Federal Reserve Bank of Dallas

Credit and Housing Price Effects of Automated Underwriting Adoption

Stephanie Johnson and Nitzan Tzur-Ilan

Working Paper 2506 **January 2025 (revised May 2026)**

Research Department

https://doi.org/10.24149/wp2506r2

Working papers from the Federal Reserve Bank of Dallas are preliminary drafts circulated for professional comment. The views in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

Credit and Housing Price Effects of Automated Underwriting Adoption[*]

Stephanie Johnson[†] and Nitzan Tzur-Ilan[‡]

November 2024
Revised: April 2026

Abstract

We study how the 1990s adoption of now widely-used automated mortgage underwriting systems affected credit, house prices and their comovement across locations. The effects go well beyond processing improvements. By implementing more complex, statistically-informed lending rules, the systems allowed households to borrow more, pushing up house prices. Furthermore, by transmitting a common set of credit standards across lenders, the new technology increased house price synchronization. Together, our results illustrate how new lending technology can generate systematic credit supply shocks, influencing house prices and increasing market interconnectedness.

JEL Classification: G21, L85, R21, R31

Keywords: mortgage credit, financial technology, automated underwriting, housing prices, credit supply, house price comovement

[*] We would like to thank Brent Ambrose, Scott Baker, Neil Bhutta, Alex Butler, Larry Cordell, Anthony DeFusco, Matthias Doepke, Martin Eichenbaum, Scott Frame, Shan Ge, Dan Greenwald, Adam Guren, Lorenz Kueng, Lara Loewenstein, John Mondragon, Eva Steiner, and James Vickery as well as seminar participants at the Federal Reserve Bank of Dallas, Duke University, Pennsylvania State University, Rice University, UT Austin, the 2020 Meeting of the Urban Economics Association, the 2021 AREUEA National Conference, the AREUEA-ASSA 2023, the U.S. Department of Justice Civil Rights Division, and the AFA-ASSA 2025. The contents of the paper are solely our responsibility. The views expressed are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Dallas or the Federal Reserve System.

[†] Stephanie Johnson, Jones Graduate School of Business, Rice University, 6100 Main Street, Houston, TX 77005. E-mail: stephanie.g.johnson@rice.edu.

[‡] Nitzan Tzur-Ilan, Federal Reserve Bank of Dallas, 2200 N Pearl St, Dallas, TX 75201. E-mail: nitzan.tzurilan@dal.frb.org.

1 Introduction

In the U.S. mortgage market, the 1990s introduction of automated underwriting systems was associated with a shift to standardized, statistically-driven rules that are now used for the majority of applications. In 2023, around two-thirds of U.S. mortgage originations were underwritten using Fannie Mae or Freddie Mac's systems. This shift was not simply a processing improvement. It coincided with a change in lending standards that was associated with meaningful shifts in borrowing capacity and house prices.

We exploit variation in the system choice of early adopters driven by lenders' pre-existing relationships with the government-sponsored enterprises to implement a difference-in-differences approach. Our setting mitigates concerns relating to endogenous adoption timing and allows us to distinguish between two dimensions of automation: processing efficiency and the introduction of new, statistically-informed underwriting rules. We show that the adoption of a system incorporating new rules relaxed borrowing constraints across the income distribution, increasing loan amounts relative to income and raising house prices.

This paper relates to a large literature studying the causes of the 2000s housing boom, in particular the relative roles of credit supply expansion and house price expectations (Mian and Sufi, 2009; Keys, Mukherjee, Seru and Vig, 2010; Adelino, Schoar and Severino, 2016; Foote, Loewenstein and Willen, 2020; Griffin, Kruger and Maturana, 2021; Albanesi, DeGiorgi and Nosal, 2022). Much of this literature exploits cross-sectional patterns in borrowing, such as whether credit expands disproportionately among lower-income households. In the absence of a clearly identified shock to credit supply, these patterns are often used to distinguish between competing explanations, with increases in borrowing by middle and high income borrowers interpreted as evidence more consistent with an expectations-driven channel.

We contribute to this debate by identifying a specific and substantial shift in lending standards occurring in the second half of the 1990s: the gradual adoption of automated underwriting systems that introduced new underwriting rules. We show that this shift increased borrowing across the income distribution and estimate the effect on house prices. While the initial source of variation is a broad-based change in lending standards, the estimated response may still reflect some feedback through expectations.

To assess the aggregate importance of this mechanism, we combine our difference-in-differences estimates with the aggregate time series of automated underwriting adoption in the late 1990s. Given rapid adoption in the 1998 refinancing boom and the magnitude of credit expansion, it is possible that this technological shift could have played a role in the early stages of the housing boom.

Finally, we suggest that automated underwriting systems can affect the geographic transmission of credit conditions. Because a small number of systems apply common lending rules across different lenders, updates to underwriting standards can propagate rapidly across markets. Consistent with this mechanism, we find that adoption increases the comovement of house prices across locations. This highlights a channel through which financial technology can increase market interconnectedness and potentially systemic risk.

Our empirical strategy leverages key institutional differences between the two dominant systems. Freddie Mac's Loan Prospector (LP) system implemented new statistically-informed underwriting rules at rollout (Maselli, 1994; Straka, 2000), whereas early versions of Fannie Mae's Desktop Underwriter (DU) system primarily automated existing manual guidelines. This distinction mitigates concerns relating to endogenous adoption timing when we make comparisons across systems within the set of early adopters. We collect the names of initial adopters from contemporaneous news articles and trade publications, and match these lenders to loan-level data. Trade journals also provide qualitative evidence on how the systems were used in practice. Importantly, lenders' choice of system was likely determined by pre-existing selling relationships with Fannie Mae or Freddie Mac, rather than borrower composition or local market conditions. This institutional feature helps us address concerns about endogenous selection into a particular system and provides a source of plausibly exogenous variation in exposure. We focus on the period immediately following adoption, before subsequent changes to DU also altered its underwriting rules.

Contemporaneous accounts indicate that the introduction of statistically-informed underwriting rules substantially increased the maximum loan size available to many borrowers, primarily by reducing the emphasis on traditional debt payment-to-income constraints. Consistent with this, we find an increase in high loan-to-income lending among early adopters of Freddie Mac's LP system relative to adopters of Fannie Mae's DU within the same ZIP code. Using Home Mortgage Disclosure Act data, we estimate that LP adoption increases high loan-to-income lending by about 10 percent, with effects that are broad-based across the income distribution.

We next estimate the effect on house prices using variation in ZIP code exposure to early LP adopters. A one standard deviation increase in LP market share leads to a cumulative increase in house prices of 1.8 percent over a three-year horizon. We interpret this as the total effect of a shift to statistically-informed underwriting standards. While part of the response could reflect expectations-driven feedback (Armona, Fuster and Zafar, 2019; Bailey, Cao, Kuchler and Stroebel, 2018; Case, Shiller and Thompson, 2012), our empirical design isolates a setting in which the initial impulse is a clearly identified change in lending standards, allowing us to attribute the resulting price dynamics to a credit supply shock.

In addition to effects on price levels, automated underwriting also has implications for the spatial transmission of credit conditions. When lenders rely on a common underwriting system, their lending standards—and subsequent updates to those standards—become synchronized across markets. We test this idea by constructing measures of lending integration between county pairs that capture shared exposure to underwriting systems (Landier, Sraer and Thesmar, 2017). We find that greater exposure to common systems increases house price comovement across locations, consistent with the propagation of common credit supply shocks.

Finally, we provide, to our knowledge, the first quasi-experimental evidence on the effect of automated underwriting on loan processing times. We find that adoption leads to a modest reduction in processing time: the average time from application to closing declines by approximately four days, with larger effects for denied applications. This relatively small efficiency gain, compared to the sizable effects on borrowing and house prices, suggests that the primary impact of automated underwriting operated through lending rules rather than processing improvements.

Our paper contributes to two main strands of the literature. First, we contribute to the mortgage and housing literature that studies the relationship between lending standards and house prices, often in the context of explaining the 2000s housing boom. One set of explanations emphasizes an expansion of credit to lower-income borrowers (Mian and Sufi, 2009; Keys et al., 2010; Demyanyk and Van Hemert, 2011). The credit expansion we study has broader effects across the income distribution. Other work notes that credit growth and subsequent defaults were not concentrated among lower-income borrowers (Adelino et al., 2016; Albanesi et al., 2022), suggesting a potential role for expectations (Foote et al., 2020; Kaplan, Mitman and Violante, 2020). This work suggests outcomes such as the distribution of credit growth across households and locations can be used to infer something about the initial driver.

Our results complement these perspectives by identifying a specific shift in lending standards in the 1990s and estimating a house price response to this shock. Our findings suggest that broad-based credit growth can also be consistent with a lending standards channel. Davis, Larson, Oliner and Smith (2023) document that mortgage risk had already increased during the 1990s. Greenwald (2018), Greenwald and Guren (2025), and Kaplan et al. (2020) model the effects of a relaxation in payment-to-income limits during the late 1990s, but do not provide a fundamental explanation for this shift.

We propose that the gradual adoption of automated underwriting systems during the late-1990s provides a natural mechanism for these observed changes in lending standards. Combining our difference-in-differences estimates with aggregate system usage data, we suggest that this technological shift is of sufficient magnitude to matter for national price growth in the early stages of the housing boom. The magnitude of our implied aggregate effect aligns closely with Greenwald (2018), who calibrates the impact of a relaxation in debt-to-income limits using observed changes in the distribution of household leverage. Overall, our results reinforce the view that expansions in lending standards can have sizable effects on house prices.[1]

[1] Other related work includes Favilukis, Ludvigson and Van Nieuwerburgh (2017); Justiniano, Primiceri and Tambalotti (2019); Favara and Imbs (2015); Acharya, Bergant, Crosignani, Eisert and McCann (2022); Di Maggio and Kermani (2017); Adelino, Schoar and Severino (2025); Loutskina and Strahan (2015); Johnson (2020); Griffin et al. (2021). Our 1990s empirical setting predates much of the reduced-form work examining the housing boom. Our sample period overlaps with Favara and Imbs (2015), who argue that bank branching deregulation, which specifically affected commercial banks, was also a driver of house price growth in

Second, we contribute to a growing literature examining how technological innovations shape credit allocation. Prior work shows that advances in statistical modeling (Fuster, Goldsmith-Pinkham, Ramadorai and Walther, 2022), the use of alternative data sources for thin-file borrowers (Berg, Burg, Gombović and Puri (2019); Lee, Yang and Anderson (2025); Di Maggio, Ratnadiwakara and Carmichael (2022); Blattner and Nelson (2021)), and reductions in information frictions (Petersen and Rajan (2002); Jiang and Zhang (2025)) and biases (Howell, Kuchler, Snitkof, Stroebel and Wong, 2024) can affect credit access and allocation. Related work studies the performance of automated underwriting relative to human underwriters (Jansen, Nguyen and Shams, 2025; Khoei, 2025).

In our setting, statistically-informed underwriting rules expand access to credit but also increase demand for housing, thereby raising asset prices when supply is inelastic. As a result, improvements in access and affordability may be partially offset by higher prices. We also suggest that automated underwriting technology can increase the comovement of lending and house prices across locations, as lenders converge on similar decision rules. This mechanism is related to the idea that banking integration increases regional comovement (Landier et al. (2017)), but highlights that common adoption of underwriting systems can generate similar effects to the emergence of large national lenders.

A related literature studies the impact of fintech on convenience and processing speed in lending and real estate transactions (Fuster, Plosser, Schnabl and Vickery (2019); Buchak, Matvos, Piskorski and Seru (2018); Berg, Fuster and Puri (2022); Buchak, Matvos, Piskorski and Seru (2020)).[2] We provide evidence that automated underwriting modestly reduces mortgage processing times, with the largest effects arising from faster denials.

the late 1990s and early 2000s.

[2]The effect of automated underwriting on processing is a related but distinct question, as such systems are also used by lenders that do not offer fully online applications.

Finally, several papers examine the broader real effects of mortgage-related technologies, including Gao, Yi and Zhang (2023), Lewellen and Williams (2021), and Foote, Loewenstein and Willen (2019). Foote et al. (2019) provide a detailed narrative of underwriting innovation and adoption during the 1990s, arguing that the aggregate relationship between mortgage size and income weakened over this period and may have implied relatively greater expansion for lower-income households. They conclude that underwriting technology did not materially affect house prices, based in part on the assumption of immediate adoption following its introduction and the observed absence of differential price growth in lower-income areas. We document that the credit expansion associated with automated underwriting adoption was broad-based across the income distribution, suggesting that its effects would not necessarily be concentrated in lower-income markets. After accounting for the gradual adoption of the technology by lenders, the rise in system usage also aligns more closely with the profile of national house price growth.

2 Institutional background

2.1 Early automated underwriting systems and their usage

In the early 1990s, mortgage underwriting in the United States remained largely manual. Loan officers and underwriters evaluated applications using guidelines set by lenders and secondary-market participants, including Fannie Mae and Freddie Mac. This process created operational challenges when application volumes rose, as during refinancing booms, because lenders had to expand underwriting capacity quickly (Straka, 2000). Manual underwriting may also have constrained the time available for more difficult files. These frictions arguably increased interest in automated underwriting, which was expected to improve efficiency for both lenders and borrowers.

Automation also made it easier to implement more complex underwriting rules. Traditionally, underwriting standards had reflected lenders' experience, such as the observed poor performance of loans with low down payments (Straka, 2000). Over time, greater standardization, better data, and improved computing power made it possible to estimate mortgage risk using more sophisticated statistical methods. Automated systems facilitated the application of such rules at scale without sharing proprietary algorithms. Lenders could therefore use external underwriting models without directly observing the underlying decision rules.

Freddie Mac was an early leader in analysis of loan performance, and it incorporated new statistically-informed rules into its LP system. A pilot program involving a number of lenders began in early 1994, and LP was publicly released in 1995. Fannie Mae's DU system followed a similar rollout timeline. However, although DU had the potential to improve processing efficiency, its early versions were not initially based on statistical models. Instead, DU largely automated Fannie Mae's existing manual underwriting guidelines (Straka (2000); Nixon (1995)).[3,4]

During our sample period, DU and LP were used primarily to determine loan eligibility rather than to provide lenders with measures of borrower risk (Temkin, Johnson and

[3]Some large lenders and mortgage insurers also developed systems around this time. The Countrywide Loan Underwriting Expert System (CLUES) was one of the earliest systems used on a large scale and was rolled out in 1993. The rules used by CLUES were developed not through statistical analysis, but by observing the decisions of expert underwriters. Countrywide started developing the system in 1991 with the primary goal of increasing the number of loans per employee – not changing lending standards (Talebzadeh, Mandutianu and Winner, 1995). PMI Mortgage Insurance Co. had been working on its Automated Underwriting Risk Analysis (AURA) system since the 1980s. Unlike CLUES, the system was statistically based and generated a risk score between 1 and 100 (Mikel and Baker, 1992).

[4]The early 1990s was a period of relaxation of lending standards. One example was the "GSE Act", a 1992 law mandating that the GSEs help promote credit access and homeownership opportunities for low-income households. For more information, see (Bhutta, 2009).

Levy, 2002).[5] For most conventional applications, LP returned one of only two recommendations: "accept" or "caution." Loans receiving a "caution" recommendation could still be found eligible, but only after manual underwriting.[6] Risk-based pricing remained uncommon during the 1990s. Most lenders used average-cost pricing within the prime market, while specialized subprime lenders charged higher rates to a pool of riskier borrowers. These subprime lenders initially showed limited interest in Fannie and Freddie's systems, which were seen as less relevant for high-risk segments. Temkin et al. (2002) discuss the GSEs' later movement into subprime lending and the potential for automated systems to support risk-based pricing, but these developments occurred after our sample period.

Although DU and LP were created by the GSEs, they were marketed as general underwriting tools and were often used for loans that lenders did not intend to sell to Fannie Mae or Freddie Mac. Some lenders reported running all applications through one of the systems and then manually underwriting those that were not accepted (LaMalfa (1998); LaMalfa (1999); Jones (1997)). Fannie Mae and Freddie Mac's guidelines had long served as an industry benchmark, not only for conforming loans but also for portfolio lending and loans sold elsewhere in the secondary market. This gave DU and LP an important competitive advantage. As Dennis and Robertson (1995) note,

"To a great extent, the underwriting guidelines of both Fannie Mae and Freddie Mac

[5] Freddie Mac was also active in promoting the use of FICO scores around the same time, publishing a study in 1992 showing that general FICO scores had substantial predictive power for mortgage default. While general credit scores ultimately became input into Loan Prospectors' recommendations, Freddie Mac also strongly recommended the use of credit scores to lenders using manual underwriting. The take-up of credit scores as an underwriting input during this period was, therefore, not unique to lenders using automated underwriting systems. Fannie Mae also soon followed suit in recommending the use of FICO scores in manual underwriting (Pierzchalski, 1996).

