "The Housing Boom and Bust: Model Meets Evidence." Working Paper No. 23694. Cambridge, MA: National Bureau of Economic Research.

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

LaCour-Little, Michael. "The Evolving Role of Technology in Mortgage Finance." *Journal of Housing Research* 11(2): 173–205.

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 & 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.

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. 1998a. "The Basics of a Better Application Score." In *Credit Risk Modeling: Design and Application*, Elizabeth Mays, ed., chap. 5, 59–80. Chicago, IL: Glenlake Publishing Company, Ltd.

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

Maselli, Peter. 1994. "Mortages in Minutes." *Mortgage Banking* 55(1): 102–107.

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.

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

Mincer, Jacob. 1991. "Education and Unemployment." National Bureau of Economic Research Working Paper No. 3838.

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

Mukoyama, Toshihiko, and Aysegul Sahin. 2006. "Costs of Business Cycles for Unskilled Workers." *Journal of Monetary Economics* 53(8): 2179–2193.

Munnell, Alicia H., Geoffrey M. B. Tootell, Lynne E. Browne, and James McEneaney. 1996. "Mortgage Lending in Boston: Interpreting HMDA Data." *American Economic Review* 86(1): 25—-53.

Rajan, Raghuram. 2010. *Fault Lines: How Hidden Fractures Still Threaten the World Economy*. Princeton, N.J.: 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.

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

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.

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 & 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/`.
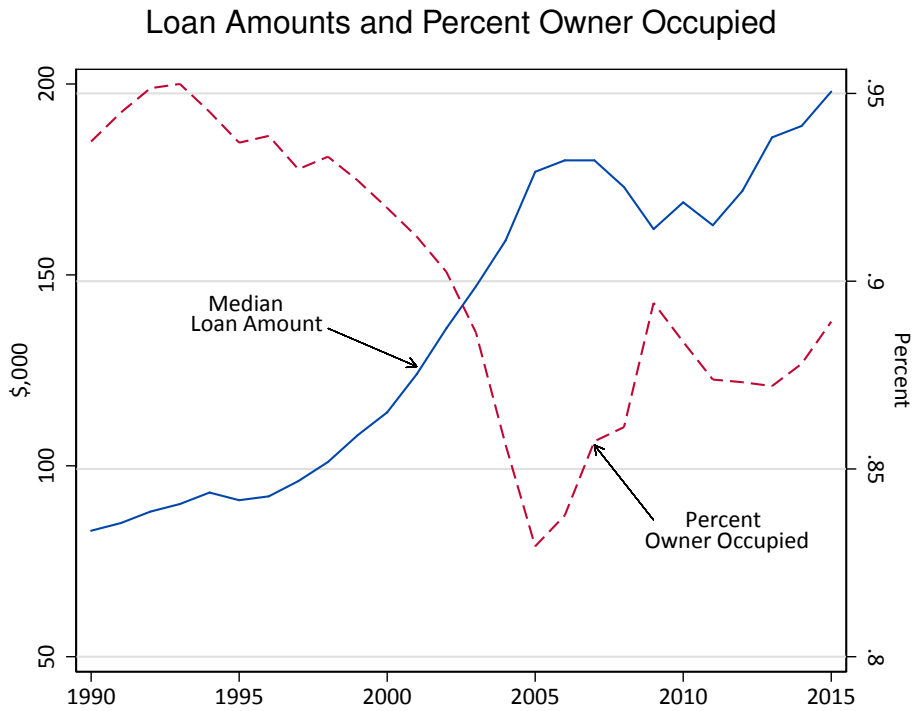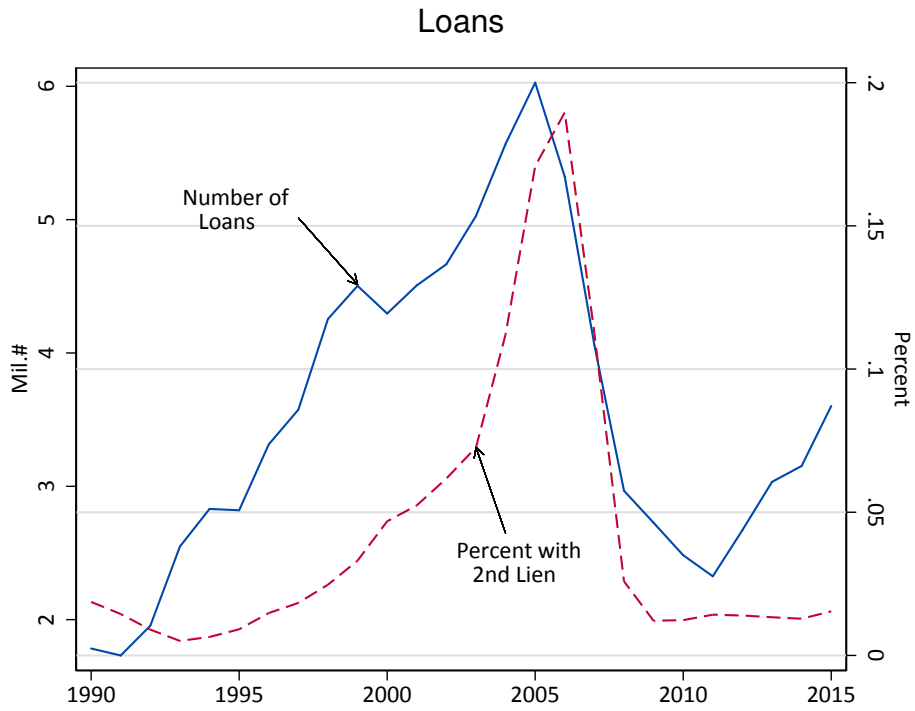
Figure 1. SUMMARY STATISTICS FROM HMDA. *Note:* The total median loan amount refers to the median value of the sum of first liens and associated piggyback loans on owner-occupied properties. The percent with second lien is the share of all loans that have a piggyback loan. *Source:* HMDA.

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:* American Housing Survey.

Mortgage Debt
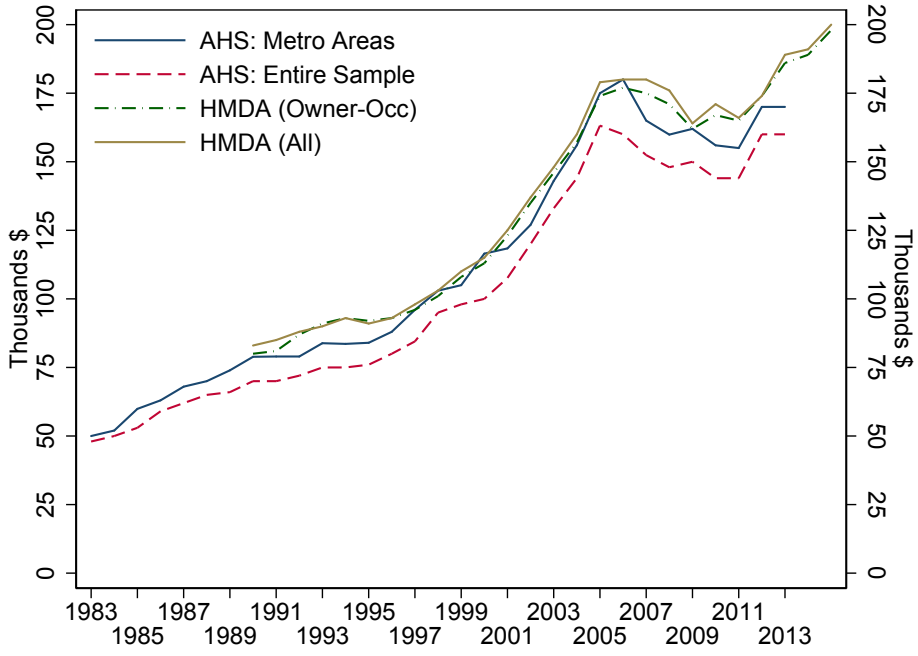
Income
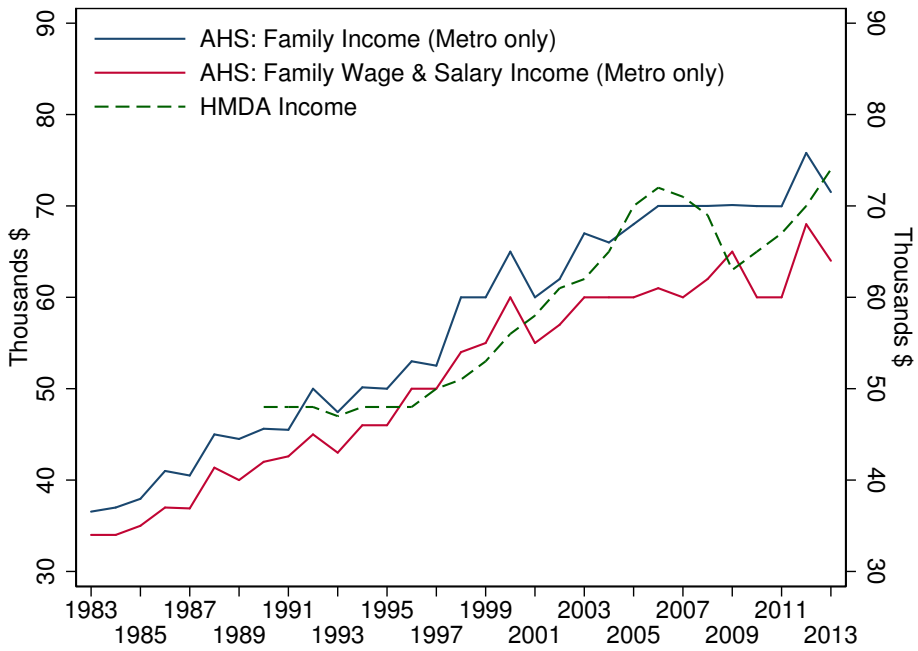
Figure 2. COMPARISON OF NEW MORTGAGE DEBT AND INCOME FROM HMDA AND THE AHS. *Source:* HMDA and AHS.
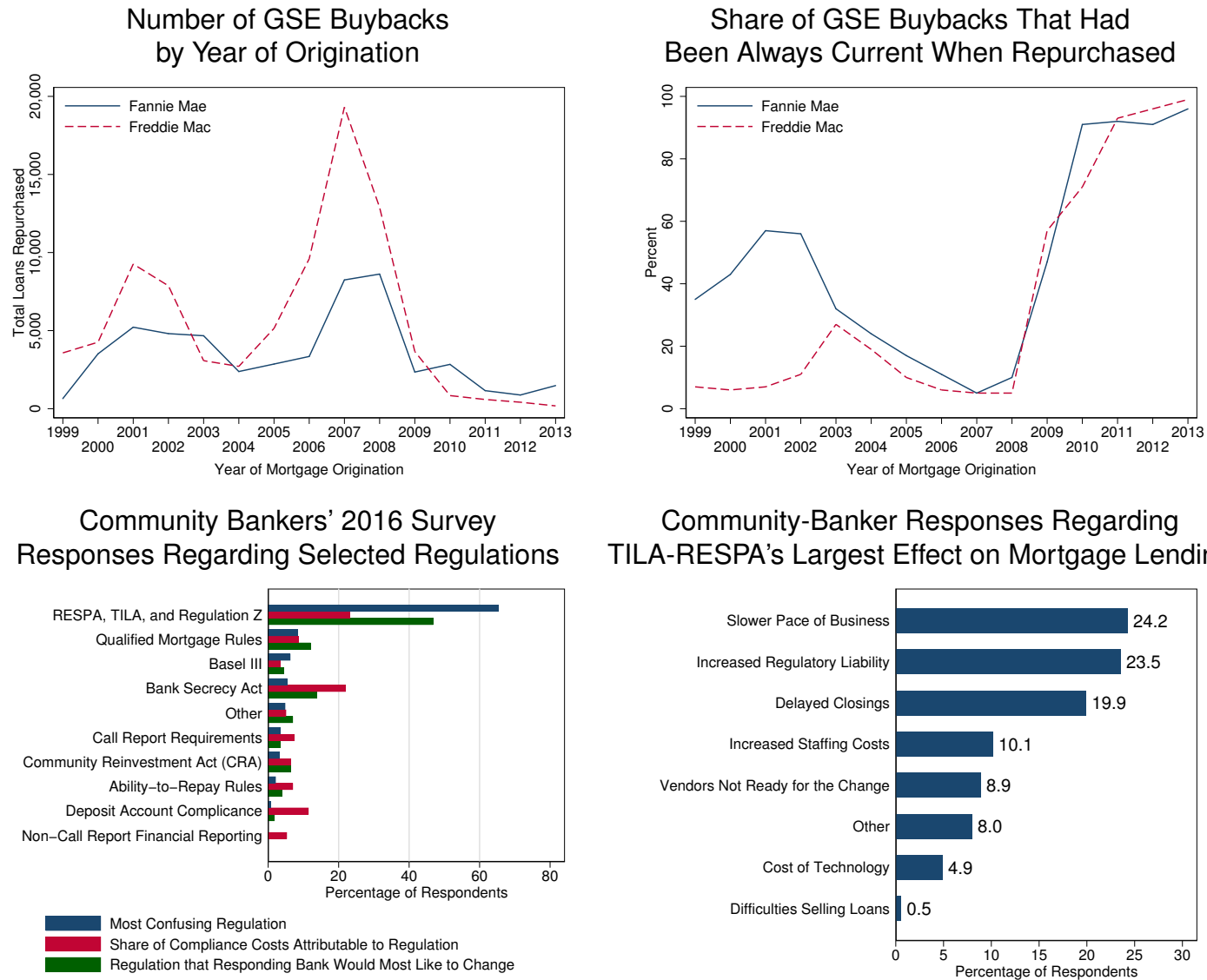
Figure 3. OUT-OF-STATE LENDING IN THE HMDA RECORDS. *Note:* Lenders are matched to their branch locations, parents and bank holding companies (if any) using their Federal Reserve Board Entity number. Mortgage companies and credit unions are excluded from the sample. *Source:* HMDA, the National Information Center, and the Federal Deposit Insurance Corporation Summary of Deposits.
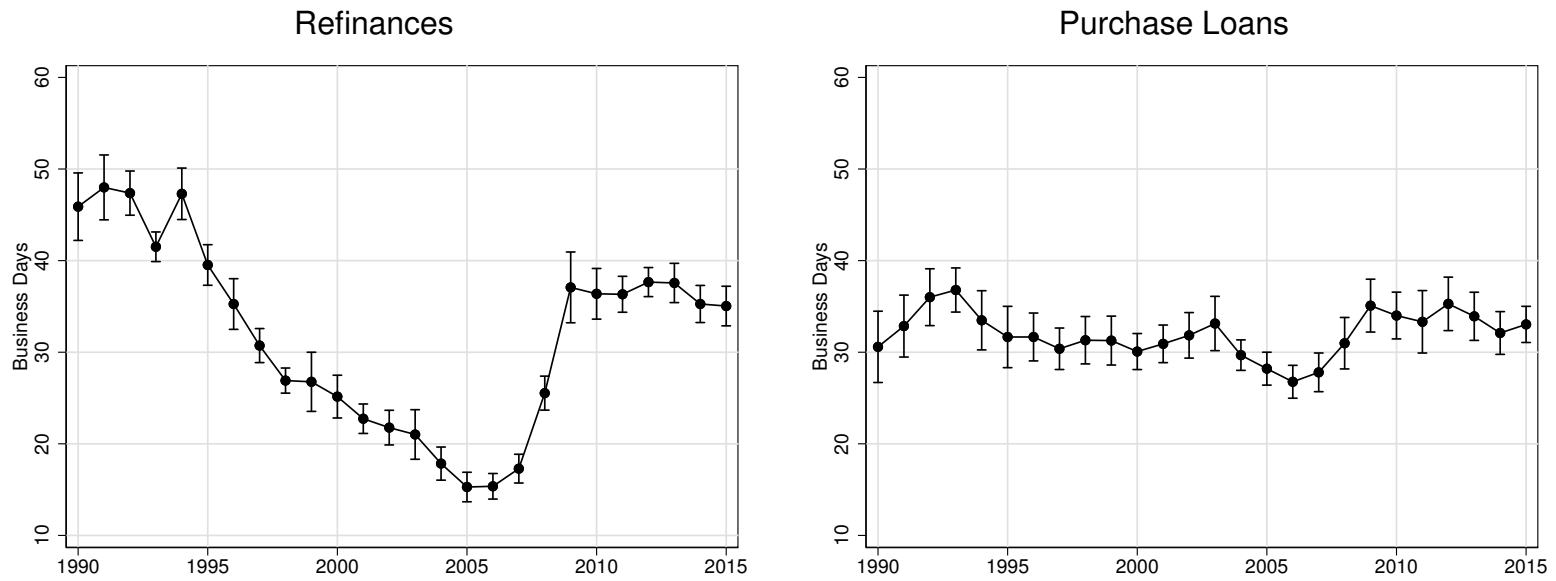
Figure 4. TIME-TO-CLOSE REGRESSIONS. *Note:* Each panel shows a plot of the average processing time by year after stripping out any variation explained by the size of the lender, the borrower's race and gender, whether the borrower has a coapplicant, and the concurrent monthly application volume. The processing times are calculated as of the year of application and include both closed loans and denials. *Source:* HMDA.
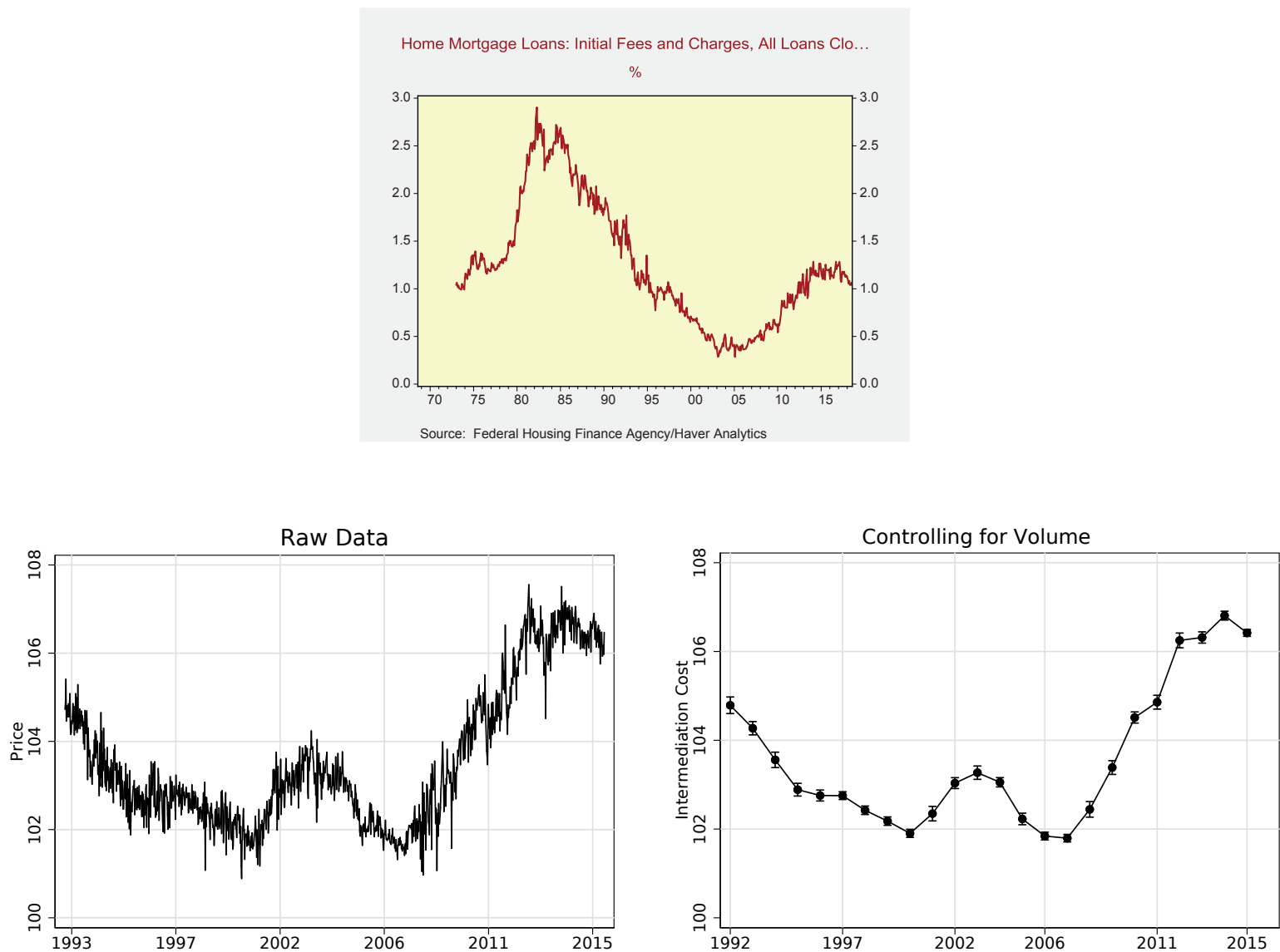
Figure 5. Costs of Intermediation. *Note:* The measure of volume used is the number of loan applications per business day in each month in HMDA. *Source:* HMDA and publicly available data used for Fuster, Lo, and Willen (2017).