[6] For government loans (i.e. FHA or VA loans) Loan Prospector generated either an "accept" or "refer" recommendation. Loans receiving a "refer" recommendation also needed to be manually underwritten in order to be eligible (Temkin et al., 2002).

> *are the core standards that most lenders attempt to follow. Even those lenders who don't intend to sell loans to these two secondary mortgage market players should attempt to follow these well-conceived underwriting guidelines." (pp. 116-117)*

The ability to certify automatically that a loan satisfied these standards was therefore valuable. In practice, an "accept" recommendation from one of the GSE systems could be interpreted as a signal that a loan met prime underwriting criteria. Competing systems could underwrite loans using Fannie or Freddie's published manual rules, but the ability to sell certain loans depended on compliance with the GSEs' proprietary automated rules, and therefore on use of their own systems.

2.2 Effect on lending standards

Freddie Mac's statistical approach to underwriting had the potential to expand credit access without a corresponding increase in default risk (Maselli, 1994). In principle, statistically-informed underwriting could broaden access by setting cutoffs more efficiently and by allowing for richer interactions across borrower characteristics. Automation made these interactions easier to implement consistently and allowed the underlying rules to remain proprietary.[7] Unlike Freddie Mac's public manual rules, LP relied on a proprietary algorithm. As a result, evidence on how lending standards changed comes in part from contemporaneous accounts by lenders using the system.

These accounts suggest that one important change was a relaxation in the allowable ratio of debt payments to income, conditional on other risk factors. Lenders participating in the 1994 LP pilot were already reporting such an expansion (Maselli, 1994). Harney (1996) describes cases in which LP accepted debt-to-income ratios as high as 72 percent,

[7] More information on the performance of scorecards can be found in Straka (2000), Gates, Perry and Zorn (2002) and Foote et al. (2019).

at a time when manual underwriting guidelines typically limited debt-to-income ratios to below 36 percent, subject to some discretion (Maselli (1994), Irwin (1992)). The back-end debt-to-income ratio is defined as the ratio of monthly debt payments and other financial obligations to gross monthly income.[8]

Both Harney (1996) and Maselli (1994) indicate that in the early years, LP approvals at high debt-to-income ratios were generally limited to borrowers with offsetting factors, such as strong credit scores, large down payments, or substantial cash reserves. By 1999, however, both Fannie Mae and Freddie Mac appear to have permitted high debt-to-income ratios in a much broader set of cases. Appendix Figure A.1 shows the debt-to-income distribution of loans acquired by the two GSEs in 1999, the earliest year for which these data are available. Although mortgage data from the mid-1990s are not sufficiently detailed to recover the exact change in underwriting standards, the evidence presented in Section 4 is consistent with a substantial relaxation in debt-to-income limits.

2.3 Adoption timing

Table A.1 reports usage of DU and LP over time. Adoption was initially limited: in the first half of 1997, less than one quarter of eligible loans were processed through either system (Foster, 1997). Usage then rose sharply in 1998, coinciding with a refinancing boom. Historically, such episodes required lenders to hire large numbers of underwriters to process increased demand, making automation particularly valuable during these periods. Consistent with this, Talebzadeh et al. (1995) note that Countrywide's development of

[8]For context, around 7 per cent of 2018 HMDA applications had a back-end debt-to-income ratio above 60 per cent (the ratio is top-coded at 60 per cent), with most being denied. We expect that a limit of 72 per cent would not be binding in most cases. In contrast, around 58 per cent of 2018 HMDA applications had a debt-to-income ratio above 36 per cent. These statistics are based on the subset of HMDA applications for which the debt-to-income ratio is known. We use 2018 data as information on the debt-to-income ratio was not collected for prior years.

CLUES had been motivated by a "serious shortage of qualified underwriters" during an earlier refinancing boom.

Trade publications suggest several reasons why adoption was gradual. Lenders emphasized that the systems were costly to implement and that the benefits would not be immediate. A representative of Flagstar Bank explained that "It isn't cheap: there are transaction costs, equipment costs, training costs. And there's a learning curve. The efficiencies are starting to materialize now" (LaMalfa, 1996). A representative of InterFirst similarly stated, "We love LP, but it's still not cost-effective" (LaMalfa, 1997). Another lender remarked that, after licensing and usage fees, Loan Prospector "doesn't appear to net any cost saving," while also acknowledging that "the Freddie Mac and Fannie Mae processes can ultimately decrease your cost in volatile periods" (LaMalfa, 1996). In November 1995, the per-loan cost of an LP approval was approximately $400 (Sullivan, 1995).[9] Integration costs appear to have been another barrier. Fannie and Freddie relied on proprietary data standards (Markus, Dutta, Steinfield and Wigand, 2008), and Oliver and McDonald (1997) report that lenders "did not make full use of AU systems (i.e., use AUS at the point of sale) owing in part to lack of integration with back-end systems".

A further impediment was an absence of reciprocity between the two GSEs. Freddie Mac would not accept DU decisions, and Fannie Mae would not accept LP decisions. Lenders seeking to compare execution across the two investors would therefore need to run loans through both systems and incur both sets of fees (Foster, 1997). Many lenders viewed the use of multiple systems as too costly. As a result, they either chose one system or continued to rely on manual underwriting, sometimes supplemented by alternative

[9] Sullivan (1995) provides some pricing details disclosed by Fannie and Freddie, but it is not possible to do a full cost comparison. Although Freddie's per loan fees were slightly higher, Fannie charged a number of additional fees that were not disclosed.

technologies (LaMalfa, 1997). More broadly, perceived switching costs and the fact that each system effectively tied a lender more closely to one secondary-market buyer helped generate persistent relationships with either Fannie Mae or Freddie Mac (DeMuth (1999); Johnson (2020)). These same frictions may also have encouraged some lenders to delay adoption strategically.

According to Fannie Mae and Freddie Mac's annual reports, reported usage of DU and LP stabilized at just over 60 percent in 2001 (Table A.1). This likely understates the prevalence of both automated underwriting and more relaxed lending rules. Beginning in the late 1990s, both GSEs entered into agreements with some large sellers to purchase loans underwritten with alternative systems. Acquisition data available from 1999 onward indicate that loans with high debt-to-income ratios continued to be purchased from these sellers even when alternative systems were used. Before such agreements were in place, Freddie Mac had projected that LP usage would eventually stabilize at 80 to 85 percent. Reported usage of DU and LP among larger community banks in 2004 was also around 85 percent, consistent with those earlier forecasts (Costanzo, 2004).

3 Data and descriptive statistics

We use mortgage lending data from the Home Mortgage Disclosure Act (HMDA). HMDA provides fairly comprehensive coverage of the U.S. mortgage market, particularly for properties located in metropolitan areas. We also use confidential supervisory data collected under the HMDA to construct a measure of processing time based on application dates and closing or decision dates. The confidential dataset contains more comprehensive information than the public version of the data. To study the effects on house prices, we use FHFA ZIP code-level house price data.[10]

3.1 Lender classification

Table 1 lists the initial users of Loan Prospector and Desktop Underwriter. These lenders were already using the systems at the time of their public release in 1995. We identify the HMDA IDs associated with these lenders based on the names and locations reported in the source articles. We also use NIC data on institutional relationships and transformations to ensure that the set of HMDA IDs we assign to each lender captures as consistent an institution as possible over our sample period.

3.1.1 Comparing initial LP adopters with initial DU adopters

Because DU initially applied Fannie Mae's manual rules, we can use the initial users of DU as a control group to quantify the effect of adopting statistical lending standards on credit and house prices. Since both groups of lenders adopted an automated system at the same time, we mitigate selection concerns related to the choice to adopt a system early. Remaining selection concerns therefore relate to the choice of LP *relative to DU*, rather than adoption timing. We show that the choice between LP and DU was likely driven by lenders' pre-existing relationships with Fannie or Freddie, rather than by anticipation of the different rules applied by LP. To further address selection concerns, we also show that our estimates are robust to using pre-existing selling relationships as an instrument for system choice.

Table 2 reports estimates from a linear probability model relating a lender's system choice to lender characteristics. Each variable is normalized by dividing by its standard

[10] We also explored other house price indices, but found them to be less reliable and to have lower effective coverage than the FHFA house price indices during the 1990s.

deviation. The main difference between initial LP and DU users is that lenders choosing LP sold a much larger share of their loans to Freddie prior to the release of the two underwriting systems. This is consistent with research conducted by Mortech in 1996, which "revealed that AU decisions are primarily based on which GSE the lender does the most business with" (Strickberger, 1999).[11]

3.1.2 Comparing initial LP and DU adopters with a matched control group

Because processing improvements could apply to both initial LP and DU users, we use a set of matched control lenders for comparison. The matching procedure targets the following variables: loan purchases as a share of total purchases and originations; refinance originations as a share of total originations; and portfolio originations (loans not sold in the year of origination) as a share of total originations. We also require matched lenders to fall in the same broad size class based on total loans. We use HUD's classification of prime and subprime lenders to ensure that matched lenders are of the same type. We match without replacement and select matches based on Mahalanobis distance.

Table A.2 compares DU and LP users with matched control lenders. Initial DU users tend to sell a larger share of loans directly to Fannie or Freddie than control lenders, though they sell a somewhat smaller share of loans overall, conditional on other variables. In the case of LP users, there is no significant relationship with any of the lender characteristics.

[11]Consistent with this, a representative of Fleet Mortgage stated that: "It is impractical for us to have two AU systems" and "we elected to go with DU first. We typically sell about 60 per cent of our business to Fannie Mae. That had a lot to do with it." (DeMuth, 1999).

3.2 ZIP exposure and statistics

To estimate the house price response, we use variation in ZIP code exposure to the initial adopters of Loan Prospector:

$$LP_z = \frac{\text{\# Loans reported in ZIP code } z \text{ by LP lenders (Table 1; Column 1)}}{\text{\# Loans reported in ZIP code } z \text{ by all HMDA reporters}} \tag{1}$$

We also condition on total exposure to early adopters of either Loan Prospector or Desktop Underwriter:

$$AUS_z = \frac{\text{\# Loans reported in ZIP code } z \text{ by all lenders in Table 1}}{\text{\# Loans reported in ZIP code } z \text{ by all HMDA reporters}} \tag{2}$$

Conditioning on AUS_z is consistent with the loan-level analysis, where we compare loans made by initial LP users with those made by initial DU users rather than with loans made by all other lenders. We compute the market shares in (1) and (2) using data from 1993, before lenders began using the systems.

Figure 1 shows the distribution of ZIP code market share held by initial Loan Prospector users, conditional on the control variables included in our regressions. Column 1 of Table 3 shows the relationship between exposure to early Loan Prospector users and ZIP code characteristics. Both dependent and independent variables are normalized by dividing by their standard deviations. There is a mechanical positive relationship with the combined market share of early LP and DU users (AUS_z). We condition on AUS_z to strengthen identification as described above. We also expect a positive relationship with the overall ZIP code share of loans sold to Freddie, given that lender-level system choices were driven by pre-existing relationships. Although it is not obvious that the average Fannie/Freddie relationship across lenders in a ZIP code should generate correlations with other ZIP code outcomes, we conservatively condition on these shares as well. Figure A.2 maps residualized ZIP code market shares of initial LP users in New England.

Although some variables in Table 3 are correlated with exposure, conditional on these variables we find no evidence of differential house price trends prior to adoption. This is consistent with our claim that the additional price growth observed in the mid-1990s is due to the early adoption of new statistical underwriting criteria, rather than to other characteristics of these lenders or the areas in which they operate. For example, if affected lenders were generally more aggressive, one might expect the locations in which they operated to exhibit stronger price growth even before system adoption. We also estimate the house price response using weights that eliminate these correlations across ZIP codes. We define the "treated" group as ZIP codes with values of LP_z in the top quartile. We then apply weights that balance covariates across "treated" and "control" ZIP codes. The dependent variable in Column 2 of Table 3 is an indicator for top-quartile exposure.

To further address any lender selection concerns, we also construct an instrument for LP_z as follows. We replace the numerator in Equation 1 with the number of loans reported in ZIP code z by lenders from Table 1 for whom more than half of their 1992 GSE loans went to Freddie Mac. That is, we compute an exposure measure under the assumption that an early adopter whose primary relationship in 1992 was with Freddie, rather than Fannie, chooses LP rather than DU. Column 3 of Table 3 shows the relationship between the instrument and ZIP code characteristics. To rule out selling relationships operating through another channel, we compute placebo analogues of LP_z and AUS_z using matched control lenders, with assignment to "LP" based on whether more than half of 1992 GSE loans went to Freddie Mac. We do not find a significant house price response when applying this approach to similar lenders that were not early AUS adopters, consistent with the IV operating through AUS selection.

4 Results

4.1 Credit response

In this section, we measure how the new rules rolled out with Loan Prospector influenced lending. Using HMDA, we compare loan characteristics of initial LP users with initial DU users before and after 1994, when these lenders started using the systems. This likely differences out any effect of automation alone. Comparing loans in the same 5-digit ZIP code and income quartile in the same year, we find that initial LP adopters lend more relative to income after adopting the system. Because we compare loans in the same ZIP code and income group at the same point in time, differences in credit outcomes are unlikely to be driven by differences in house price growth.

We focus on home purchase loans originated to owner-occupiers, and follow lending outcomes from 1992 through 1996. We create consistent institutions over the period to account for mergers and acquisitions. Our choice of sample period reflects limitations of the data and setting. We start our sample in 1992 for two reasons. Several early adopters are not present in 1991 HMDA data in a consistent form and 1992 is also the first year to use the 1990 census tract boundaries. A consistent census tract definition is important for the mapping to ZIP codes. We end the sample period in 1996 as DU also started to apply very similar rules to LP around 1997, which would affect the interpretation of our difference-in-differences estimates. An additional limitation is that we do not observe the debt-to-income ratio directly in HMDA. Instead, we observe loan size and income and can therefore compute the loan-to-income ratio. Unfortunately, we also do not observe the interest rate and cannot convert loan size to payment size.[12]

We define $HighLTI_i$ equal to one if the loan-to-income ratio exceeds 2.5 and zero otherwise. We include ZIP-income quartile-year-purchaser fixed effects, and lender-income quartile-purchaser fixed effects. For loan i reported by lender l in income quartile n in zip z with purchaser type p in year t:

$$HighLTI_i = \alpha_{l,n,p} + \gamma_{z,n,p,t} + \beta LP_l \cdot Post_t + \epsilon_i \tag{3}$$

where LP_l is an indicator equal to one for early LP users and zero for early DU users. β is interpreted as the effect of LP adoption on the high LTI loan share. If the loan is sold, purchaser type indicates the type of institution that purchased it. There is also a value assigned for loans that are not sold. This means we compare loans made to borrowers with similar income in the same ZIP code and of the same type (e.g. Fannie, Freddie, Ginnie, portfolio). These highly granular fixed effects help address the concern that differences in credit outcomes are due to demand correlations rather than the underwriting system.

We focus on a loan-to-income ratio cutoff of 2.5, as this approximately corresponds to the manual underwriting debt-to-income cutoff of 36 per cent for a household with a relatively high level of existing debt payments.[13]

Column 1 of Table 4 shows that the high loan-to-income share increases by around 2

[12]Other datasets, such as McDash, do include the debt-to-income ratio. However, coverage is poor in the early 1990s, especially after conditioning on non-missing debt-to-income. McDash also does not allow us to identify the lender, which is important for our identification strategy.