Number of GSE Buybacks by Year of Origination

Fannie Mae
Freddie Mac

Total Loans Repurchased

Year of Mortgage Origination

Share of GSE Buybacks That Had Been Always Current When Repurchased

Fannie Mae
Freddie Mac

Percent

Year of Mortgage Origination

Community Bankers' 2016 Survey Responses Regarding Selected Regulations

RESPA, TILA, and Regulation Z
Qualified Mortgage Rules
Basel III
Bank Secrecy Act
Other
Call Report Requirements
Community Reinvestment Act (CRA)
Ability–to–Repay Rules
Deposit Account Complicance
Non–Call Report Financial Reporting

Percentage of Respondents

Most Confusing Regulation
Share of Compliance Costs Attributable to Regulation
Regulation that Responding Bank Would Most Like to Change

Community-Banker Responses Regarding TILA-RESPA's Largest Effect on Mortgage Lending

Slower Pace of Business — 24.2
Increased Regulatory Liability — 23.5
Delayed Closings — 19.9
Increased Staffing Costs — 10.1
Vendors Not Ready for the Change — 8.9
Other — 8.0
Cost of Technology — 4.9
Difficulties Selling Loans — 0.5

Percentage of Respondents

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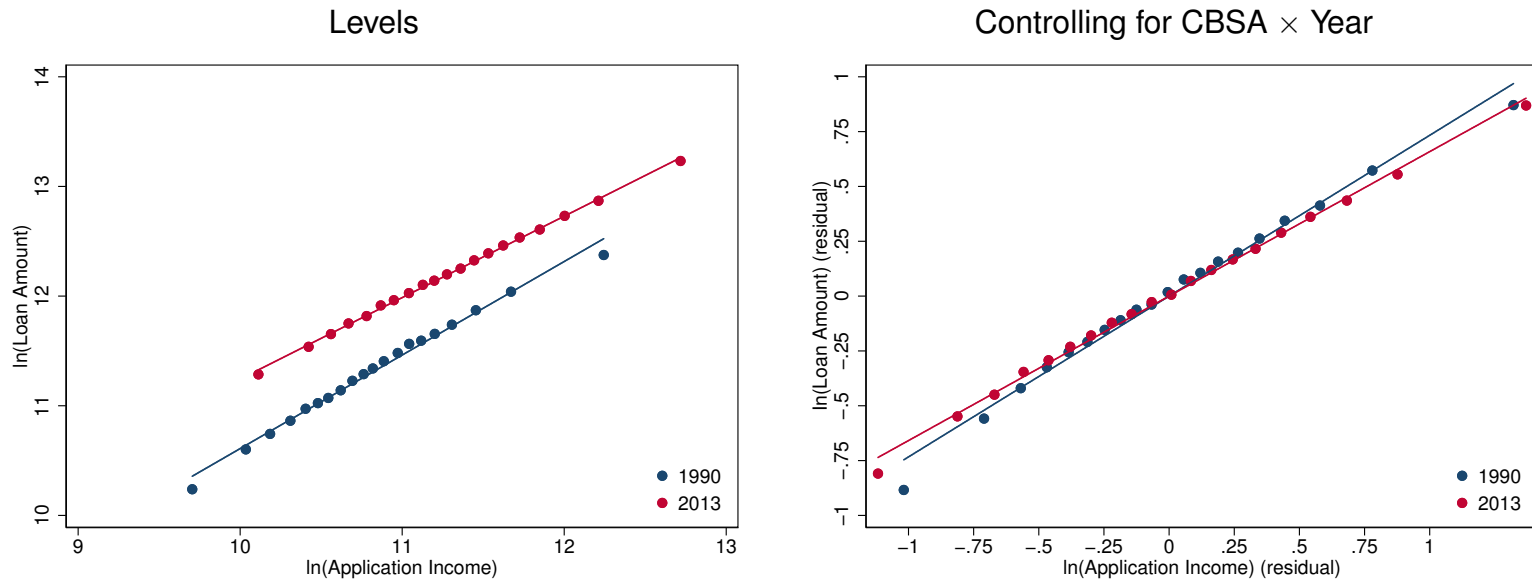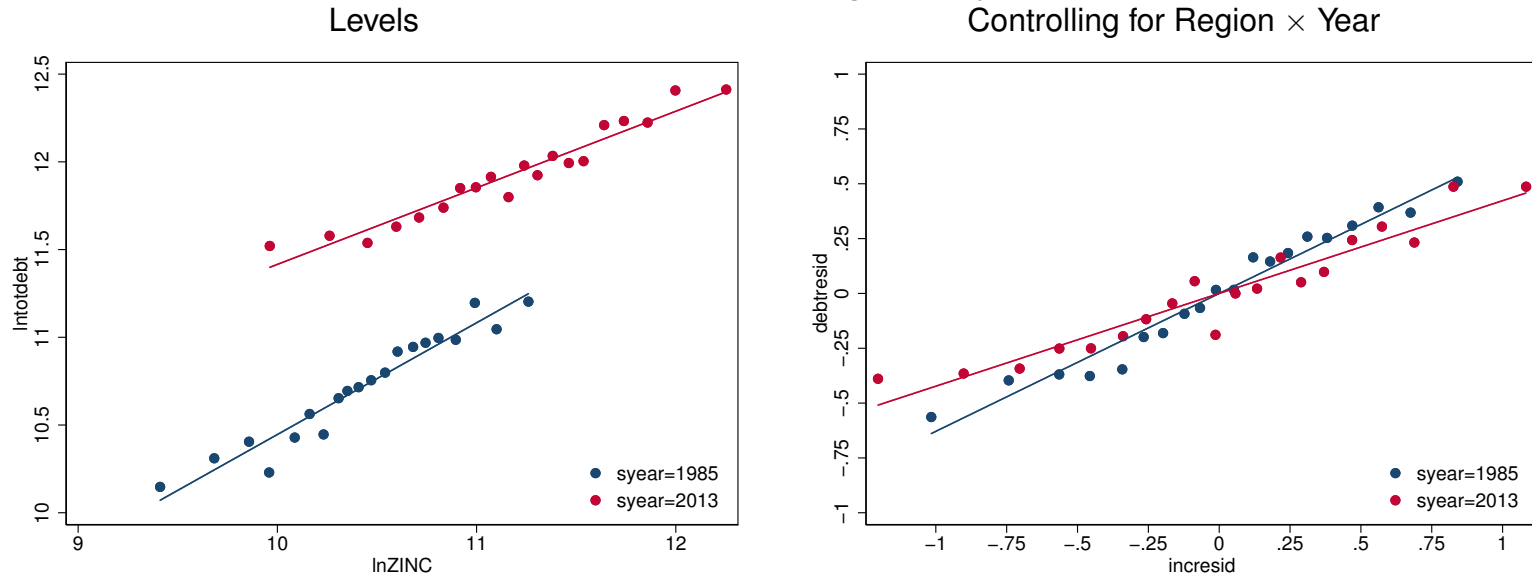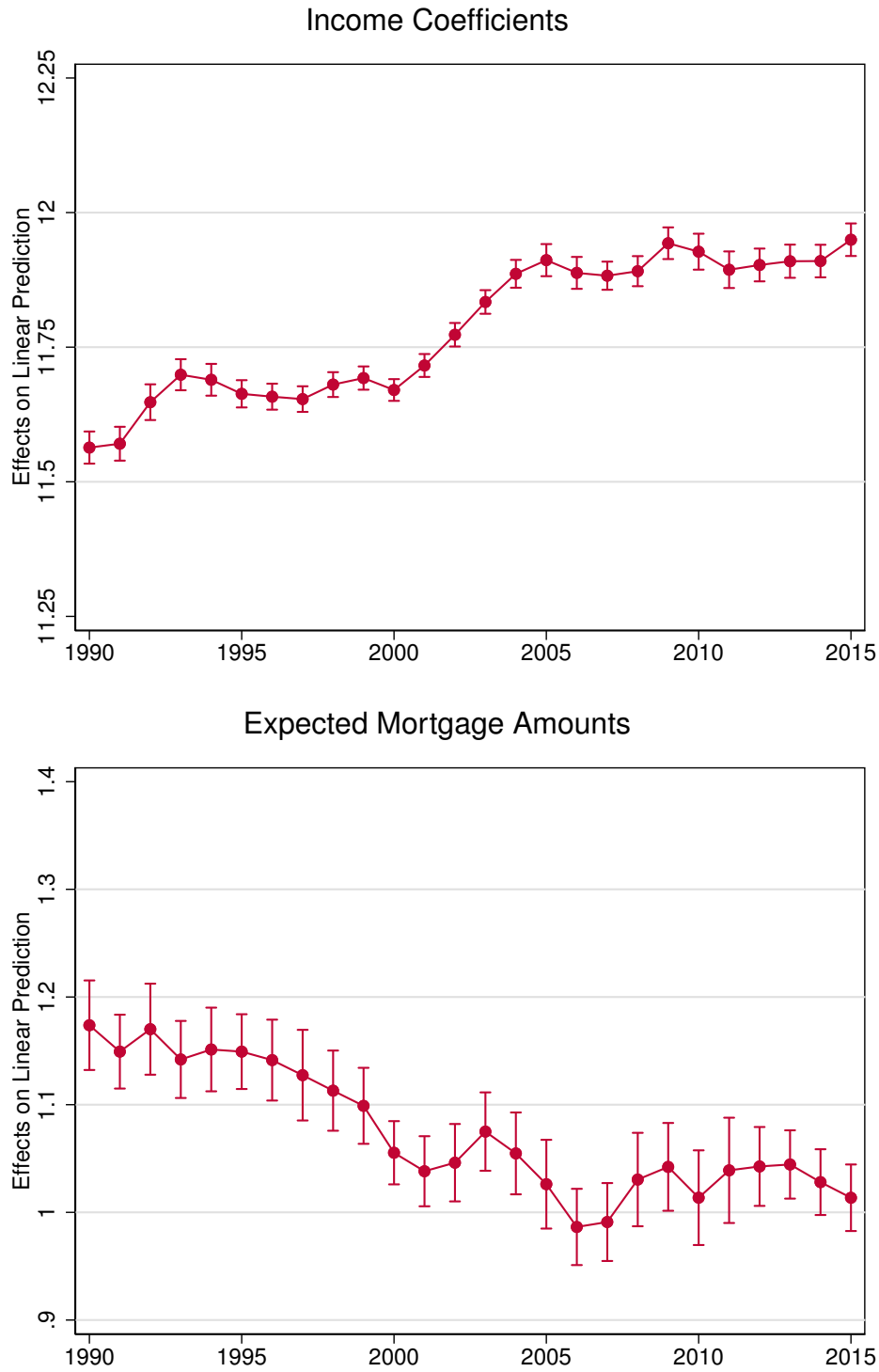
Figure 7. Binned Scatter Plots of New Mortgage Debt and Income at the Individual Level. *Note:* Each dot plots the average loan value for a given income quantile. The two left panels are binned scatter plots of residuals from regressions of the natural log of loan amounts and income on geographic area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

Figure 8. 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 where we instrument for HMDA income using the most recently available Census tract income from the Census and ACS. The regressions include CBSA by year fixed effects and 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:* HMDA, decennial Census, and the American Community Survey.
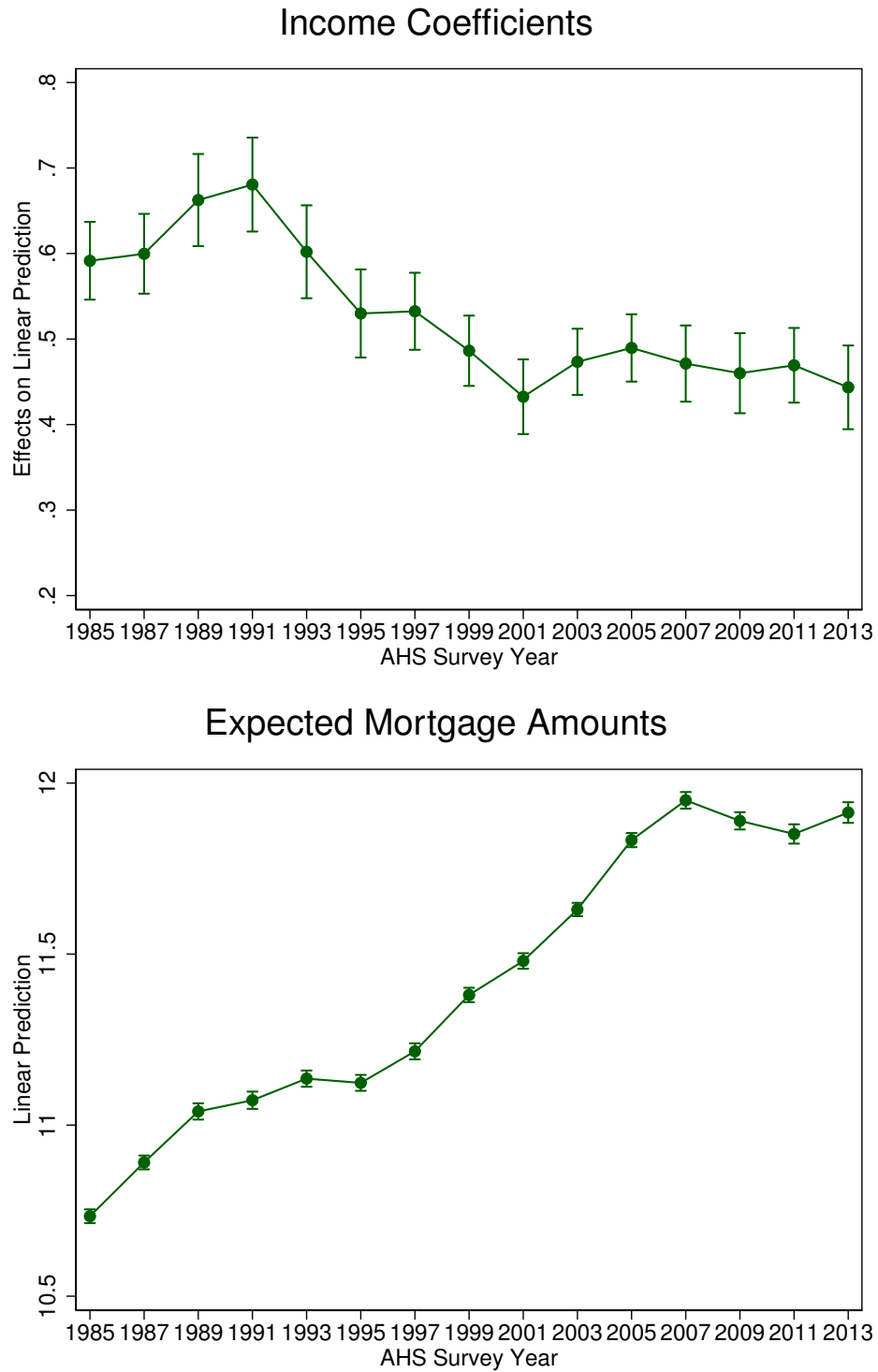
Figure 9. Canonical Regression using Individual-Level AHS Mortgage Balances and Income Levels. *Note:* More here . *Source:* AHS.
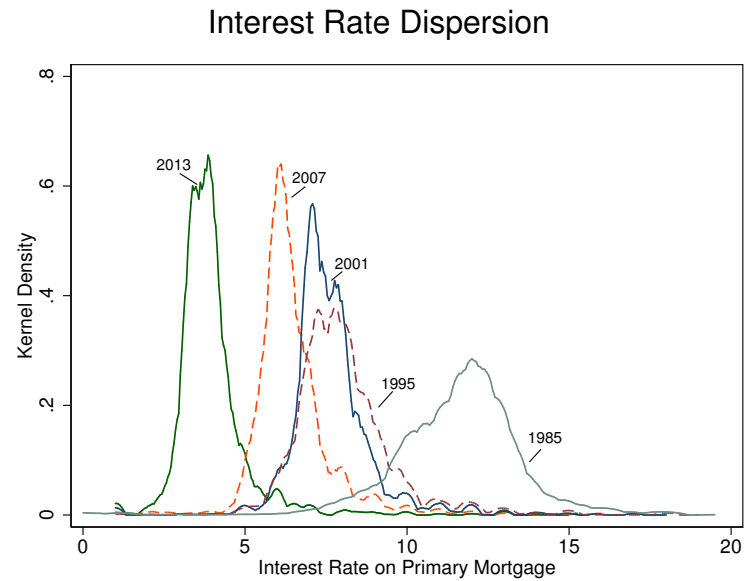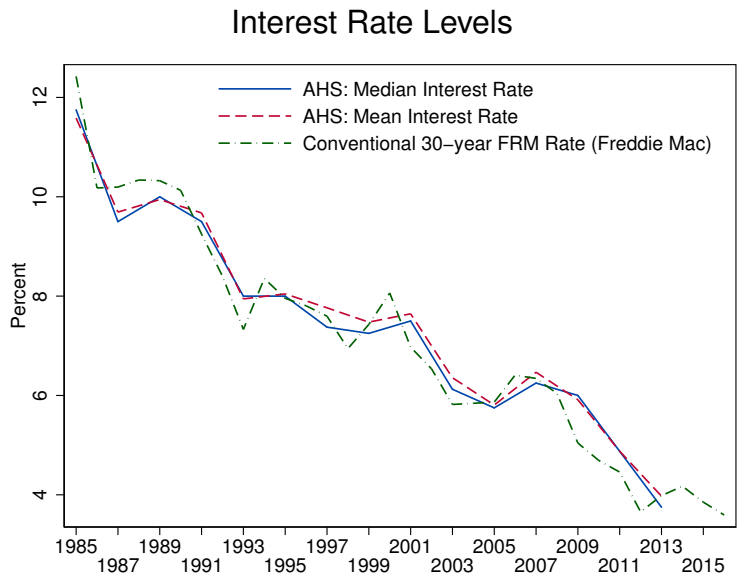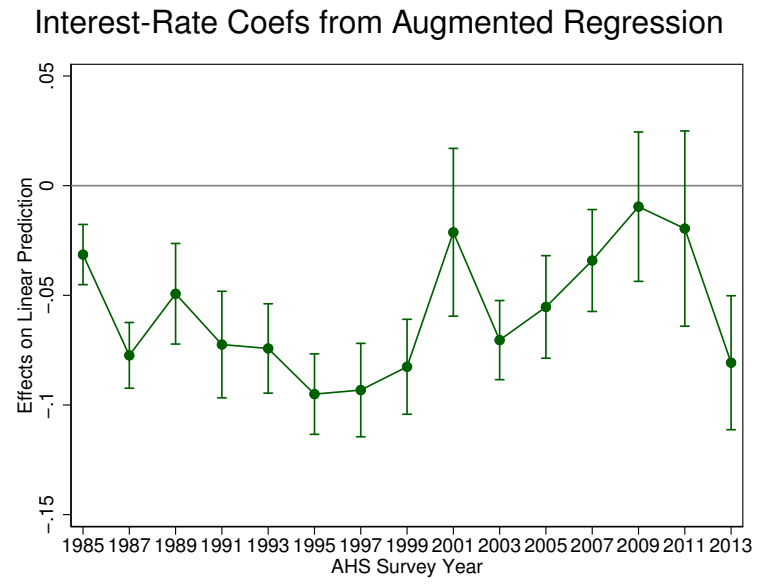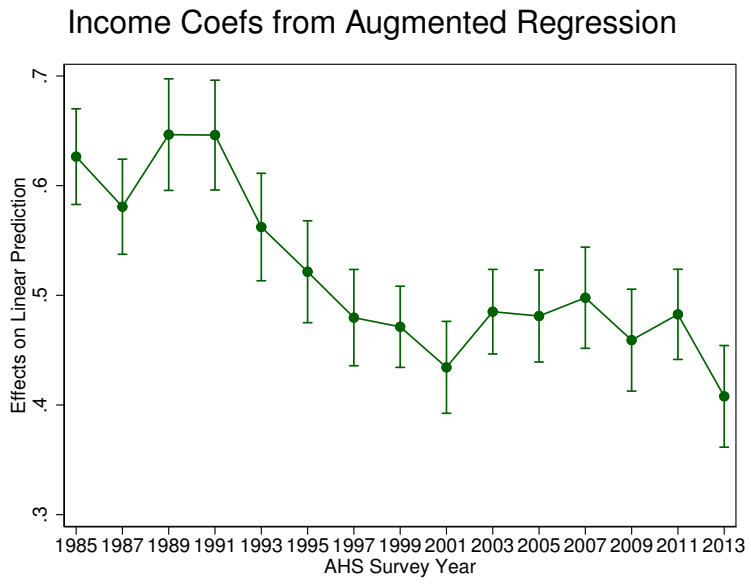
Figure 10. Interest Rates and Mortgage Lending. *Source:* AHS and Freddie Mac.