[13]To inform the cutoff, we looked at data on components of the debt-to-income ratio from the mid 1990s. Property taxes in the 1995 Survey of Consumer Finances average around 1.3 per cent of the property value, and homeowner's insurance in the 1995 American Housing Survey averages around 0.45 per cent of the property value. The average 30 year mortgage rate in 1995 was around 8 per cent. At a loan-to-value ratio

percentage points relative to an average high loan-to-income share of 22 percent. Column 2 shows percentage point effects on the loan-to-income ratio using the same specification. Loan size increases by about 5 percent of income, relative to an initial average loan size of double the borrower's income. Figures A.3A and A.3B plot the responses over time and show that the magnitude of the effects increases. Peak effects in 1996 are 2.6 and 6.6 percentage points respectively.[14]

In Columns 3 and 4 of Table 4 we test whether the credit response differs by income level. We find that the new rules do not expand credit differentially to low income borrowers. If housing demand and down payments scale with income, we might expect both the numerator and the denominator of the debt-to-income ratio to increase at a similar rate with income. In recent HMDA data, high debt-to-income borrowing spans the entire income distribution (Appendix Figure A.4).[15] This could also partly reflect an expansion in borrowing capacity conditional on other risk factors, such as credit score. Overall, we do not find evidence of larger effects for low income households and conclude that the shift to statistical lending standards led to a broad-based increase in credit access.

Although we find lending outcomes move similarly prior to adoption, data limitations

---

of 80 per cent, the debt-to-income ratio is then approximately:

$$\begin{aligned} DTI &= \frac{OtherObligations}{Income} + \left(\frac{1.3 + 0.45}{0.8} + 1200 \times \frac{\frac{0.08}{12} \cdot (1 + \frac{0.08}{12})^{360}}{(1 + \frac{0.08}{12})^{360} - 1} \right) \cdot LTI \\ &= \frac{OtherObligations}{Income} + 11 \cdot LTI \end{aligned}$$

A household with an LTI of 2.5 will therefore have a DTI of around 36 per cent if their other financial obligations, such as debt payments and child support, are around 8.5 per cent of income. This is about the 75th percentile of other financial obligations for households in the 1995 Survey of Consumer Finances who bought their home in the last 5 years.

[14] Appendix Figure A.5 shows that application denial rates also decline. As denied applications may be reported less consistently than originations, especially during our early 1990s sample period, we prefer to focus on loan-to-income outcomes for originated loans.

[15] Unfortunately, debt-to-income was only added recently and is not reported for our sample period.

mean we cannot extend the analysis far back in time. We therefore conduct additional tests to better establish causality.

To address concerns regarding lenders' selection of a system, we use an instrumental variable approach. Within the set of early adopters, we predict adoption of LP (rather than DU) using lenders' pre-existing selling relationships. Whether a lender sold the majority of its 1992 loans to Freddie is a strong predictor of ultimately choosing LP rather than DU in 1994. Column 5 of Table 4 shows the first stage and Columns 6 and 7 show the second stage for each outcome. The estimates are very similar to OLS.

Although the automated systems could in principle be used for most loans, we expect effects coming through the change in lending algorithm to be small for FHA and VA loans and loans sold to Fannie Mae. To our knowledge, even if lenders' behavior were informed by the systems more broadly, these loan types would have been subject to additional underwriting constraints that would limit the possible relaxation. Appendix Table A.3 therefore splits the sample into loans that could be affected by the change in lending algorithm and loans that are unlikely to be affected.

Columns 1 and 2 of Appendix Table A.3 exclude FHA and VA loans and loans sold to Fannie. Columns 3 and 4 restrict the sample to these loan types for which the effect is expected to be small. We find that the response is much smaller and indeed close to zero for these loan types. In particular, there is no significant change in the share of loans with a loan-to-income ratio that would likely put them above the traditional 36 percent debt-to-income cutoff. This provides further evidence that the credit response reflects a relaxation of debt-to-income limits through through adoption of the automated system, not through demand effects or through a correlation with the lenders' business strategy.

4.2 Processing time response

Next we quantify effects of automation on processing time. We adopt a different empirical approach given that both DU and LP adopters should experience similar processing benefits. We therefore need to compare initial users of either system with other lenders. Comparing lenders who chose to participate in Fannie or Freddie's pilot programs with other lenders raises the possibility of selection. To mitigate this, we construct a control group of three matched lenders for each initial DU and LP user. There is a large pool of potential control lenders to choose from. The matching procedure targets three variables with the goal of finding lenders with similar business models in 1993: the share of refinance loans, the share of originations held in portfolio and the share of loans that were purchased (rather than originated). We also only match to lenders in the same size class, based on the number of loans. For purchase application i in income quartile n submitted to lender l with action a taken in year t we estimate:

$$Time_i = \alpha_{l,n,a} + \gamma_{g(l),a,t} + \delta_{z,t} + \beta AUS_l \cdot Post_t + \alpha_1 X_i + \epsilon_i \tag{4}$$

where $Time_i$ is the number of days between application and closing for originated loans, and the number of days between application and denial for denied applications. Actions a include origination and denial. Lender group $g(l)$ includes lender l as well as the three matched control lenders. AUS_l is an indicator equal to one for initial DU or LP adopters and zero for control lenders. $Post_t$ is an indicator equal to one for 1994-1996. We condition on log loan amount and log income and restrict the sample to conventional loans. In specifications with only originated loans we also further interact fixed effects with loan purchaser type.

Table 5 shows estimates of β from Equation 4 for early LP adopters. The time from application to closing declines by 3.6 days relative to matched lenders. In Column 2 we estimate Equation 4 separately for loans going to Freddie Mac and find a somewhat larger reduction of 4.6 days.[16] In Column 3, we restrict the sample to loans not going to Freddie Mac. We also find a statistically significant reduction in the processing time in this sample, consistent with broad usage of the systems for other loans as described in contemporaneous trade publications. Column 4 shows a substantial reduction in time from application to denial of 11.3 days. Table 6 repeats the analysis for early DU adopters. The effects are somewhat smaller but still broadly similar.

The overall effect of AUS is somewhat smaller than the 9 day reduction in processing time for purchase applications to 'Fintech' lenders documented by Fuster et al. (2019), and considerably smaller than potential reductions from automated underwriting that were projected in the mid 1990s (Maselli, 1994). Fuster et al. (2019) focus on lenders with a fully online application process, which is a different technology. For example, it is likely that most 'non-Fintech' lenders in their sample use an automated underwriting system given the more recent setting.

Our historical setting is useful for measuring the effect of underwriting automation on processing time, despite the fact that more recent data provide loan-level information about automated underwriting system usage. Most lenders now use automated underwriting to some extent and its usage for a given application is likely subject to selection bias. To illustrate this, we re-estimate the processing time relationship with 2018-2019 HMDA data and replace AUS_l with an indicator for whether an automated system was

[16] AmerUs, an early LP user, reported that the system allowed it to reduce the time between application and closing by 6 days (Hewitt, 1996).

actually used to underwrite application i (AUS_i).

Appendix Table A.4 shows that using an automated system is associated with a four day *increase* in processing time on average. Conditional on denial, processing time *increases* by 11 days. Conditional on origination, automation is associated with a 3 day reduction in processing time – similar to what we find in our historical setting.

This difference could be because denials following automated underwriting are more likely to occur for reasons that emerge later in the process. To explore this further, we separately estimate the relationship between automated approval and automated non-approval in Columns 2, 4, and 6. If the system does not recommend approval but the loan is ultimately made, this is associated with an additional 6 days to closing, perhaps because the loan then needed to go through an additional underwriting process. The use of an automated system is always associated with longer time to denial for denied loans, but this is more extreme for loans the system initially approved. This could be consistent with the systems being more likely to be used for loans that are not clear cut denials.

Overall, using our historic setting, we find that AUS allows for much faster denials and modestly reduces processing time for originated loans as well. However, the average reduction is less than one week, suggesting that the automated task likely accounts for a modest proportion of work and time needed to successfully close a loan.

4.3 House price response

Next we estimate the effect of the new rules applied by the automated systems on house prices. To establish a causal relationship, we take a difference-in-differences approach using the fact that initially only LP applied new rules, with DU automating the application of existing rules. We compare ZIP codes with different exposure to initial LP adopters within the same census division:

$$\log(Price_{z,t}) = \delta_z + \gamma_{div,t} + \sum_{\substack{k=1990 \\ k \neq 1993}}^{1996} \mathbb{1}_{t=k} \left(\beta_k \frac{LP_z}{SD(LP_z)} + \alpha_k X_z \right) + \epsilon_{z,t} \tag{5}$$

where $\log(Price_{z,t})$ is the log of the FHFA ZIP code house price index and LP_z is the measure of ZIP code exposure to early LP adopters defined in Section 3. Controls X_z include the ZIP code characteristics in Table 3 and AUS_z defined in Section 3. All controls are interacted with year dummies.[17] We divide the exposure measure by its standard deviation, so the coefficient of interest β_k is interpreted as the cumulative house price response to a one standard deviation increase in exposure.

Figure 2 plots the cumulative house price response estimated using Equation 5. A one standard deviation increase in 1993 exposure to Loan Prospector (4.4pp) raises prices by around 1.8 percent by 1996. Prices do not move differently prior to the rollout of Loan Prospector. We also plot the relationship between 1993-1996 house price growth and the ZIP code exposure measure conditional on the same characteristics. The relationship between the exposure measure and house price growth looks broadly linear (Figure 3).[18]

There are two main threats to a causal interpretation. First, early adopters' choice of Loan Prospector over Desktop Underwriter could be endogenous. For example, these lenders could be more expansionary, leading to stronger price growth where they are located. Second, even with exogenous adoption the exposure measure could still be cor-

[17] Interacting the controls with year dummies ensures that they enter the equation in the same way as the exposure measure to properly address omitted variable bias.

[18] In Appendix Figure A.7 we show the cumulative house price response over a longer horizon for transparency. We find the response grows for several years after the original shock. We believe price momentum could be a plausible explanation, for example, through feedback channels such as adaptive expectations (Armona et al. (2019); Bailey et al. (2018); Case et al. (2012)). We explore this channel further in Appendix A, but acknowledge the long-run evidence is primarily suggestive given the length of time since adoption.

related with other ZIP-level factors driving differential house price growth (though this would need to also align with the adoption timing given the lack of pre-trend). We address these concerns using several approaches.

To address the lender selection concern, we use lenders' 1992 GSE selling relationships to construct an instrument for the market share of early Loan Prospector adopters. We compute an analogous ZIP code exposure measure as if all early adopters who sold more than half of their 1992 GSE loans to Freddie chose Loan Prospector (and other early adopters chose Desktop Underwriter). Column 1 of Table 7 shows the first stage. Columns 2 and 3 show that the OLS and IV estimates are very similar to each other.

While our main specification does linearly control for the ZIP-level characteristics in Table 3, to further address concerns about correlation, we define a binary treatment indicator and weight ZIP codes to obtain covariate balance across treated and control groups (Hainmueller, 2012). We define the treated group as ZIP codes in the top exposure quartile and the control group as the remaining ZIP codes. Average values of covariates across treated and control groups before and after weighting are shown in Table A.7. Column 4 of Table 7 shows that price growth is 1.99 percentage points higher in the top quartile of exposure. The average difference in exposure between the treatment and control groups is 1.25 standard deviations. This suggests a 1.6 percentage point increase in price growth for a one standard deviation increase, consistent with Columns 2 and 3.

Columns 5 and 6 of Table 7 contain two placebo tests. In Column 5, we construct variables analogous to LP_z and AUS_z using the set of matched control lenders, described in Section 3, and estimate the same specification. We find no statistically significant effect on house prices. This helps to address the concern that early LP adopters have a different business model than early DU adopters and that this interacts with some other changes occurring around 1994 to generate a divergent house price trajectory.

There may be a concern that 1992 GSE selling relationships, which we use as an instrument above, are linked to house prices through some channel other than the automated underwriting system itself. For example, through a correlation with lender characteristics that matter for local house price growth. If this were the case, we would expect 1992 GSE relationships to also matter for house prices when an identical approach is applied to the set of matched control lenders. Column 6 shows the relationship between this placebo version of the instrument and house prices. We again find no statistically significant effect on house prices, consistent with the instrument operating through the statistical lending standards introduced by LP, but not (initially) by DU.

Appendix Table A.5 reports two additional specifications. Column 1 compares early LP adopters with matched control lenders, rather than with early DU adopters, and yields a similar house price response. Column 2 compares early DU adopters with matched control lenders. If automation alone raised house prices, we would expect the estimated response in Column 1 to be larger than in the DU comparison and the estimate in Column 2 to be positive and statistically significant. Instead, the coefficient in Column 1 is similar to the DU comparison, and the estimate in Column 2 is statistically insignificant. Together, these results suggest that automation itself did not raise house prices; rather, the response we document appears to be driven by the shift in lending standards. However, we note that these specifications are potentially more exposed to selection into early AUS adoption than our main results.

4.3.1 Magnitude

To better compare our house price response with other studies, we estimate the short-run elasticity of ZIP code house prices to credit using ZIP code exposure to early LP adopters as an instrument. This requires us to estimate ZIP-level credit responses. Given changes to HMDA location reporting rules over our sample, we do not estimate the response of aggregate lending and instead focus on the average loan-to-income ratio and share of high loan-to-income lending.[19] Table 8 shows first and second stage output. A one percent increase in average loan-to-income leads to a 0.43 percent increase in house prices. A one percentage point increase in the share of loans with a loan-to-income ratio above 2.5 leads to a 1.82 percent increase in house prices. The magnitude of the response is somewhat larger than, but still broadly similar to, estimates of the response of house prices to mortgage credit obtained in other settings.[20]

4.3.2 Potential effect of systems adoption on national house prices

We document a relative local house price boom in locations with greater exposure to initial Loan Prospector users. Assuming adoption by other lenders stimulated a similar house price response as we see for early adopters, automated underwriting technology could have played some role in the early phases of the 2000s housing boom. We com-

[19] Prior to 1996, depository institutions were only required to report locations for properties in MSAs where the lender had a home or branch office. This did not include offices of affiliates such as brokers or correspondents, or non-branch locations which accepted applications. Non-depositories were considered to have a branch office in any MSA where they had at least 5 reportable loans or applications for home purchase or home improvement in the previous calendar year, and therefore reported locations for a larger share of loans. From 1996 onward, banks or thrifts with at least $250 million in assets (or that were subsidiaries of a holding company with banking or thrift assets of at least $1 billion) no longer received exemptions from reporting property locations and location reporting improved considerably.

[20] Favara and Imbs (2015) find an elasticity of house prices to loan-to-income ratio on impact of 0.12, with a peak elasticity of 0.2 two years out. Di Maggio and Kermani (2017) find an elasticity of house prices to loan amount of 0.33.

bine aggregate adoption data from Fannie and Freddie with the house price response estimated using our natural experiment to compute a back-of-the-envelope aggregate response from 1993 to 2005.

A lender adopting in year t is assumed to contribute the cumulative house price response through 2005 that Section 4.3 would assign at horizon $2005 - t$, as if the response were comparable to the early-adopter shock we estimate. We assume the cumulative response stops growing after three years (1996 for 1994 adopters; Figure 2). To reflect later rule alignment, we treat DU as not applying the new rules before 1997 and include DU usage from 1997 onward together with Loan Prospector. Incremental adoption is small after 2000, so for simplicity we hold usage flat from that point (Table A.1). Figure 4 shows the resulting adoption profile combining Table A.1 with these assumptions.

Our main estimation sample is limited to the set of ZIP codes for which an annual repeat sales index is available back to 1990. It is possible that the house price response could be larger in this sample than for the U.S. as a whole, as these tend to be more urban ZIP codes with higher population and home values.[21] This suggests we should adjust our estimates to make them more representative prior to aggregating. We adapt our main specification to allow the house price response to vary with housing supply elasticity (Saiz, 2010). We then divide all ZIP codes into supply elasticity deciles weighted

[21] To explore this, we estimate the house price response from 1990-2000 and compare it with estimates obtained using the census median home value instead of the FHFA price index. Columns 1 and 2 of Table A.6 show that broadly similar responses are obtained using Decennial Census home values on our main sample of ZIP codes. Column 3 shows that switching to the more broadly available dependent variable reduces the magnitude of the price response by about 25 percent. In Column 4 we drop the housing supply elasticity control to expand the sample further. This reduces the magnitude by about 45 percent relative to the main sample.

by number of housing units. The aggregate price effect for supply elasticity decile i is:

$$\sum_{k=1994}^{2005} s_k \times ME^i_{1994+(2005-k)} \tag{6}$$

where s_k is incremental adoption in year k from the aggregate usage series (with $s_k = 0$ for $k > 2000$ under our flat-usage assumption), and ME^i_t is the year t marginal effect of early Loan Prospector exposure evaluated for elasticity decile i. We set $ME^i_t = ME^i_{1996}$ for all $t \geq 1996$, reflecting a three-year cap on the cumulative response. We obtain a national response by aggregating over the distribution of housing supply elasticity, weighting by the number of housing units in each bin. We conservatively assign locations without an estimated supply elasticity to the top bin, which has the smallest price response. Figure A.6 plots the implied aggregate price response across the distribution of housing supply elasticity.