GOLD MEASURE WORKSHEET—Version 2.0

Freddie Mac

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	RUs	Bankruptcy Score	RUs
Equifax Beacon Score, Trans Union Empirica Score and TRW-FICO Score (See instructions)		Equifax DAS Score, Trans Union Delphi Score and TRW-MDS Score (See instructions)	
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:** _____

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.

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:	RUs
■ 30–180 days delinquent:	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:	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:** _____

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	Property type is:	RUs
Fixed-Rate:	15-Year	–6	2 Unit	5
	20-Year	–4	3–4 Unit	11
	25-Year	–1	Condominium	5
ARM:	Rate-Capped	6		
	Payment-Capped	8		

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 MEDAL WORKSHEET. *Source:* Avery et al. (1996).

		Automated Underwriting				
		1995 Model		2000 Model		
		Accept	Caution	Accept	Caution
Manual Underwriting	Accept	51.6 [1.04]	24.0 [0.21]	27.5 [1.75]	44.7 [0.66]	6.9 [3.52]
	Caution	48.4 [0.96]	20.8 [0.21]	27.7 [1.52]	42.7 [0.75]	5.8 [2.55]
	Total	100 [1]	44.8 [0.21]	55.2 [1.63]	87.3 [0.70]	12.7 [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).
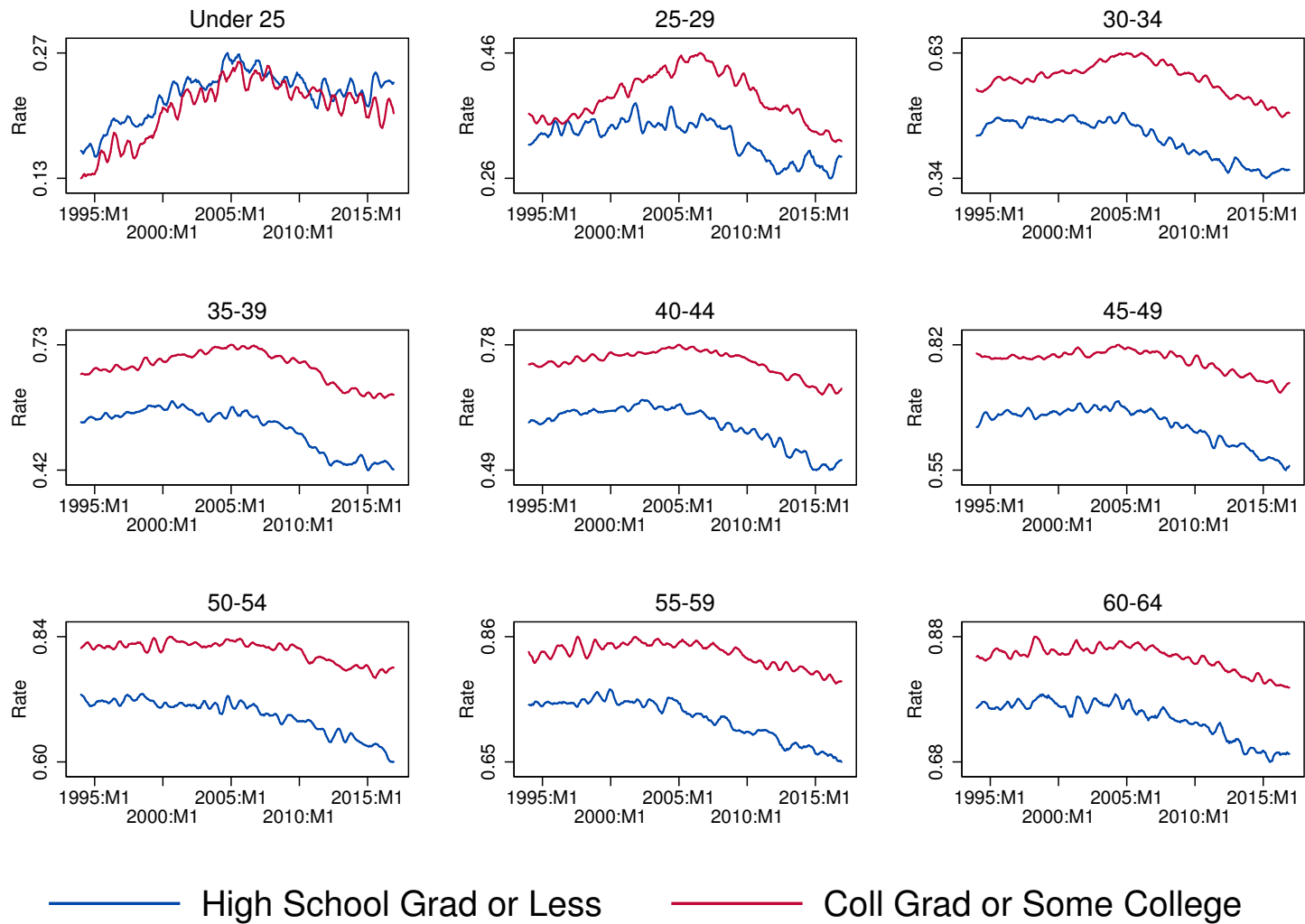
Figure 12. Homeownership Rates by Age of Householder and Education in the Current Population Survey/Housing Vacancy Survey. *Note:* Data are six-month moving averages of monthly rates. *Source:* Current Population Survey/Housing Vacancy Survey.
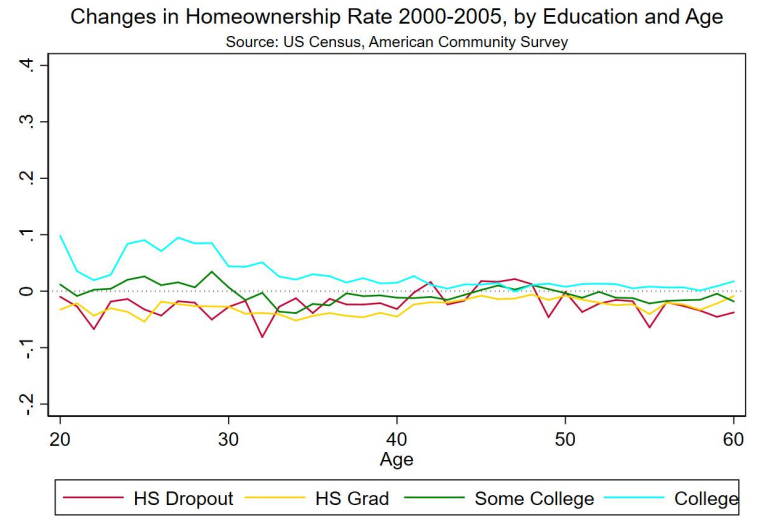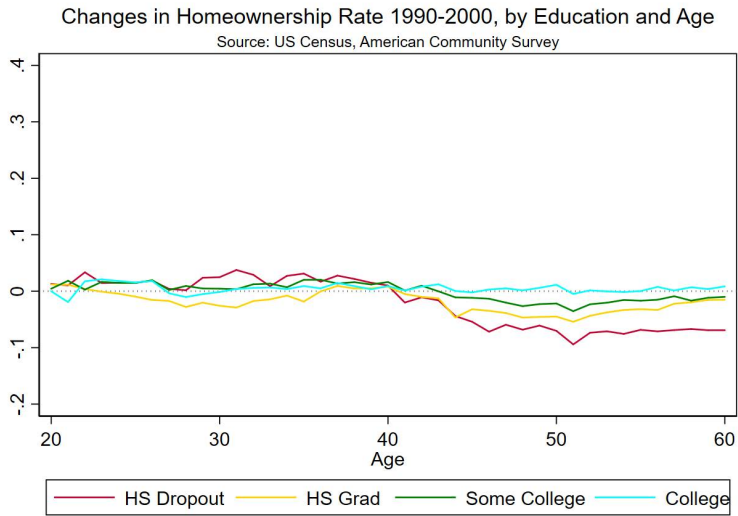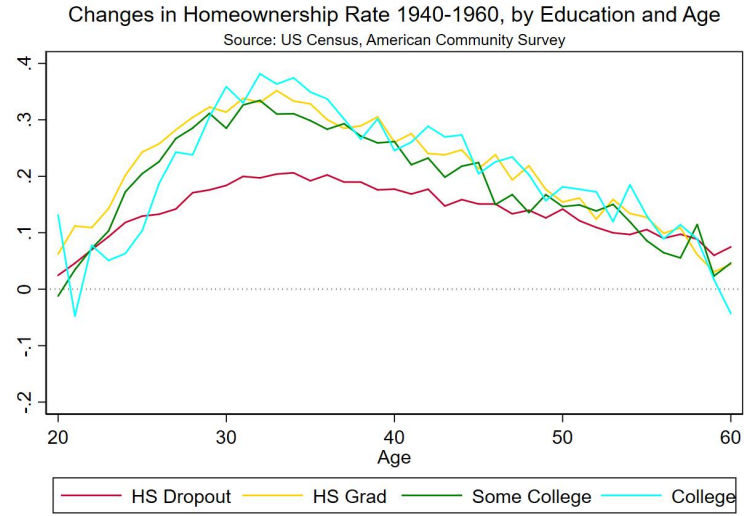
Figure 13. CREDIT EXPANSIONS IN HISTORY. *Source:* Decennial Census and ACS.

| AHS Surveys | Initial Units | Remaining Units | Total Losses | | Conversions to Rented or Vacant | | Losses Other than Conversions to Rented or Vacant | | | | | | |
			Number	As % of Init. Units	Number	As % of Init. Units	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.
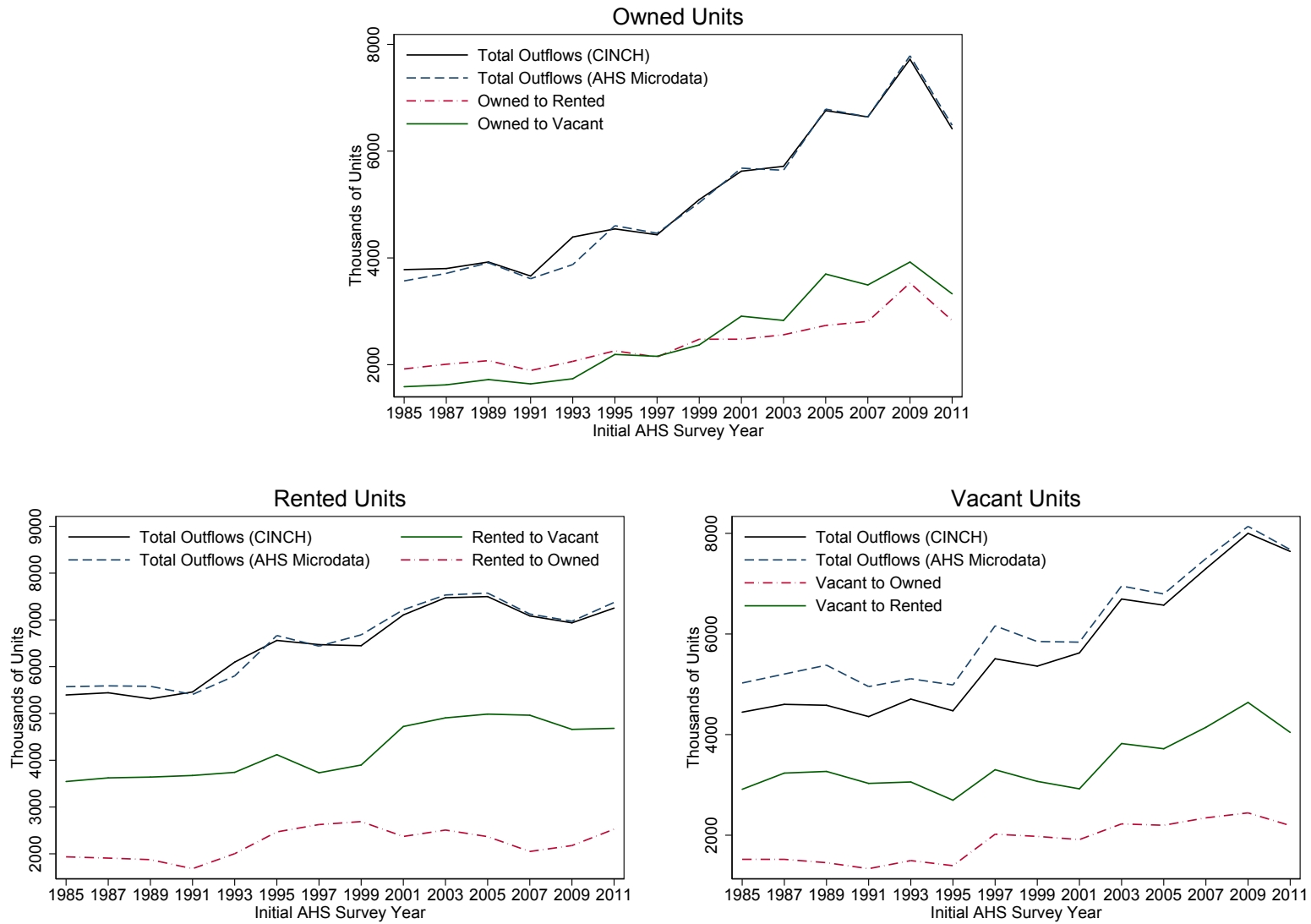
Figure 14. Gross Flows of Housing Units out of Owner-Occupied, Rented, and Vacant Status. *Note:* The black solid lines in each panel display total outflows according to CINCH, while the dashed blue lines depicts our estimates of these series. The estimates are not in perfect agreement, due in part to our use of beginning-year weights only when constructing flows; CINCH combines a sample properties weights from both the beginning and ending years. *Source:* AHS.

Period	Gross Flow: Rented to Owned	Gross Flow: Owned to Rented	Difference: Net Flow from Rented to Owned
1985-87	968	960	7
1987-89	955	1,004	-50
1989-91	938	1,038	-100
1991-93	841	946	-104
1993-95	1,002	1,031	-29
1995-97	1,235	1,130	104
1997-99	1,313	1,072	241
1999-01	1,345	1,239	106
2001-03	1,186	1,239	-53
2003-05	1,254	1,280	-26
2005-07	1,184	1,368	-184
2007-09	1,025	1,406	-381
2009-11	1,088	1,765	-677
2011-13	1,264	1,414	-151

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:* AHS and Bureau of the Census.