We calculate that gradual adoption of Fannie and Freddie's automated underwriting systems predicts 16 percent cumulative house price growth from 1993 to 2005. This is broadly consistent with Greenwald (2018), who models the effect of a relaxation of debt-to-income limits from 36 percent to 58 percent and finds a 17 percent increase in the price-to-rent ratio to the peak of the boom in 2006Q1. Greenwald (2018)'s exercise is motivated by an observed relaxation of debt-to-income limits in mortgage data. In this paper, we argue gradual adoption of Fannie and Freddie's systems provides a plausible explanation for the debt-to-income relaxation Greenwald (2018) observes.

Overall, this extrapolation from our difference-in-differences estimates suggests that the rollout of automated underwriting could have had an important effect on house prices. However, care should be taken when interpreting the magnitude. Our calculation relies on the strong assumptions that our difference-in-differences estimates can be applied to later adopters and can be aggregated to the national level.[22]

4.4 House price correlation

If two lenders adopt an automated underwriting system and use it to underwrite all applications, the lending standards applied by these two lenders become the same as each other (and other lenders who use the same system for all loans). A change in underwriting rules then immediately propagates to all lenders using the system. Automated underwriting may therefore affect house price comovement in a similar way to banking integration (Landier et al., 2017).

Even without automated underwriting, common product rules should themselves contribute to comovement. For example, if all loans were made in accordance with Fannie, Freddie or FHA rules, the result could be similar to that of a highly concentrated banking system, even if individual lenders remain small. Automated underwriting could strengthen integration in addition to this loan product effect for two main reasons.

First, if system rules are favorable or efficiency gains are large, lenders may be incentivized to apply the system beyond those loans to which they would have applied Fannie or Freddie's manual rules. Narrative evidence indicates that early adopters used the systems broadly, including on portfolio loans (Section 2). Post-2018 HMDA, which reports

[22] A primary concern with aggregation in our setting relates to the supply of funds. If lending standards are relaxed but the supply of funds is fixed, it is expected that aggregate interest rates will rise. If true, we might still expect to see divergence in local house prices due to differences in local lending standards, but with a smaller effect on national prices. Justiniano et al. (2019) consider this argument and note that the empirical facts suggest an expansion in the supply of funds over the housing boom period. Such an expansion of funds, for example due to securitization or a global savings glut, would allow changes in lending standards to have large aggregate effects on house prices. We also note that agency MBS has broad appeal both in the U.S. and internationally. A large expansion in GSE lending criteria need not necessarily imply a large increase in mortgage rates.

system usage, shows Fannie and Freddie's systems are applied to loans that are not sold to the GSEs. Second, the manual underwriting process involved more discretion on the part of lenders, even for fairly standardized products. By removing this idiosyncratic component, automation would tend to lead to stronger correlation. Together, reduced discretion and wider application of a common rule set could raise correlation in credit conditions, and, in turn, in house prices, across places linked through shared system adoption.

We evaluate the effect of greater lender integration coming from either LP or DU. It is still possible that DU could increase comovement through reduced discretion and application to a broader set of loans, though incentives to apply the system more broadly would likely be limited to processing efficiencies initially.

Given the large number of pairwise ZIP code combinations, we focus on county correlations. Our dependent variable $C_{i,j,t}$ is correlation of annual house price growth between counties i and j for over the five years t to $t+4$. We use non-overlapping windows covering 1984-2003, so each county pair contributes four observations, two in the pre-adoption period (1984-1988; 1989-1993) and two in the post-adoption period (1994-1998; 1999-2003).

Next we quantify the increase in integration between counties i and j implied by a set of lenders adopting the same automated underwriting system. We measure integration using 1993 lender shares so exposure is predetermined. We interpret estimates as showing that county pairs more connected through early AUS adopters saw stronger house price correlation after they started using the systems. Although both DU and LP gradually spread to other lenders over time, the early structure could still remain relevant. Following Landier et al. (2017) we first compute a co-Herfindal index treating initial LP users as distinct lenders. Let S_i^k denote lender k's 1993 market share in county i:

$$H^1_{i,j} = \sum_{k \in LP} S_i^k S_j^k + \sum_{l \notin LP} S_i^l S_j^l \tag{7}$$

We then recompute the index treating initial LP users as if they were the same lender:

$$H^2_{i,j} = (\sum_{k \in LP} S_i^k)(\sum_{k \in LP} S_j^k) + \sum_{l \notin LP} S_i^l S_j^l \tag{8}$$

The increase in integration attributable to treating LP adopters as one lender is the difference between the two indexes:

$$\Delta H_{i,j} = H^2_{i,j} - H^1_{i,j} = (\sum_{k \in LP} S_i^k)(\sum_{k \in LP} S_j^k) - \sum_{k \in LP} S_i^k S_j^k \tag{9}$$

We repeat the procedure for initial DU adopters. Note that $\Delta H_{i,j}$ reflects common system usage by different lenders. It does not reflect integration arising from common lenders, as this component is differenced out.

To separate integration due to system usage rather than the prevalence of Fannie, Freddie or FHA loan products, we also construct a loan product integration index:

$$H^p_{i,j} = S_i^{Fannie} \cdot S_j^{Fannie} + S_i^{Freddie} \cdot S_j^{Freddie} + S_i^{FHA} \cdot S_j^{FHA} + \sum_l \tilde{S}_i^l \tilde{S}_j^l \tag{10}$$

Where S_i^{Fannie} and $S_i^{Freddie}$ are the share of loans in county i sold to Fannie Mae or Freddie Mac, S_i^{FHA} is the FHA share and $\tilde{S}_i^l$ is lender l's share of the market in county i computed using only loans not in those three categories in the numerator. This captures integration arising from common rules applied to Fannie, Freddie and FHA loans, and potentially common rules for other loans due to the presence of the same lender in multiple counties.

We compute this measure for each time period.

For county i (in Combined Statistical Area $a(i)$) and county j (in Combined Statistical Area $a(j)$) and year t, we estimate:

$$
\begin{aligned}
Corr_{i,j,t} = \alpha_{i,j} + \gamma_{a(i),a(j),t} + \beta_1 Post_t \cdot \Delta H^{LP}_{i,j} + \beta_2 Post_t \cdot \Delta H^{DU}_{i,j} \\
+ Controls_{i,j,t} + \epsilon_{i,j,t}
\end{aligned}
$$

Where $Corr_{i,j,t}$ is the annual house price growth correlation between county i and county j for years t to $t+4$. The post indicator is equal to 0 for periods starting in $t = 1984$ to $t = 1989$ ($t = 1989$ includes house price growth up until 1993) and 1 for later years. Controls include the per capita income growth and population growth correlation between county i and county j in the time period starting with year t and the loan product integration measure discussed above. Standard errors are clustered by Combined Statistical Area of county i and county j. All specifications include fixed effects for time period by CSA of county i by the CSA of county j. So for example, when looking at determinants of the correlation between a county i in Houston-Pasadena CSA and a county j in Dallas-Fort Worth CSA, we use variation across Houston-Dallas county pairs. We standardize the three control variables, so each coefficient is interpreted as a one standard deviation increase. We scale $\Delta H^{LP}_{i,j}$ and $\Delta H^{DU}_{i,j}$ by 10,000. So if two locations i and j respectively contain two different lenders A and B each with a market share of 10%, who are both LP adopters, then $\Delta H^{LP}_{i,j} = 10,000 * 0.1 * 0.1 = 100$.

Table 9 reports the results. We present estimates with and without the later 1999-2003 period for transparency. DU implemented new rules similar to LP by the late 1990s, so we expect that DU-linkages could induce greater comovement in the long run than in the short run. We find that a one unit increase in 1993 LP integration raises house price correlation by 0.27 percentage points in the short run and 0.39 percentage points in the extended sample. As described above, in the case where each county contains a lender with 10% market share adopting LP, the implied short-run increase in house price correlation due to common system adoption by these two lenders is 27pp. A one standard deviation increase, 3.89, increases house price correlation by around 1 percentage point, comparable to an 0.6 standard deviation increase in population correlation. Adding the loan product integration control, $H^p_{i,j}$ has little effect on the coefficients on system integration. We find that a one standard deviation increase in loan product integration increases house price correlation by about 1.13 percentage points. This is equivalent to a 1.5 standard deviation increase in population correlation.

Table 9 shows that DU integration only raises house price correlation in the extended sample. The negligible short-run effect could be consistent with DU not applying substantial rule changes before the late 1990s. There may also be differences across systems in the relevance of 1993 integration measures for the long-run network. Unfortunately it is difficult for us to quantify this using the data available to us. Figure 5 plots separate coefficients for each time period. The systems integration measures are not associated with a change in correlation prior to systems adoption.

5 Conclusion

We use the 1990s rollout of Fannie and Freddie's automated underwriting systems to study the effects of automated mortgage underwriting on processing time, lending standards, house prices and price comovement. Automation ultimately implied the adoption of more complex, statistically-informed underwriting rules. Freddie Mac's system Loan

Prospector allowed households to take on larger mortgage payments relative to their income, and Fannie Mae later incorporated similar rules into Desktop Underwriter. We show that locations with early exposure to these rules experienced a local house price boom starting around 1995.

Overall, we find that automated underwriting systems do not simply streamline processing, but can facilitate coordinated changes in lending standards, potentially contributing to housing boom or bust episodes, and increasing house price comovement across locations where lenders use the same systems. This has important implications. The U.S. mortgage market is currently highly concentrated in terms of system usage, with most loan applications being run through either Fannie or Freddie's system.

References

Acharya, Viral V, Katharina Bergant, Matteo Crosignani, Tim Eisert, and Fergal McCann, "The Anatomy of the Transmission of Macroprudential Policies," *The Journal of Finance*, 2022, *77* (5), 2533–2575.

Adelino, Manuel, Antoinette Schoar, and Felipe Severino, "Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class," *Review of Financial Studies*, 2016, *29*, 1635–1670.

_ , _ , and _ , "Credit Supply and House Prices: Evidence from Mortgage Market Segmentation," *Journal of Financial Economics*, 2025, *163*, 103958.

Albanesi, Stefania, Giacomo DeGiorgi, and Jaromir Nosal, "Credit growth and the financial crisis: A new narrative," *Journal of Monetary Economics*, 2022, *132*, 118–139.

American Banker, "Freddie Rolls Out High-Tech Underwriting System," *American Banker*, February 1995, *160* (33).

Armona, Luis, Andreas Fuster, and Basit Zafar, "Home Price Expectations and Behavior: Evidence from a Randomized Information Experiment," *Review of Economic Studies*, July 2019, *86* (4), 1371–1410.

Bailey, Michael, Rachel Cao, Theresa Kuchler, and Johannes Stroebel, "The Economic Effects of Social Networks: Evidence from the Housing Market," *Journal of Political Economy*, 2018, *126* (6), 2224–2276.

Baum-Snow, Nathaniel and Lu Han, "The Microgeography of Housing Supply," *Journal of Political Economy*, 2024, *132* (6), 1897–1946.

Berg, Tobias, Andreas Fuster, and Manju Puri, "Fintech lending," *Annual Review of Financial Economics*, 2022, *14* (1), 187–207.

_ , Valentin Burg, Ana Gombović, and Manju Puri, "On the Rise of FinTechs: Credit Scoring Using Digital Footprints," *The Review of Financial Studies*, 09 2019, *33* (7), 2845–2897.

Bhutta, Neil, *Regression Discontinuity Estimates of the Effects of the GSE Act of 1992*, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, 2009.

Blattner, Laura and Scott Nelson, "Data and Disparities in Consumer Credit," *Working Paper*, 2021.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru, "Why is Intermediating Houses so Difficult? Evidence from iBuyers," *Working Paper*, 2020.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, "Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks," *Journal of Financial Economics*, 2018, *130* (3), 453–483.

Case, Karl, Robert Shiller, and Anne Thompson, "What Have they been Thinking? Home Buyer Behavior in Hot and Cold Markets," *NBER Working Paper No. 18400*, 2012.

Chodorow-Reich, Gabriel, Adam M Guren, and Timothy J McQuade, "The 2000s Housing Cycle with 2020 Hindsight: A Neo-Kindlebergerian View," *Review of Economic Studies*, 2024, *91* (2), 785–816.

Costanzo, Chris, "Automated Underwriting is now Business as Usual," *Community Banker*, Apr 2004, *13* (4), 49.

Davis, Morris A, William D Larson, Stephen D Oliner, and Benjamin R Smith, "A Quarter Century of Mortgage Risk," *Review of Finance*, 2023, *27* (2), 581–618.

DeMuth, Jerry, "The Selling of Two Systems," *Mortgage Banking*, Apr 1999, *59* (7), 16.

Demyanyk, Yuliya and Otto Van Hemert, "Understanding the subprime mortgage crisis," *The review of financial studies*, 2011, *24* (6), 1848–1880.

Dennis, Marshall W. and Michael J. Robertson, *Residential Mortgage Lending*, 4 ed., Prentice Hall, 1995.

Di Maggio, Marco and Amir Kermani, "Credit-Induced Boom and Bust," *Review of Financial Studies*, 2017, *30* (11).

— , Dimuthu Ratnadiwakara, and Don Carmichael, "Invisible Primes: Fintech Lending with Alternative Data," *NBER Working Paper No. w29840*, 2022.

Favara, Giovanni and Jean Imbs, "Credit Supply and the Price of Housing," *American Economic Review*, 2015, *105* (3), 958–992.

Favilukis, Jack, Sydney Ludvigson, and Stijn Van Nieuwerburgh, "The Macroeconomic Effects of Housing Wealth, Housing Finance, and Limited Risk Sharing in General Equilibrium," *Journal of Political Economy*, 2017, *125* (1).

Foote, C., L. Loewenstein, and P. Willen, " Technological Innovation in Mortgage Underwriting and the Growth in Credit: 1985–2015," *Boston Fed Research Department Working Paper*, 2019, (19-11).

— , — , and — , "Cross-Sectional Patterns of Mortgage Debt during the Housing Boom: Evidence and Implications," *The Review of Economic Studies*, 2020.

Foster, Doug, "The Debate over Automated Underwriting," *Mortgage Banking*, May 1997, *57* (8), 10–15.

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery, "The Role of Technology in Mortgage Lending," *The Review of Financial Studies*, 2019, *32* (5), 1854–1899.

— , Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther, "Predictably Unequal? The Effects of Machine Learning on Credit Markets," *The Journal of Finance*, 2022, *77* (1), 5–47.

Gao, Janet, Hanyi Livia Yi, and David Zhang, "Algorithmic Underwriting in High Risk Mortgage Markets," *Available at SSRN 4602411*, 2023.

Gates, Susan, Vanessa Perry, and Peter Zorn, "Automated Underwriting in Mortgage Lending: Good News for the Underserved?," *Housing Policy Debate*, 2002, *13* (2), 369–

391.

Greenwald, Daniel, "The Mortgage Credit Channel of Macroeconomic Transmission," *Working Paper*, 2018.

Greenwald, Daniel L and Adam Guren, "Do Credit Conditions Move House Prices?," *American Economic Review*, 2025, *115* (10), 3559–3596.

Griffin, John, Samuel Kruger, and Gonzalo Maturana, "What drove the 2003–2006 house price boom and subsequent collapse? Disentangling competing explanations," *Journal of Financial Economics*, September 2021, *141* (3), 1007–1035.

Hainmueller, Jens, "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies," *Political Analysis*, 2012, *20* (1), 25–46.

Harney, Kenneth, "Burdened by a Heavy Debt Load? A Mortgage is Not Out of Reach.," *The Washington Post*, 1996.

Hewitt, Janet Reilly, "Giving Away the Gain," *Mortgage Banking*, May 1996, *56* (8), 26.

Howell, Sabrina T., Theresa Kuchler, David Snitkof, Johannes Stroebel, and Jun Wong, "Lender Automation and Racial Disparities in Credit Access," *The Journal of Finance*, 2024, *79* (2), 1457–1512.