Net Change in Owner-Occupied Units

Flows into Owner-Occupied Status
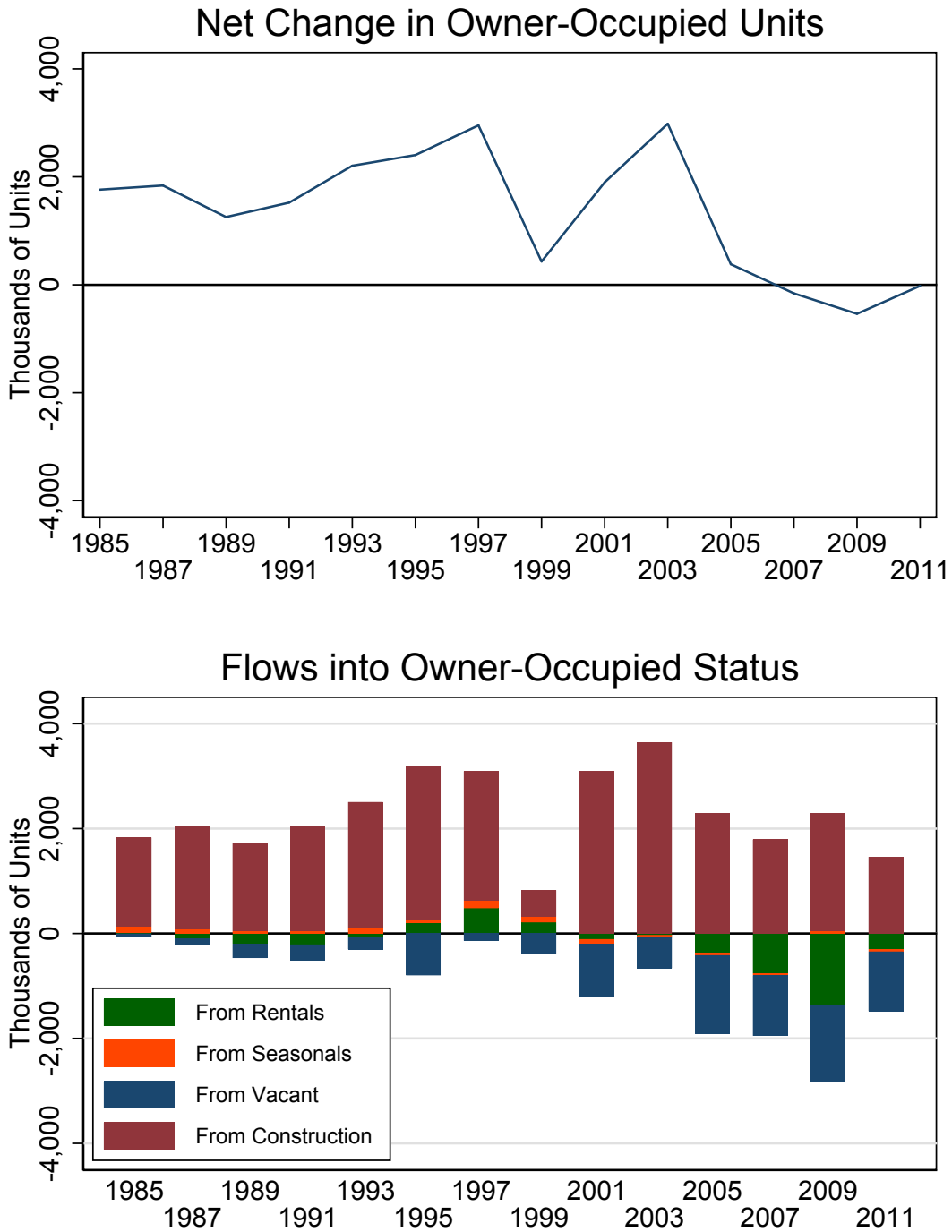
Figure 15. NET AND GROSS FLOWS OF OWNER-OCCUPIED HOMES. *Note:* In the top panel, the significant deceleration of owner-occupied growth in the two-year period after 1999 is a figment of the data; unfortunately, there is a break in the CPS/HVS series for owner-occupied and renter-occupied homes due to a change in sample weights in that year. In the bottom panel, the height of the stacked bar above the y-axis, less the height of the bar below this axis, equals the net change in owner-occupied units depicted in the upper panel. The green bars in the lower panel depict the net flow of existing homes from rented to owned status, that is, the gross flow of rented-to-owned homes less the gross flow in the other direction. *Source:* AHS Microdata and Bureau of the Census.

A   Internet Appendix

A.1   Identifying Piggyback Loans in HMDA

The procedure for identifying piggyback loans in HMDA has 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 the following:

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 the borrower will be occupying the property or not.

7. The income, race, and sex of the borrower.

The second stage is similar to the first, but in place of some of the loan-level characteristics, it 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 variey 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 10th, 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, 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
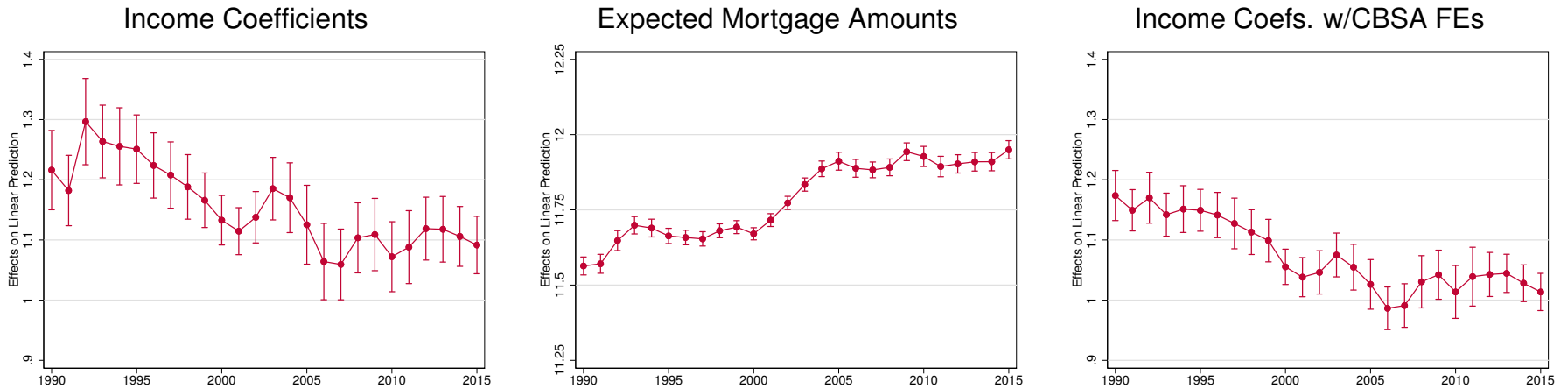
2

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.3 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.
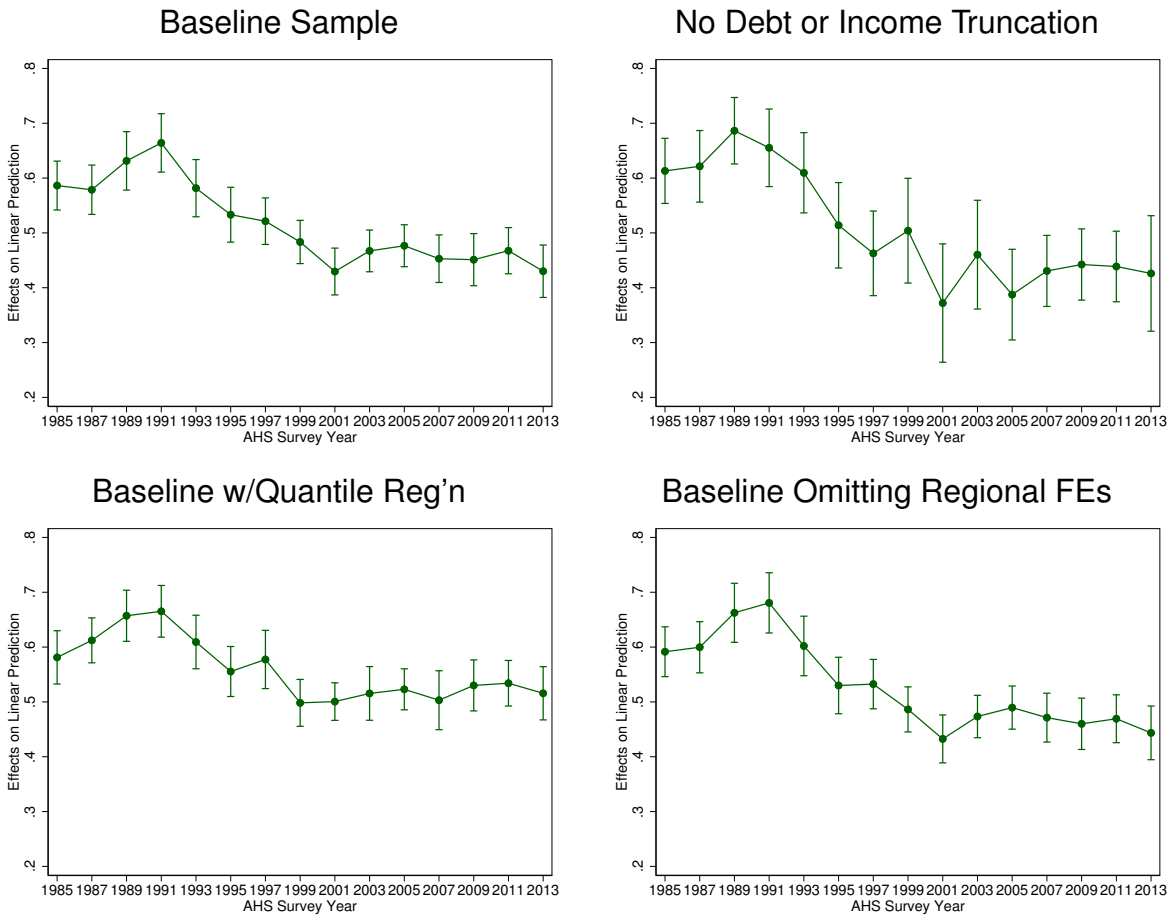
Ordinary Least Squares

Income Coefficients

Expected Mortgage Amounts

Income Coefs. w/CBSA FEs
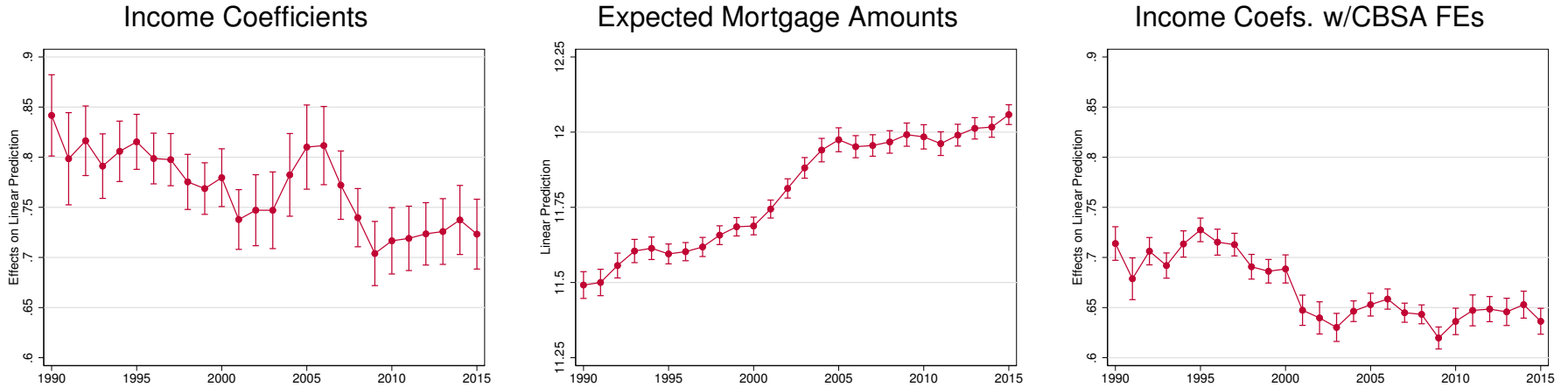
Using Tract-Level Census Income as an IV for Individual-Level HMDA Income