Irwin, Robert, *Tips and Traps When Mortgage Hunting*, McGraw-Hill, 1992.

Jansen, Mark, Hieu Quang Nguyen, and Amin Shams, "Rise of the Machines: The Impact of Automated Underwriting," *Management Science*, 2025, *71* (2), 955–975.

Jiang, Erica Xuewei and Anthony Lee Zhang, "Collateral Value Uncertainty and Mortgage Credit Provision," *Journal of Financial Economics*, 2025, *169*, 104054.

Johnson, Stephanie, "Mortgage Leverage and House Prices," *Working Paper*, 2020.

Jones, James, "Automated Underwriting Makes it Possible to Increase Origination Volume," *American Banker*, September 1997, p. 8.

Justiniano, A, G Primiceri, and A Tambalotti, "Credit Supply and the Housing Boom," *Journal of Political Economy*, 2019, *127* (3), 1317–1350.

Kaplan, Greg, Kurt Mitman, and Giovanni L Violante, "The Housing Boom and Bust: Model Meets Evidence," *Journal of Political Economy*, 2020, *128* (9).

Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans*," *The Quarterly Journal of Economics*, 02 2010, *125* (1), 307–362.

Khoei, Sarah, "When Models Fail: Evidence from Automated Underwriting in Auto Loan Markets," 2025.

LaMalfa, Tom, "Wholesale Giants 1995," *Mortgage Banking*, 1996, *57*, 42–59.

_ , "Wholesale Giants 1996," *Mortgage Banking*, 1997, *57*, 42–59.

_ , "Wholesale Giants 1997," *Mortgage Banking*, 1998, *58*.

— , "Wholesale Giants 1998," *Mortgage Banking*, 1999, *59.*

Landier, Augustin, David Sraer, and David Thesmar, "Banking Integration and House Price Co-movement," *Journal of Financial Economics*, 2017, *125* (1), 1–25.

Lee, Jung Youn, Joonhyuk Yang, and Eric T. Anderson, "Using Grocery Data for Credit Decisions," *Management Science*, 2025, *71* (4), 2753–2777.

Lewellen, Stefan and Emily Williams, "Did Technology Contribute to the Housing Boom? Evidence from MERS," *Journal of Financial Economics*, 2021, *141* (3), 1244–1261.

Loutskina, Elena and Philip E. Strahan, "Financial Integration, Housing and Economic Volatility," *Journal of Financial Economics*, 2015, *115*, 25–41.

Markus, M, Andrew Dutta, Charles Steinfield, and Rolf Wigand, "The Computerization Movement In The US Home Mortgage Industry: Automated Underwriting From 1980 To 2004," in "Computerization Movements and Technology Diffusion: From Mainframes to Ubiquitous Computing," Information Today, 01 2008, pp. 115–144.

Maselli, Peter, "Mortgages in Minutes," *Mortgage Banking*, 1994, *55* (1), 102.

Mian, Atif and Amir Sufi, "The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis," *The Quarterly Journal of Economics*, 2009, *124* (4), 1449–1496.

Mikel, Pat and Terri L. Baker, "Here and Now High-Tech," *Mortgage Banking*, June 1992, *52* (9), 26.

Nixon, Brian, "The New World of Automated Lending," *Savings and Community Banker*, March 1995, *4* (3), 16.

Oliver, Geoffrey and Laura McDonald, "Managing Production for Profits," *Mortgage Banking*, October 1997, *58* (1), 152–157.

Petersen, Mitchell A. and Raghuram G. Rajan, "Does Distance Still Matter? The Information Revolution in Small Business Lending," *The Journal of Finance*, 2002, *57* (6), 2533–2570.

Pierzchalski, Larry, "Guarding against risk," *Mortgage Banking*, June 1996, *56* (9).

PR Newswire, "Fannie Mae announces first nationally-available technology, saving time and costs, to process any conventional mortgage; also names 19 lenders on Desktop Originator and Underwriter system," April 1995.

Saiz, Albert, "The Geographic Determinants of Housing Supply," *Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

Straka, John W., "A Shift in the Mortgage Landscape: The 1990s Move to Automated Credit Evaluations," *Journal of Housing Research*, 2000, *11* (2), 207–232.

Strickberger, Matt, "Freddie Challenges Mortech's AU Market Share Data," *National Mortgage News*, March 1999, *23* (28), 15.

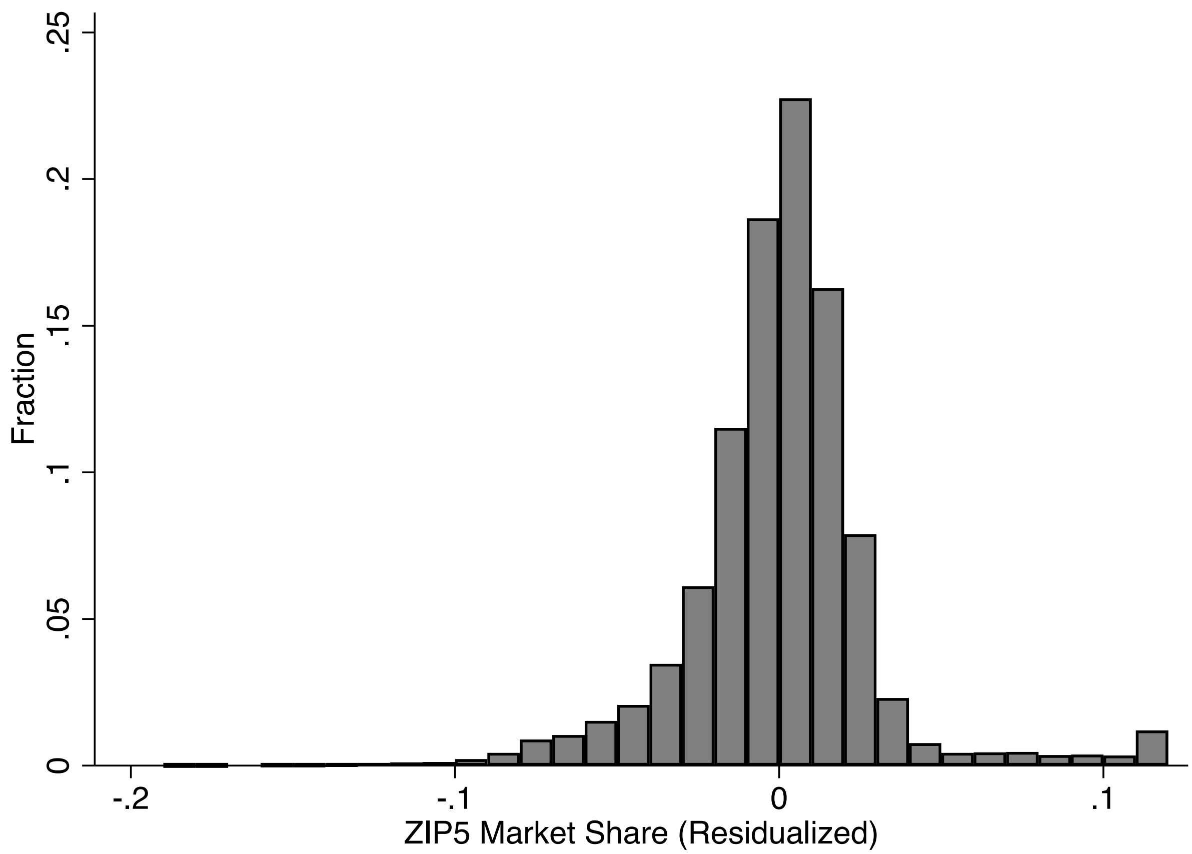
Sullivan, Orla O, "GSEs Detail Prices on AU," *National Mortgage News*, November 1995, *20* (7), 1.

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

Temkin, Kenneth, Jennifer Johnson, and Diane Levy, "Subprime Markets, the Role of the GSEs and Risk-based Pricing," Technical Report, U.S. Department of Housing and Urban Development March 2002.

FIGURE 1

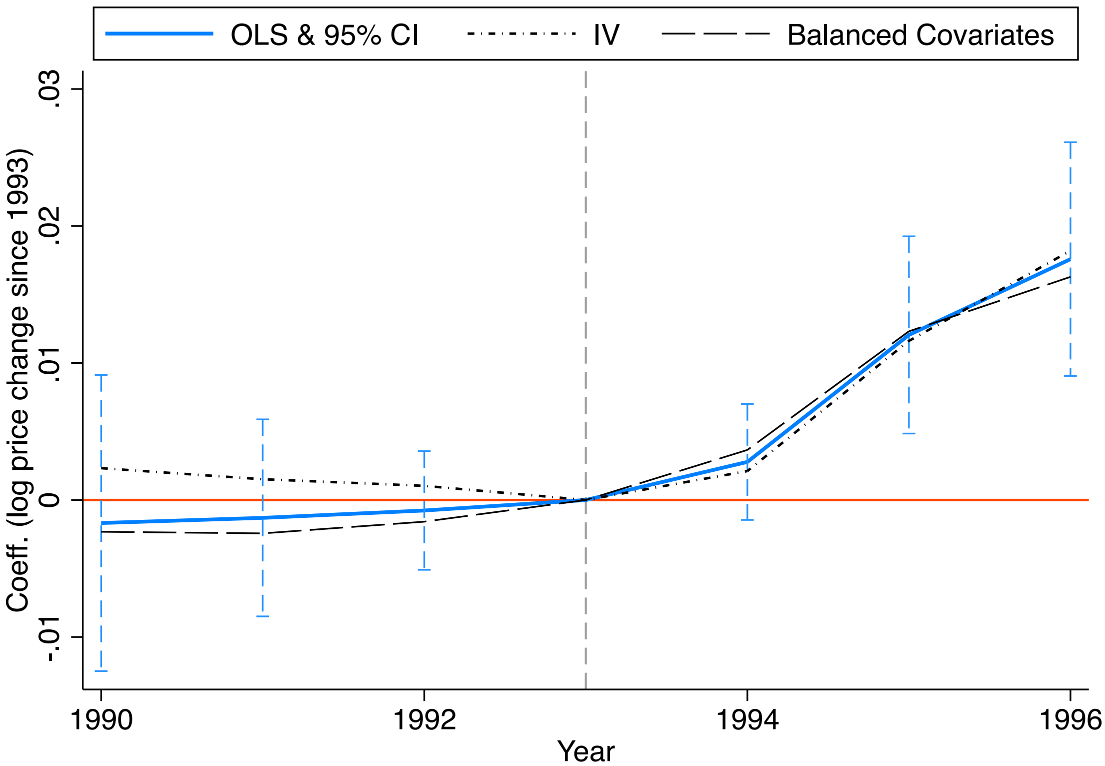
ZIP CODE EXPOSURE TO INITIAL LOAN PROSPECTOR USERS



NOTES: Figure 1 plots the residualized exposure measure. The raw exposure measure is: $LP_z = \frac{\text{\# Loans reported in ZIP } z \text{ by LP lenders in Table 1}}{\text{\# Loans reported in ZIP } z \text{ by all HMDA reporters}}$, computed using 1993 HMDA loans. We condition on the following variables, at the zip code level unless otherwise indicated: (county) coastal indicator, number of lenders, large lender market share, thrift market share, commercial bank market share, share of originations sold to either Fannie or Freddie, share of originations sold to Freddie, log median household income, Saiz (2010) housing supply elasticity, combined share of initial LP and DU users. The sample includes ZIP codes in metropolitan counties with non-missing house price data. The measure is winsorized at the 99th percentile. Sources: HMDA and authors' calculations.

FIGURE 2

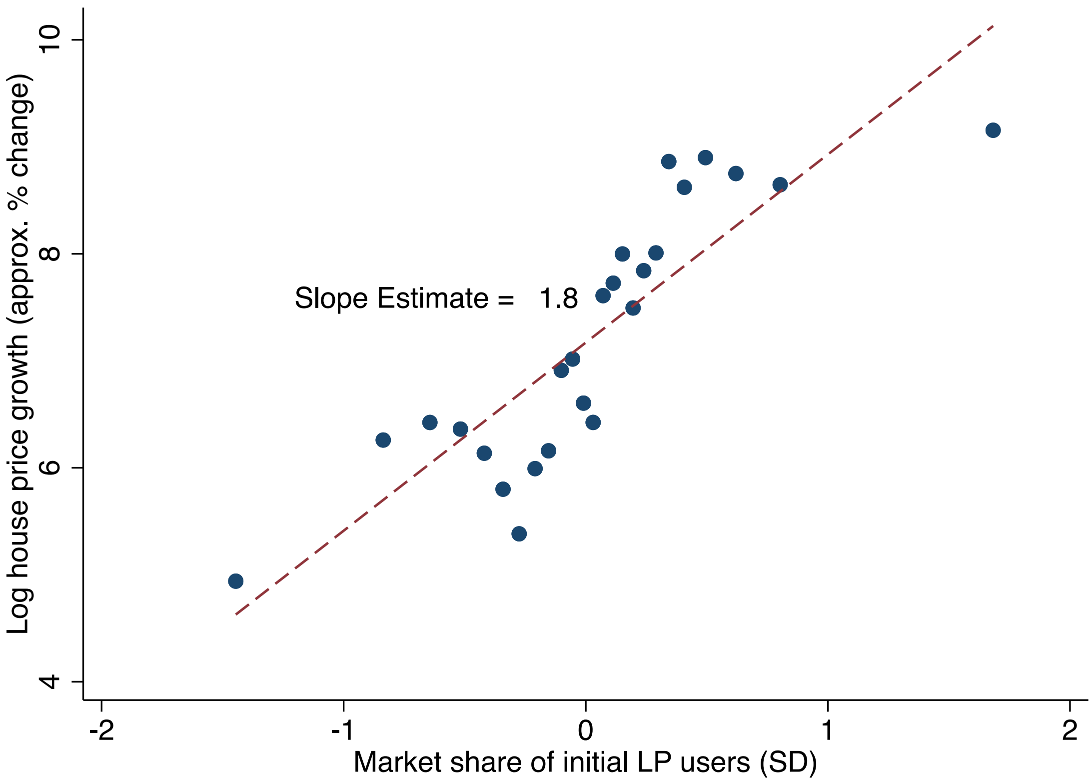
SWITCHING TO STATISTICAL UNDERWRITING STANDARDS RAISES HOUSE PRICES



NOTES: Figure 2 shows the cumulative house price response (estimates of $\{\beta_k\}$ from Equation 5). We include census division by year fixed effects and condition on the following variables interacted with year dummies, at the zip code level unless otherwise indicated: (county) coastal indicator, number of lenders, large lender market share, thrift market share, commercial bank market share, share of originations sold to either Fannie or Freddie, share of originations sold to Freddie, log median household income, Saiz (2010) housing supply elasticity, the combined ZIP code share of initial DU and LP users AUS_z. The figure also shows the estimates obtained when using LP_z^{IV} as an instrument for LP_z, where LP_z^{IV} is constructed assuming that early adopters with a primary relationship with Freddie Mac in 1992 choose LP, whereas other lenders choose DU. Additionally, we show estimates obtained using weights to achieve covariate balance. We use top quartile exposure to instrument for continuous exposure LP_z while applying weights to achieve balance across top quartile locations and other locations. This is so that the magnitude remains comparable to the other approaches. The sample includes ZIP codes in metropolitan counties with FHFA house price data available continuously from 1990 to 2005. Coefficients are interpreted as cumulative log price changes relative to 1993. Standard errors are clustered by CBSA. Sources: FHFA HPI; HMDA 1990 decennial census; BEA; NOAA list of coastal counties; and authors' calculations.