Income Coefficients

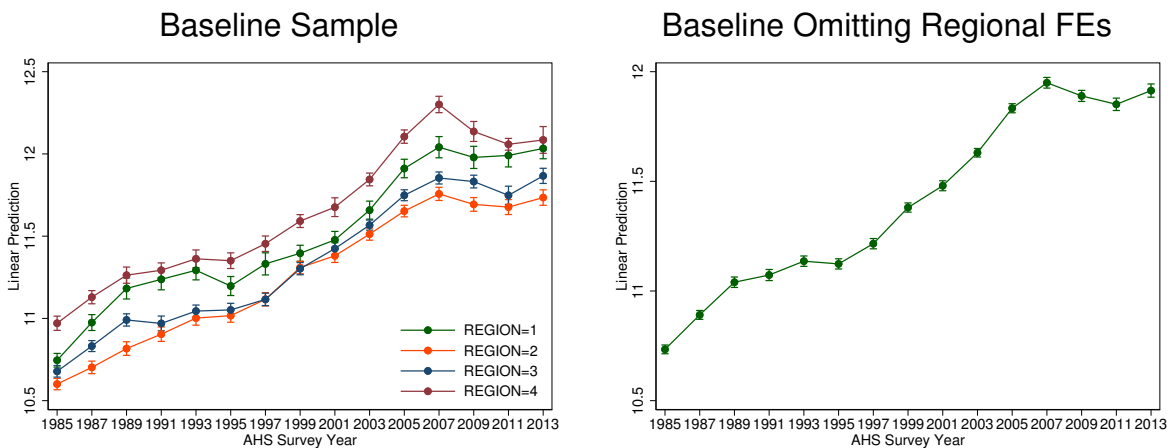### Expected Mortgage Amounts

Income Coefs. w/CBSA FEs

Figure A.1. 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.
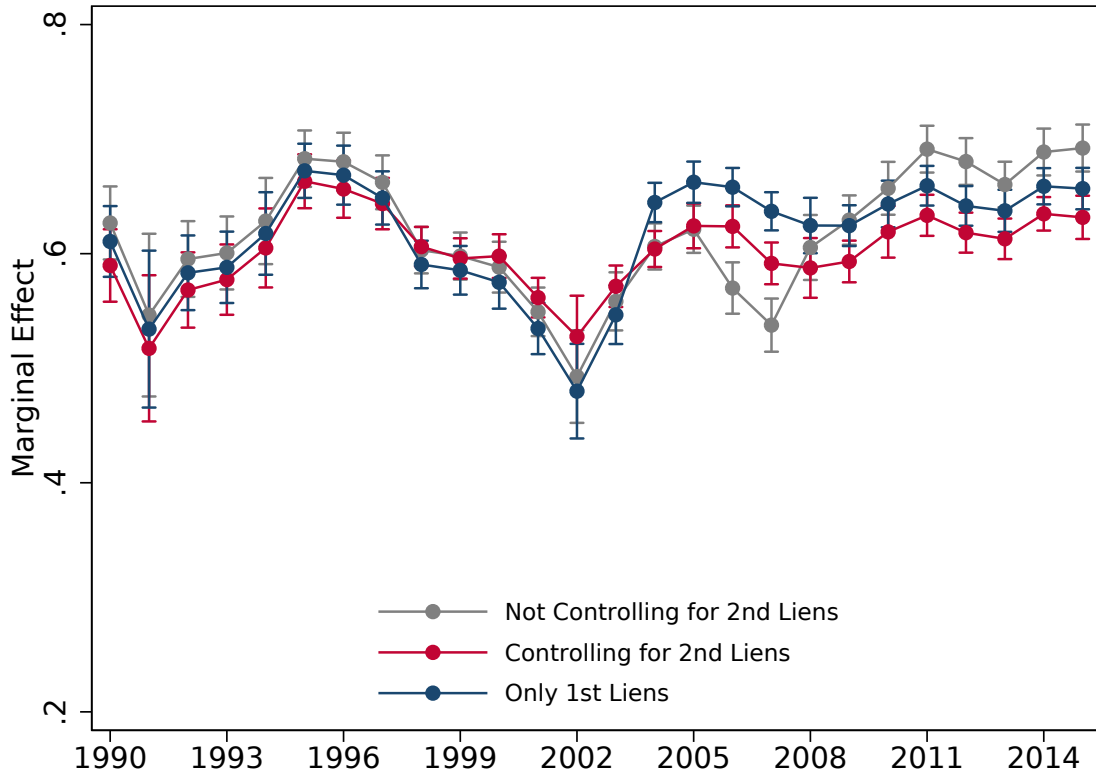
Income Coefficients

Baseline Sample

No Debt or Income Truncation

Baseline w/Quantile Reg'n

Baseline Omitting Regional FEs

Expected Mortgage Amounts

Baseline Sample

Baseline Omitting Regional FEs

Figure A.2. CANONICAL REGRESSION USING INDIVIDUAL-LEVEL AHS MORTGAGE BALANCES AND INCOME LEVELS. *Note:* More here . *Source:* AHS.

Figure A.3. 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 Author's algorithm to identify second liens back to 1990, and correct the loan amount to account for both mortgages. *Source:* HMDA.

	AGI Growth Rate	2006 AGI Level	Accounting for 2nd Liens in HMDA			
			AGI Growth Rate	2006 AGI Level	Wage & Salary Growth Rate	2006 Wage & Salary Level
Model 1						
HMDA Purchase Growth Rate	–.21**		–0.21**		0.23**	
	(0.09)		(0.09)		(0.11)	
Model 2						
HMDA Purchase Growth Rate	0.65***	–0.32***	0.65***	–0.32***	0.84***	–0.33***
	(0.08)	(0.02)	(0.08)	(0.02)	(0.11)	(0.03)
Model 3						
HMDA Growth Rate: Avg. Purchase Mortgage Amount	0.25***	–0.03***	0.23***	–0.08***	0.27***	–0.08***
	(0.03)	(0.01)	(0.03)	(0.01)	(0.04)	(0.01)
HMDA Growth Rate: No. of Purchase Mortgages	0.40***	–0.29***	0.41***	0.23***	0.57***	–0.25***
	(0.07)	(0.02)	(0.07)	(0.02)	(0.10)	(0.02)

Table A.1. Decomposing the Negative Correlation Between Growth in Purchase Mortgages and Income Growth: 2002–2006. *Note:* Standard errors are included in parentheses. Model 1 replicates the negative sign of the correlation between purchase-mortgage growth and AGI growth reported in Mian and Sufi (2009). Model 2 adds the level term to the regression to show that relationship between purchase-mortgage originations and AGI was never negative in levels. Model 3 follows Adelino, Schoar, and Severino (2016) by dividing purchase-mortgage originations into the average size of each purchase mortgage and the number of purchase mortgages. This regression shows a significant decline in the slope of the positive relationship between purchase-mortgage originations and income only when considering the number of purchase mortgages, not the average purchase amount. *Source:* HMDA and IRS Statistics of Income.
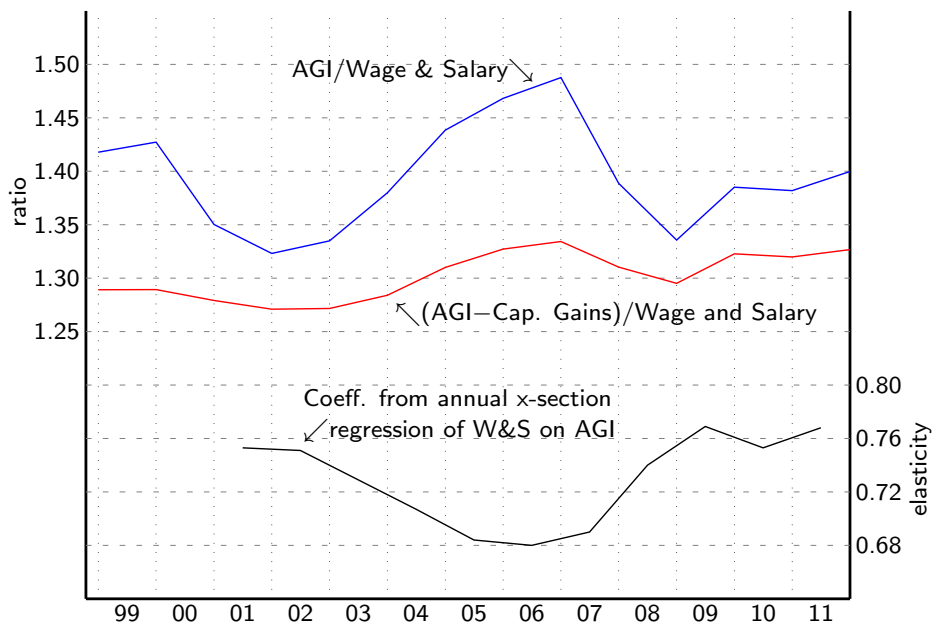
Figure A.4. THE RELATIONSHIP BETWEEN ADJUSTED GROSS INCOME (AGI) AND WAGE AND SALARY INCOME (W&S). *Note:* The blue line at top is the ratio of aggregate AGI to aggregate wage and salary income. The red line is the same ratio after capital gains are excluded from AGI. The black line at bottom depicts coefficients from yearly cross-sectional regressions of wage and salary income on AGI at the zip code-level. Aggregates are generated by summing zip code-level data. *Source:* IRS Statistics of Income.