FIGURE 3

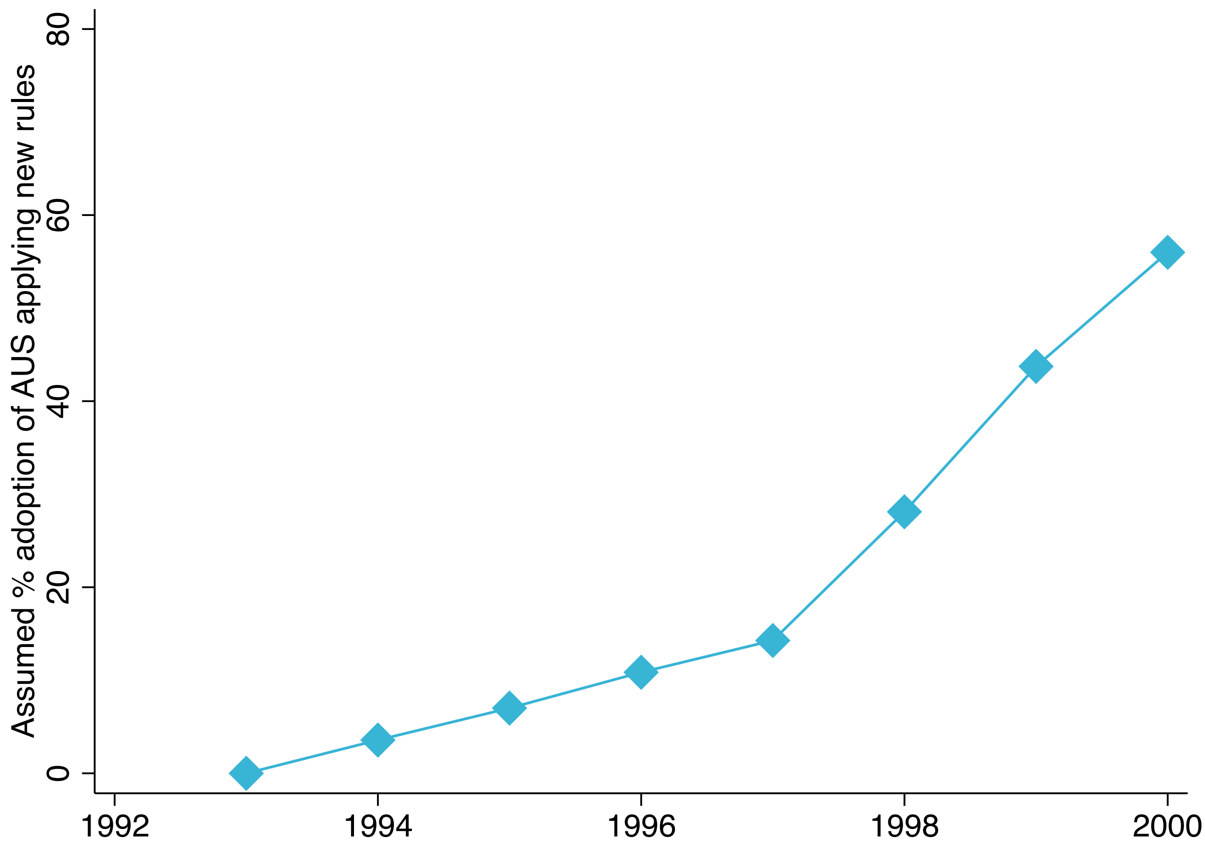
RELATIONSHIP BETWEEN LP EXPOSURE AND 1993-1996 HOUSE PRICE GROWTH



NOTES: This binned scatter plot shows the ZIP code log house price change from 1993-1996 by the market share of initial Loan Prospector users. We condition on the following variables, at the ZIP code level unless otherwise indicated: (county) coastal indicator, number of lenders, large lender market share, thrift market share, commercial bank market share, share of originations sold to either Fannie or Freddie, share of originations sold to Freddie, log median household income, Saiz (2010) housing supply elasticity, the combined ZIP code share of initial DU and LP users AUS_z. The sample includes ZIP codes in metropolitan counties with with FHFA house price data available continuously from 1990 to 2005. We use census division fixed effects. Sources: FHFA HPI; HMDA; 1990 decennial census; NOAA list of coastal counties.

FIGURE 4

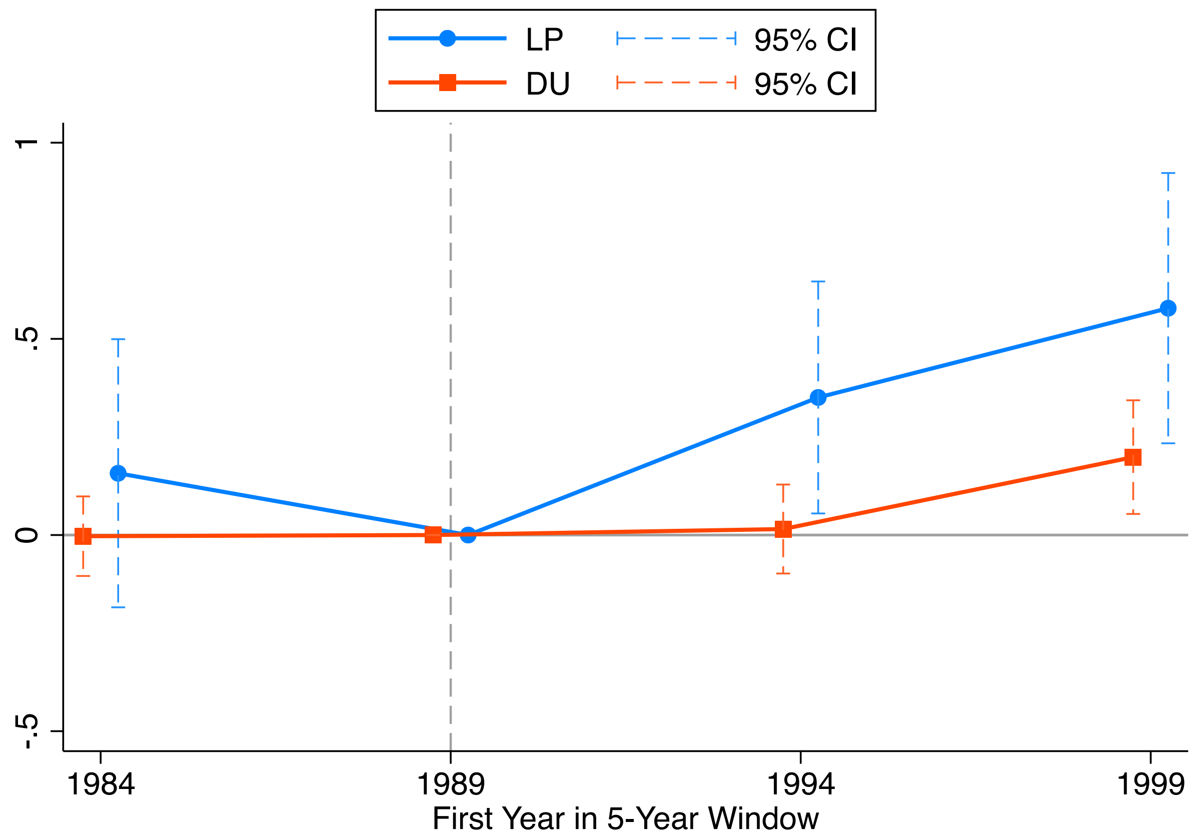
DIFFUSION OF NEW RULES THROUGH AUTOMATED UNDERWRITING ADOPTION



NOTES: Data on systems adoption is from the sources listed in Table A.1 from 1995-2000 and the market shares of early adopters for 1994. We weight data for Fannie and Freddie's systems using data on Fannie and Freddie's mortgage acquisitions from HUD's GSE public use database. We use linear interpolation to compute values for years where no data is available. We assume that DU incorporates new rules starting in 1997.

FIGURE 5

EFFECT OF AUTOMATED UNDERWRITING ADOPTION ON COUNTY HOUSE PRICE CORRELATION



NOTES: This figure shows the effect of systems-induced integration on county house price growth correlation. The base period is 1989-1993. For county i (in Combined Statistical Area $a(i)$) and county j (in Combined Statistical Area $a(j)$) and year t:

$$\begin{aligned} Corr_{i,j,t} = \alpha_{i,j} + \gamma_{a(i),a(j),t} + \beta_t^1 \cdot \Delta H_{i,j}^{LP} + \beta_t^2 \cdot \Delta H_{i,j}^{DU} \\ + Controls_{i,j,t} + \epsilon_{i,j,t} \end{aligned}$$

where $Corr_{i,j,t}$ is the annual growth rate correlation between county i and county j over a 5-year period starting in year t, where $t \in \{1984, 1989, 1994, 1999\}$. The base period is $t = 1989$, which includes house price growth from 1989 up until 1993. Controls include the annual population growth and per capita income growth correlation computed over the same 5-year period as the dependent variable and an additional measure of lending integration based on the shares of loans sold to Fannie Mae, Freddie Mac and the share of FHA loans. $\Delta H_{i,j}^{LP}$ and $\Delta H_{i,j}^{DU}$ are trimmed at the 1st and 99th percentile. Standard errors are clustered by Combined Statistical Area of county i Combined Statistical Area of county j. All specifications include fixed effects for time by CSA of county i by CSA of county j.

TABLE 1

INITIAL USERS OF FREDDIE MAC AND FANNIE MAE'S UNDERWRITING SYSTEMS

Loan Prospector (Freddie Mac)	Desktop Underwriter (Fannie Mae)
Citicorp Mortgage	American City Mortgage Corp.
First Security Savings Bank (Flagstar Bank)	BancBoston Mortgage Corp.
Midland Financial Mortgages (Amerus)	BrooksAmerica Mortgage Corp.
Mission Hills Mortgage Corp.	Crestar Mortgage Corp.
Monument Mortgage	Crossland Mortgage Corp.
Mortgage America	Fleet Mortgage Corp.
Old Kent Mortgage Co.	Headlands Mortgage Co.
PHH Mortgage Corp.	ICM Mortgage Corp. (Pulte Mortgage)
Standard Federal Bank (InterFirst)	National Pacific Mortgage Corp.
	Phoenix Mortgage and Investment
	Residential Funding Corp. (GMAC-RFC)
	Seattle Mortgage Co.
	State Savings Bank
	Temple-Inland Mortgage Corp.
	Trustmark National Bank
	Universal American Mortgage Co.
	Washtenaw Mortgage Co.

NOTES: This table shows initial Loan Prospector and Desktop Underwriter users. The two lists were obtained from American Banker (1995) and (PR Newswire, 1995). We track mergers, acquisitions and name changes over the sample period using data from the National Information Center (NIC). We exclude West Jersey Community Bank from the main analysis as it was acquired by Sovereign Bank early in our sample period. Flagstar Bank was an initial user of both LP and DU. We exclude Flagstar from the set of Desktop Underwriter adopters as they reported relying mainly on Loan Prospector up until at least the late 1990s (LaMalfa, 1997, 1998, 1999).

TABLE 2

SYSTEM CHOICE IS DRIVEN BY PRE-EXISTING BUSINESS RELATIONSHIPS

Dependent variable: Indicator equal to 1 for LP users and 0 for DU users.

	(1)	(2)
Share sold to Freddie	0.34**	0.54**
	(0.13)	(0.23)
Average loan-to-income ratio		0.09
		(0.20)
Portfolio share		0.07
		(0.13)
Thrift or thrift subsidiary		-0.12
		(0.13)
Share bottom quartile income		0.12
		(0.30)
Share LTI $>$ 2.5		-0.15
		(0.17)
Conventional share of originations		0.11
		(0.16)
Refinance share of originations		-0.21
		(0.21)
Number of Observations	27	27
Adjusted R-squared	0.18	0.13

NOTES: This table shows estimated coefficients from $\text{LP}_l = \alpha + \beta X_{l,1992-1993} + \epsilon_l$. LP_l is an indicator equal to 1 for initial Loan Prospector users listed and zero for initial Desktop Underwriter users (see Table 1). Flagstar Bank is classified as a Loan Prospector user as it reported relying mainly on Loan Prospector during the period we analyze. Share sold to Freddie is $\frac{\text{\#Loans Sold to Freddie}}{\text{\#Loans Sold to Fannie or Freddie}}$. Portfolio share is the share of loans originated by the institution that were not sold in the the calendar year of origination. 10%, 5% and 1% significance levels are denoted by *, ** and ***. Sources: HMDA and authors' calculations.

TABLE 3

ZIP CODE CHARACTERISTICS AND EXPOSURE TO INITIAL LOAN PROSPECTOR USERS

	LP_z	High LP_z	LP_z^{IV}
	(1)	(2)	(3)
Share of population living in NOAA coastal county	0.11***	0.00	0.10***
	(0.04)	(0.05)	(0.04)
MSA housing supply elasticity	-0.06	0.00	-0.04
	(0.05)	(0.05)	(0.06)
# HMDA Respondents	-0.14***	0.00	-0.12***
	(0.03)	(0.03)	(0.04)
Market share of large HMDA respondents	-0.02	-0.00	0.01
	(0.05)	(0.05)	(0.05)
Share of originations by thrift	0.03	-0.00	0.08*
	(0.04)	(0.05)	(0.04)
Share of originations by commercial bank	-0.07**	-0.00	-0.04
	(0.03)	(0.04)	(0.03)
Share of originations sold to either Fannie or Freddie	-0.10**	0.00	-0.16***
	(0.05)	(0.04)	(0.05)
Share of originations sold to Freddie in calendar year	0.14***	-0.00	0.19***
	(0.06)	(0.04)	(0.06)
Log median HH income	0.01	-0.00	0.01
	(0.02)	(0.02)	(0.02)
Market share of early LP or DU users by # of loans	0.75***	0.00	0.72***
	(0.16)	(0.05)	(0.17)
Division FE	X	X	X
Weighted for Covariate Balance		X	
Number of ZIP5	6,426	6,426	6,426
Number of Counties	587	587	587
Number of ZIP3	551	551	551
Within R-squared	0.62	-0.00	0.60
Number of Observations	6,426	6,426	6,426

NOTES: Column 1 shows estimated coefficients from: $LP_z = \alpha_d + \beta X_z + \epsilon_z$. LP_z is the 1993 ZIP code market share of Loan Prospector users listed in Table 1 by number of HMDA loans (see Equation 1). The dependent variable in Column 2 is a binary indicator equal to one if the ZIP code is in the top quartile of exposure. In Column 3, the dependent variable is an analogous measure constructed by assigning early AUS adopters to DU or LP based on the share of their GSE loans going to Freddie in 1992. All variables are normalized by dividing by the standard deviation. Standard errors are clustered by CBSA. The sample is restricted to zip codes in metropolitan areas with non-missing FHFA house price data. 10%, 5% and 1% significance levels are denoted by *, ** and ***. Sources: HMDA; 1990 decennial census; NOAA list of coastal counties; authors' calculations.

TABLE 4

LOAN-LEVEL CREDIT EFFECT OF STATISTICAL UNDERWRITING STANDARDS ADOPTION

	High LTI	LTI	High LTI	LTI	LP adopter × Post	High LTI	LTI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LP adopter × Post	1.94***	5.01***	2.60**	5.50***		2.44***	5.06***
	(0.66)	(1.17)	(1.05)	(1.87)		(0.64)	(1.14)
LP adopter × Income Q1 × Post			-1.30	-3.58			
			(2.30)	(4.40)			
LP adopter × Income Q2 × Post			-1.11	0.71			
			(1.75)	(2.89)			
LP adopter × Income Q3 × Post			-0.48	-0.25			
			(1.49)	(2.60)			
Primary Freddie Relationship × Post					0.93***		
					(0.02)		
OLS/2SLS Stage	OLS	OLS	OLS	OLS	1st Stage	2nd Stage	2nd Stage
Lender × Income × Purchaser Type FE	X	X	X	X	X	X	X
ZIP5 × Income Q × Yr × Purchaser Type FE	X	X	X	X	X	X	X
N	675,032	675,032	675,032	675,032	603,660	603,660	603,660
Kleibergen-Paap Wald F stat						2,120	2,120

NOTES: This table compares loans made by initial adopters of Freddie Mac's Loan Prospector (LP) system, which rolled out new underwriting standards, with initial adopters of Fannie Mae's Desktop Underwriter (DU) system, which automated existing manual rules. The sample period is 1992-1996. The dependent variable in Columns 2, 4 and 7 is the loan-to-income ratio (loan size divided by the borrower's income). The dependent variable in Columns 1, 3 and 6 is an indicator equal to 1 if the loan-to-income ratio exceeds 2.5. The dependent variable in Column 5 is the early LP adopter indicator interacted with the post period dummy. *Primary Freddie Relationship* is an indicator equal to 1 if the lender sold more than half of its 1992 GSE loans (i.e. Fannie or Freddie loans) to Freddie. The sample in Columns 5-7 is somewhat smaller because lenders that cannot be matched to a 1992 HMDA ID are dropped. We exclude loans insured by the Farmer Home Administration. Standard errors are clustered by lender × income quartile.

TABLE 5

EFFECT OF LOAN PROSPECTOR ADOPTION ON PROCESSING TIME

Dependent variable: Time in days from application to closing/denial

	Originated			Denied
	All Orig.	Freddie	Non-Freddie	
	(1)	(2)	(3)	(4)
Early LP User X Post	-3.62***	-4.63***	-3.04***	-11.35***
	(0.94)	(0.94)	(1.12)	(2.23)
Lender $\times$ Income Quartile $\times$ Action FE		X		X
Lender $\times$ Income Quartile $\times$ Action $\times$ Purchaser FE	X		X	
Matched Lender Group $\times$ Year $\times$ Action FE		X		X
Matched Lender Group $\times$ Year $\times$ Action $\times$ Purchaser FE	X		X	
ZIP $\times$ Year FE		X		X
ZIP $\times$ Year $\times$ Purchaser FE	X		X	
Number of Observations	531,727	181,671	350,056	35,662

NOTES: This table shows estimates of β from Equation 4. Processing time is the time to origination for originated loans and the time to denial for denied applications. The sample includes originated loans and denied applications reported by the initial LP users in Table 1, and a group of matched control lenders. The sample excludes applications for non-conventional loans. Columns 1 to 3 are restricted to originations and Column 4 is restricted to applications that end in a denial. Column 2 is restricted to loans sold to Freddie Mac in the year of origination. Column 3 is restricted to loans not sold to Freddie Mac in the year of origination. 10%, 5% and 1% significance levels are denoted by *, ** and ***. Standard errors are clustered by lender $\times$ income quartile. Sources: Confidential HMDA; 1990 decennial census; NOAA list of coastal counties; and authors' calculations.

TABLE 6

EFFECT OF DESKTOP UNDERWRITER ADOPTION ON PROCESSING TIME

Dependent variable: Time in days from application to closing/denial

	Originated			Denied
	All Orig.	Fannie	Non-Fannie	
	(1)	(2)	(3)	(4)
Early DU User X Post	-3.01***	-3.37***	-2.54***	-6.19***
	(0.86)	(0.92)	(0.95)	(1.44)
Lender × Income Quartile × Action FE		X		X
Lender × Income Quartile × Action × Purchaser FE	X		X	
Matched Lender Group × Year × Action FE		X		X
Matched Lender Group × Year × Action × Purchaser FE	X		X	
ZIP × Year FE		X		X
ZIP × Year × Purchaser FE	X		X	
Number of Observations	1,111,824	428,612	683,212	151,957

NOTES: This table shows estimates of β from Equation 4. Processing time is the time to origination for originated loans and the time to denial for denied applications. The sample includes originated loans and denied applications reported by the initial DU users in Table 1, and a group of matched control lenders. The sample excludes applications for non-conventional loans. Columns 1 to 3 are restricted to originations and Column 4 is restricted to applications that end in a denial. Column 2 is restricted to loans sold to Fannie Mae in the year of origination. Column 3 is restricted to loans not sold to Fannie Mae in the year of origination. 10%, 5% and 1% significance levels are denoted by *, ** and ***. Standard errors are clustered by lender × income quartile. Sources: Confidential HMDA; 1990 decennial census; NOAA list of coastal counties; and authors' calculations.

TABLE 7

HOUSE PRICE RESPONSE: 1993-1996

	LP_z	Δ log Price				
	(1)	(2)	(3)	(4)	(5)	(6)
Primary Freddie Relationship	0.89***					
	(0.03)					
Early LP Market Share		1.76***	1.82***			
		(0.44)	(0.49)			
Top Quartile Early LP Market Share				1.99***		
				(0.23)		
Matched LP Control (Placebo)					-0.07	
					(0.36)	
Primary Freddie Relationship (Placebo)						-0.05
						(0.41)
OLS/2SLS Stage	1st Stage	OLS	2nd Stage	OLS	OLS	OLS
Division FE	Y	Y	Y		Y	Y
Covariate Balance				Y		
N	6,426	6,426	6,426	6,426	6,426	6,426
Kleibergen-Paap Wald F stat			720			

NOTES: Columns 1 to 3 show OLS and IV estimates of the effect of statistical lending standards adoption on house prices. In Column 4 we use a binary exposure variable. We define ZIP codes with top quartile exposure as the 'treated' group and other ZIP codes as 'control' group. We weight each ZIP code such that covariates are balanced across the two groups. The difference in average exposure across the two groups is 1.25 standard deviations. Columns 5 and 6 show placebo tests. *Early LP Market Share* (LP_z) is the zip code exposure measure defined in Equation 1. *Primary Freddie Relationship* is an alternative version of the ZIP code exposure measure. Rather than using observed choice of LP (rather than DU), it uses an indicator equal to 1 if the lender sold more than half of its pre-AUS release (1992) GSE loans (i.e. Fannie or Freddie loans) to Freddie. *Matched LP Control (Placebo)* is a placebo exposure measure computed using matched lenders with a similar business model to each early adopter of Loan Prospector. *Primary Freddie Relationship (Placebo)* is computed analogously to *Primary Freddie Relationship*, but uses the set of matched control lenders rather than early AUS adopters. This is to test whether 1992 Freddie relationships are acting through some channel other than choice of AUS. We include census division fixed effects and condition on the following variables, at the ZIP code level unless otherwise indicated: (county) coastal indicator, number of lenders, large lender market share, thrift market share, commercial bank market share, share of originations sold to either Fannie or Freddie, share of originations sold to Freddie, log median household income, Saiz (2010) housing supply elasticity, the combined ZIP code share of initial DU and LP users AUS_z. The sample includes ZIP codes in metropolitan counties with FHFA house price data available continuously from 1990. Standard errors are clustered by CBSA. Sources: FHFA HPI; HMDA 1990 decennial census; BEA; NOAA list of coastal counties; and authors' calculations.

TABLE 8

RESPONSE OF HOUSE PRICE TO CREDIT

	Δ log LTI (1)	Δ log Price (2)	Δ High LTI share (3)	Δ log Price (4)
Δ Log LTI		0.43**		
		(0.17)		
Δ High LTI Share				1.83***
				(0.50)
Early LP	4.11***		0.96***	
	(1.43)		(0.27)	
Stage	1st Stage	2nd Stage	1st Stage	2nd Stage
Division FE	Y	Y	Y	Y
N	6,423	6,423	6,423	6,423
Kleibergen-Paap Wald F stat		8		12

NOTES: This table shows IV regressions of log house price changes on instrumented credit measures. All changes are computed from 1993 to 1996. Columns 1 and 2 show the response of house prices to the average loan-to-income ratio of home purchase originations (the top and bottom 1 per cent of income and loan size distributions each year are dropped before computing LTI). Columns 3 and 4 show the response of house prices to the share of high loan-to-income lending. The dependent variable in Column 3 is the change in the share of home purchase originations with a loan-to-income ratio above 2.5. All specifications include census division fixed effects and standard errors are clustered by CBSA. We condition on the following variables, at the ZIP code level unless otherwise indicated: (county) coastal indicator, number of lenders, large lender market share, thrift market share, commercial bank market share, share of originations sold to either Fannie or Freddie, share of originations sold to Freddie, log median household income, Saiz (2010) housing supply elasticity, the combined ZIP code share of initial DU and LP users AUS_z. The sample is restricted to zip codes in metropolitan areas with non-missing FHFA house price data. 10%, 5% and 1% significance levels are denoted by *, ** and ***. Sources: HMDA; FHFA HPI; 1990 decennial census; NOAA list of coastal counties; and authors' calculations.

TABLE 9

EFFECT OF AUTOMATED UNDERWRITING ADOPTION ON COUNTY HOUSE PRICE AND LENDING CORRELATION

	House Price Correlation			
	1984-1998		1984-2003	
	(1)	(2)	(3)	(4)
Integration (LP) × Post	0.27**	0.27**	0.39***	0.39***
	(0.11)	(0.11)	(0.10)	(0.10)
Integration (DU) × Post	0.02	0.02	0.12***	0.11***
	(0.04)	(0.05)	(0.04)	(0.04)
Population corr. (1SD)	1.66***	1.65***	1.96***	1.96***
	(0.33)	(0.33)	(0.34)	(0.34)
Per capita income corr. (1SD)	0.22	0.22	0.75**	0.75**
	(0.41)	(0.41)	(0.34)	(0.34)
Loan type integration (1SD)		1.13**		1.68***
		(0.50)		(0.51)
County Pair FE	X	X	X	X
CSA (i) × CSA (j) × Year FE	X	X	X	X
Sample	1984-1998	1984-1998	1984-2003	1984-2003
Number of Observations	596,742	596,742	795,656	795,656

NOTES: This table shows the effect of systems-induced integration on county house price growth correlation. The base period is 1989-1993. For county i (in Combined Statistical Area $a(i)$) and county j (in Combined Statistical Area $a(j)$) and year t:

$$Corr_{i,j,t} = \alpha_{i,j} + \gamma_{a(i),a(j),t} + \beta_1 Post_t \cdot \Delta H_{i,j}^{LP} + \beta_2 Post_t \cdot \Delta H_{i,j}^{DU} + Controls_{i,j,t} + \epsilon_{i,j,t}$$

where $Corr_{i,j,t}$ is the annual growth rate correlation between county i and county j over a 5-year sample period starting in year t, where $t \in \{1984, 1989, 1994, 1999\}$. The post indicator is equal to 0 for $t \in \{1984, 1989\}$ (1989 includes house price growth up until 1993) and 1 for later years. Controls include the annual population growth and per capita income growth correlation computed over the same 5-year period as the dependent variable. In Columns 2 and 4 we also control for an additional measure of lending integration based on the shares of loans sold to Fannie Mae, Freddie Mac and the share of FHA loans. $\Delta H_{i,j}^{LP}$ and $\Delta H_{i,j}^{DU}$ are trimmed at the 1st and 99th percentile. Standard errors are clustered by Combined Statistical Area of county i and Combined Statistical Area of county j. All specifications include fixed effects for time by the CSA of county i by the CSA of county j.

Internet Appendix

Credit and House Price Effects of Automated Underwriting Adoption

A Expectations Feedback

For transparency, Figure A.7 plots the cumulative price response to 2005. Our main results focus on the response to 1996 as longer-run effects are more speculative. In this appendix we show that the larger long-term price response can potentially be explained by feedback to expectations. Combining adaptive expectations with the short-run response can approximately replicate the observed price response profile. Intuitively, households in more exposed locations observe higher recent house price growth following system adoption. If these households naively extrapolate, past growth lowers their perceived cost of housing and increases housing demand on top of the initial direct effect of automated underwriting.

We assume that the price response up until 1995 is the direct effect of adopting automated underwriting. Starting in 1996, we update expectations according to an adaptive rule. We also estimate the housing supply response using the Census Building Permits Survey and incorporate this when updating house prices. By assumption, the house price response is zero up until 1993, and then matches the estimated price effect to 1995.

A.1 Housing demand

We assume nominal housing demand of unconstrained households with income y_i is given by:

$$P \cdot H_i = \frac{\alpha_i y_i}{\theta + r + \delta - g^e} \tag{A.1}$$

That is, each (unconstrained) household has a constant housing budget share equal to α_i. The denominator on the right hand side of A.1 is the user cost: θ is property taxes and insurance as a share of property value; r is the mortgage interest rate, which we also assume is equal to the opportunity cost of home equity for convenience; δ is the depreciation rate; and g^e is expected annual house price growth.

Next we compute aggregate nominal housing demand. For this, we also need to consider the demand of households who are constrained. We denote the observed housing budget share by $\hat{\alpha}_i$. For unconstrained households $\alpha_i = \hat{\alpha}_i$, but for constrained households $\alpha_i > \hat{\alpha}_i$. For simplicity we assume $y_i = \bar{y}\ \forall i$:

$$P \cdot H = P \sum_{i=1}^{N} H_i = N \frac{\bar{y} \frac{\sum_{i=1}^{N} \hat{\alpha}_i}{N}}{\theta + r + \delta - g^e} = N \frac{\alpha \bar{y}}{\theta + r + \delta - g^e} \tag{A.2}$$

From here on we use α to denote the aggregate observed housing budget share.

A.2 Using the short-run price response to back out the change in fundamentals

When Loan Prospector allows households to borrow more relative to their income, this increases α and, therefore, housing demand. We take our estimated short-run price response and use it to back out the implied change in α for 'treated' locations. We distinguish between 'treated' locations ($\tilde{\alpha}$) and control locations (α). We start with the following decomposition of the change in log nominal housing demand:

$$\Delta \log(PH) = \Delta \log P + \Delta \log H \tag{A.3}$$

We then substitute our estimated short-run price and housing supply responses as follows:

$$\Delta_{1993-1995} \log(PH) = \hat{\beta}^P_{1995} + \hat{\beta}^H_{1995} \tag{A.4}$$

Where $\hat{\beta}^{P}_{1995}$ is from Equation 5 and $\hat{\beta}^{H}_{1995}$ is described when we discuss calibration below. Assuming that there is no short-run effect on the user cost, $\theta + r + \delta - g^{e}$, we use $\Delta \log(PH) = \Delta \log(\alpha)$ from Equation A.2 combined with a calibrated value for $\alpha_{1993}(= \tilde{\alpha}_{1993})$ to obtain $\tilde{\alpha}_{1995}$. We then keep $\tilde{\alpha}_{t} = \tilde{\alpha}_{1995}$ permanently. That is, we assume our natural experiment generates a one-off permanent relaxation of borrowing constraints between 1993 and 1995 for the 'treated' group. We further assume that the 'control' group (initial DU users) eventually shifts to using the same lending standards as the 'treatment' group, consistent with narrative evidence.

A.3 House price growth expectations

Beginning in 1996, we update households' growth expectations. We assume that households form expectations at the start of year t about price growth during the year as follows:

$$g^{e}_{t} = \lambda \sum_{j=1}^{t-t_0} (1-\lambda)^{j-1} g_{t-j} \tag{A.5}$$

As we will be making a comparison with our difference-in-differences estimates, we next rewrite Equation A.5 to separate out the part of expectations that comes from the 'treatment'. We denote expectations in the 'treated' location by $\tilde{g}^{e}_{t}$ and write it as a function of our estimated price response Effect_{t} (expressed as an annual percentage change):

$$\begin{aligned} \tilde{g}^{e}_{t} &= \lambda \sum_{j=1}^{t-t_0} (1-\lambda)^{j-1} g_{t-j} + \lambda \sum_{j=1}^{t-t_0} (1-\lambda)^{j-1} \text{Effect}_{t-j} \\ \Rightarrow \tilde{g}^{e}_{t} &= g^{e}_{t} + \Delta_{t} \end{aligned} \tag{A.6}$$

That is, expectations in the 'treated' location are equal to expectations in the 'control' location plus $\Delta_t = \lambda \sum_{j=1}^{t-t_0}(1-\lambda)^{j-1}\text{Effect}_{t-j}$. Recalling our estimated responses are cumulative log changes, we compute the annual percentage price response as:

$$\text{Effect}_t = e^{\hat{\beta}_t - \hat{\beta}_{t-1}} - 1 \tag{A.7}$$

A.3.1 Implied difference-in-differences house price response

Next, we compute the cumulative implied (difference-in-differences) effect on nominal housing demand relative to 1993.[1]

$$\Delta \log(PH)_t = \log\left(\frac{\tilde{\alpha}_t}{\alpha_t}\right) + \left[\log(\theta_t + r_t + \delta - g_t^e) - \log(\theta_t + r_t + \delta - g_t^e - \Delta_t)\right] \tag{A.8}$$

The first term on the right hand side of Equation A.8 is the direct effect due to Loan Prospector adoption. This diminishes over time due to eventual adoption in the 'control' group. The term in square brackets reflects feedback to growth expectations. Finally, we combine Equations A.3 and A.8 with our estimated supply response to back out the cumulative log price response at the end of year $t \geq 1996$:

$$\Delta \log P_t = \Delta \log(PH)_t - \Delta \log H_t = \Delta \log(PH)_t - \hat{\beta}_t^H \tag{A.9}$$

[1] In 'treated' locations the cumulative log change in demand is: $\log(\frac{\tilde{\alpha}_t}{\tilde{\alpha}_{1993}}) + \log(\theta_{1993} + r_{1993} + \delta - g_{1993}^e) - \log(\theta_t + r_t + \delta - g_t^e - \Delta_t)$. In 'control' locations the cumulative log change in demand is: $\log(\frac{\alpha_t}{\alpha_{1993}}) + \log(\theta_{1993} + r_{1993} + \delta - g_{1993}^e) - \log(\theta_t + r_t + \delta - g_t^e)$. Taking the difference between treated and control locations gives Equation A.8.

A.4 Calibration

Growth expectations (λ, g^e_t)

Our goal here is to show that our estimated response profile *could* be generated from a one-off relaxation of constraints due to Loan Prospector adoption. We perform a search for the value of the expectations parameter λ that best matches our response profile using a least squares criterion. For each value of λ, we first compute g^e_t using the FHFA U.S. All Transactions HPI before computing $\Delta \log P_t$ as described above.[2] We find that $\lambda = 0.091$ best matches the data response profile.

Non-growth user cost components (θ_t, r_t, δ)

We use data on property tax and insurance costs from the American Housing Survey to calibrate θ_t. We divide each household's total property tax and insurance costs by their property value and set θ_t equal to the average value for owner-occupiers. $\theta_{1993} = 0.018$ and is broadly similar over the sample period. We assume an interest rate (and opportunity cost of home equity) equal to the prevailing Freddie Mac interest rate on a 30-year fixed rate mortgage. This is 7.31% in 1993.[3] We assume an annual depreciation rate of $\delta = 0.02$.

Initial housing budget share (α_{1993})

We compute $\alpha_{1993} = \left(\frac{P \cdot H}{y}\right)_{1993} \cdot (\theta_{1993} + r_{1993} + \delta - g^e_{1993})$. We set $\left(\frac{P \cdot H}{y}\right)_{1993}$ equal to the 1993 AHS average house price to income ratio of 3.12. We use θ_{1993}, r_{1993}, δ and g^e_{1993} as

[2] When computing g^e_t using observed price data, we use a modified formula is needed so that weights sum to one when working with a finite price history: $g^e_t = \frac{\lambda}{1-(1-\lambda)^{t-t_0}} \sum_{j=1}^{t-t_0} (1-\lambda)^{j-1} g_{t-j}$. As we have national price data back to 1975, the difference relative to Equation A.5 is not very large by the mid 1990s.

[3] The timeseries can be downloaded here: `https://fred.stlouisfed.org/series/MORTGAGE30US`.

described above. This gives us $\alpha_{1993} = 0.18$.

Housing supply response ($\hat{\beta}^H_t$)

We use county Census data on permits issued for new housing units to estimate the supply response using Equation 5 with $\log(Permits_{c,t})$ as the dependent variable and county level controls. We use building permits data as we only observe the outstanding housing stock at the 1990 and 2000 censuses.[4] Figure A.9A shows an increase in the flow of building permits.

We use the path of housing supply from Figure A.9B to calibrate the model. To obtain this we convert the flow estimates from Figure A.9A into the cumulative percentage change in implied total housing units. The cumulative increase in outstanding housing stock for a one standard deviation increase in exposure from 1993 to 2003 is 0.24%. Incorporating the county level price response of 1.2% for a one standard deviation increase in exposure, this implies a decade housing supply elasticity of around 0.2 (We estimate the county-level price response using Equation 5 and comparable county-level controls). While this is lower than the MSA elasticities estimated in Saiz (2010); our period of analysis is also shorter than the 30-year period used there. We note that Baum-Snow and Han (2024) have estimated more modest local supply elasticities from 2000 to 2010.[5]

A.5 Results

Figure A.10 compares our estimated response from Figure A.7 (solid line) with the response generated by applying adaptive expectations to the short-run effect (dashed line).

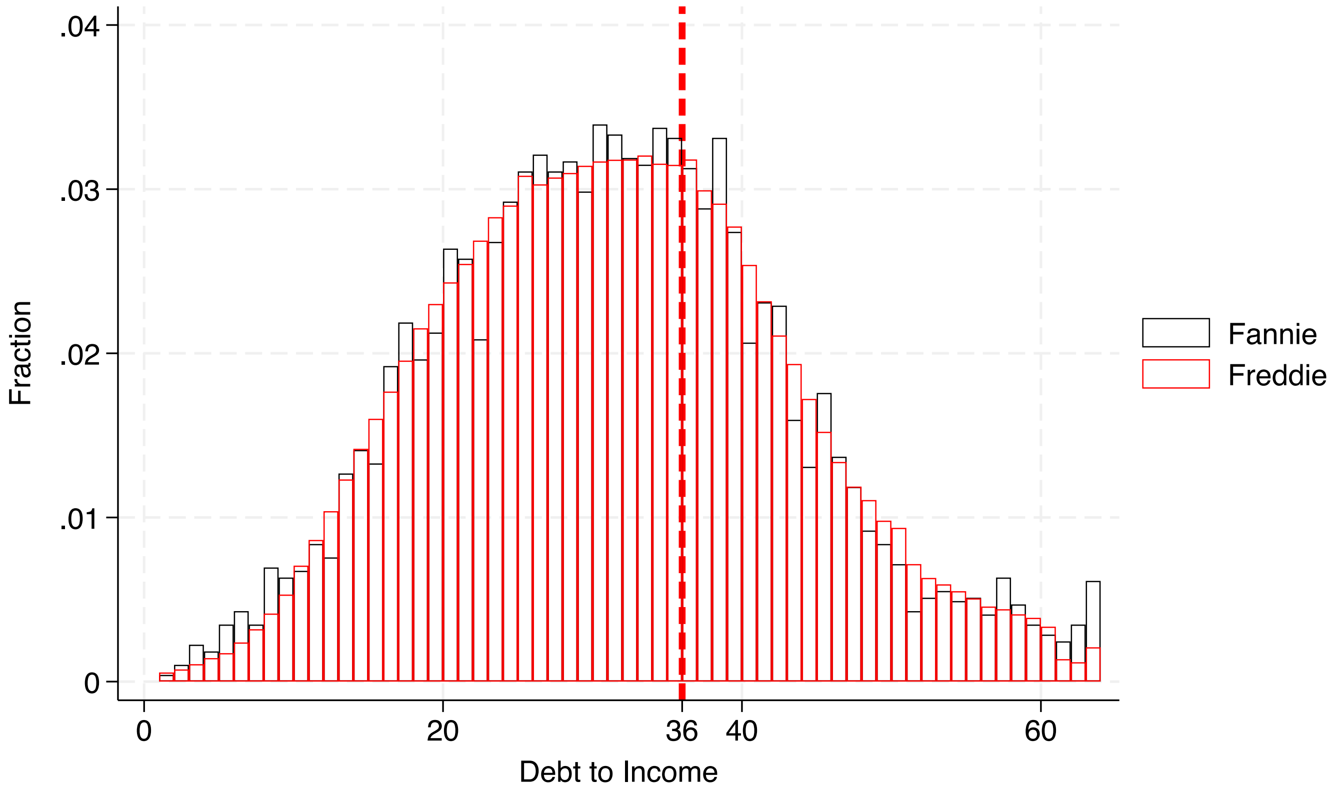
[4] Figure A.8 shows the relationship between actual and imputed growth in housing units from the 1990 to 2000 censuses.

[5] Baum-Snow and Han (2024) explore the elasticity of housing supply at the neighborhood level and find a unit elasticity of 0.41.

We are broadly able to match the shape of the data response. Although the model generated response is conditional on a number of strong simplifying assumptions, we believe this exercise enhances the plausibility of our estimates. It also illustrates how our large estimated long-run response could reflect a fundamental direct response combined with feedback through an expectations channel, consistent with Chodorow-Reich, Guren and McQuade (2024).

FIGURE A.1

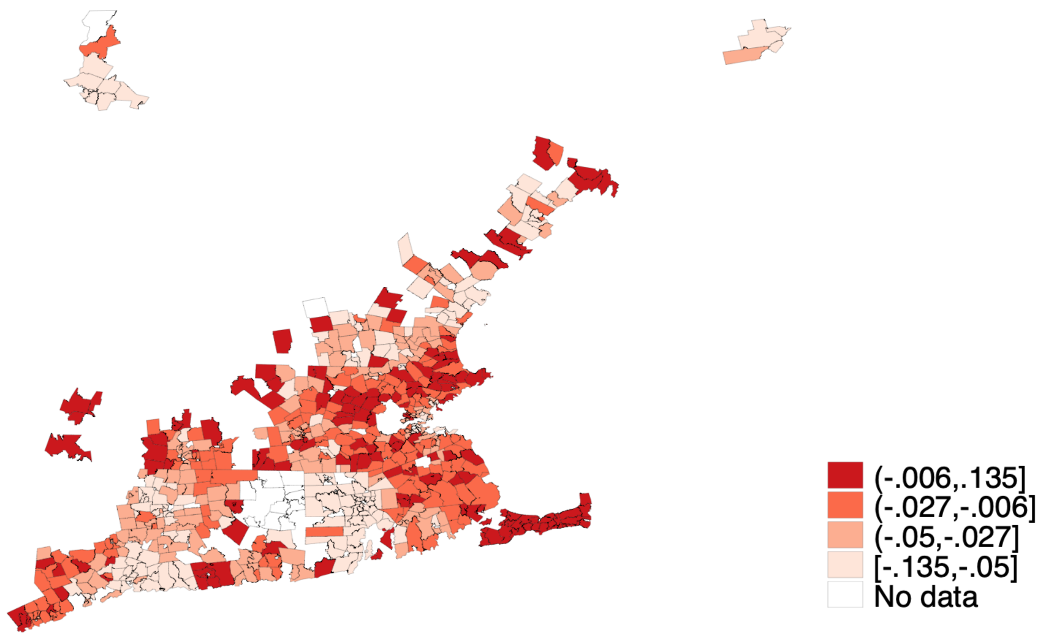
DTI distribution of loans in public GSE data originated between Jan and Jun 1999



Notes: Figure A.1 plots the distribution of the back-end debt-to-income ratio using Fannie Mae's Single Family Loan Performance Dataset and Freddie Mac's Loan Level Dataset. We also include loans that were excluded from the originally released datasets to obtain more comprehensive coverage. We also drop loans sold by Norwest due to some unique characteristics of the Norwest distribution (Norwest reached an agreement to exclusively sell to Freddie around this time period in exchange for using an alternative underwriting system instead of LP).

FIGURE A.2

ZIP CODE EXPOSURE VARIATION (CT, ME, MA, NH, RI, VT)

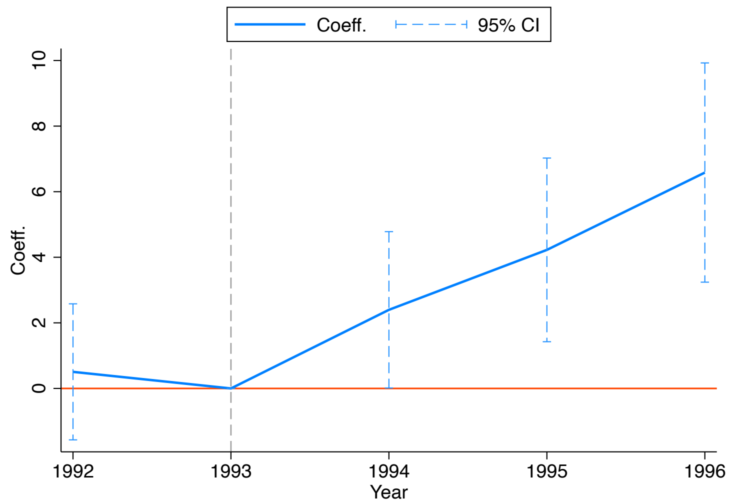


NOTES: Variation in residualized exposure across 5-digit zip codes. The exposure measure is only plotted for zip codes with a sufficiently long price history to be included in our sample.

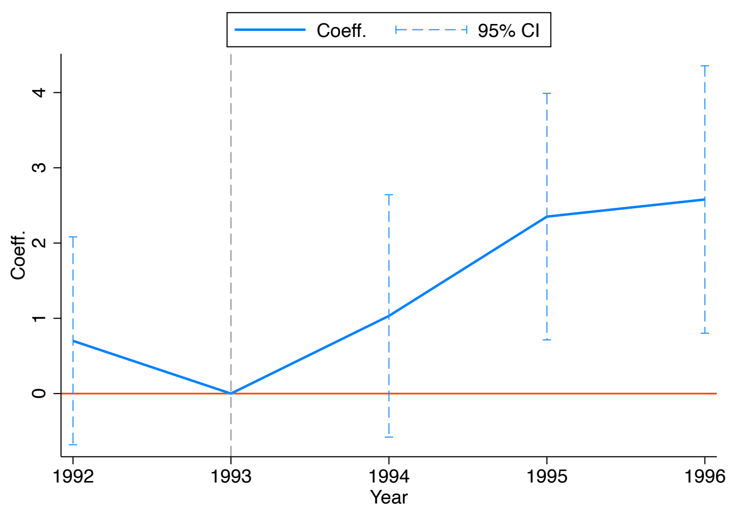
FIGURE A.3

STATISTICAL UNDERWRITING STANDARDS INCREASE BORROWING RELATIVE TO INCOME (WITHIN THE SAME ZIP5, INCOME GROUP AND LOAN TYPE)

Panel A. Loan-to-income ratio (pp)

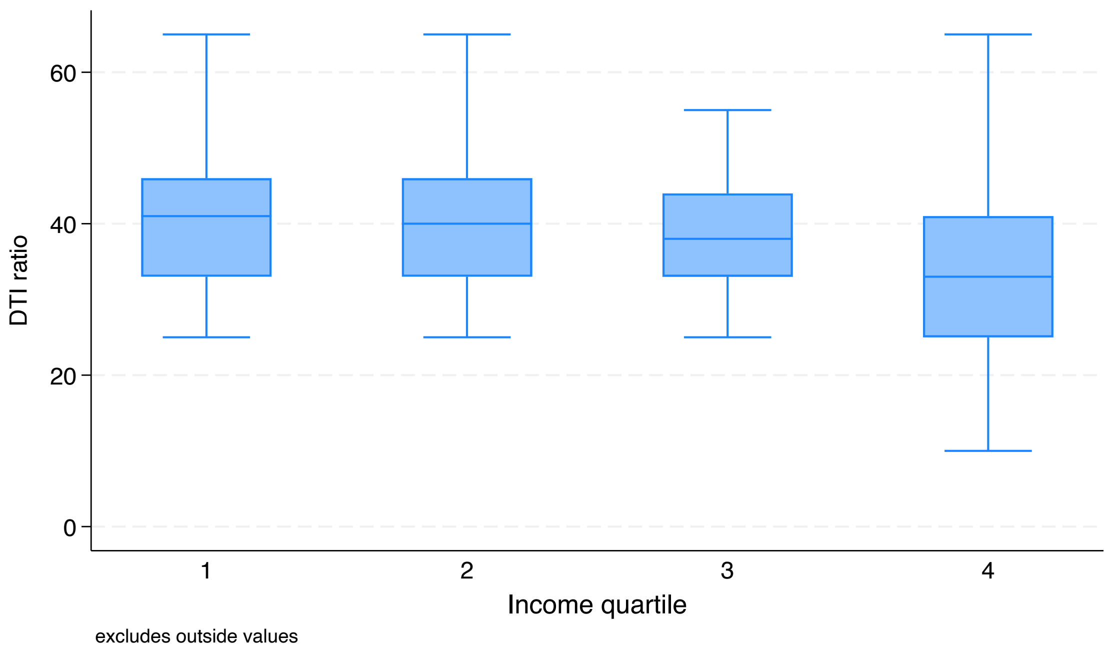


Panel B. Share loan-to-income ratio > 2.5 (pp)



NOTES: This figure plots estimates $\{\beta_k\}$ from a difference-in-differences specification that includes ZIP5 (z) $\times$ year (t) $\times$ income quartile (n) $\times$ purchaser type (p) fixed effects and lender (l) $\times$ income quartile $\times$ purchaser type fixed effects: $Y_i = \alpha_{l,n,p} + \gamma_{z,n,p,t} + \sum_{k \neq 1993} \mathbb{1}_{t=k} \beta_k LP_l + \epsilon_i$. The dependent variable in Figure A.3A is loan size divided by income. Figure A.3B plots the percentage point response of the share of loans with a LTI > 2.5. The sample is HMDA purchase originations to owner-occupiers reported by initial users of Loan Prospector or Desktop Underwriter. Standard errors are clustered by lender $\times$ income quartile. Sources: HMDA and authors' calculations.

FIGURE A.4
DEBT-TO-INCOME DISTRIBUTION BY INCOME QUARTILE

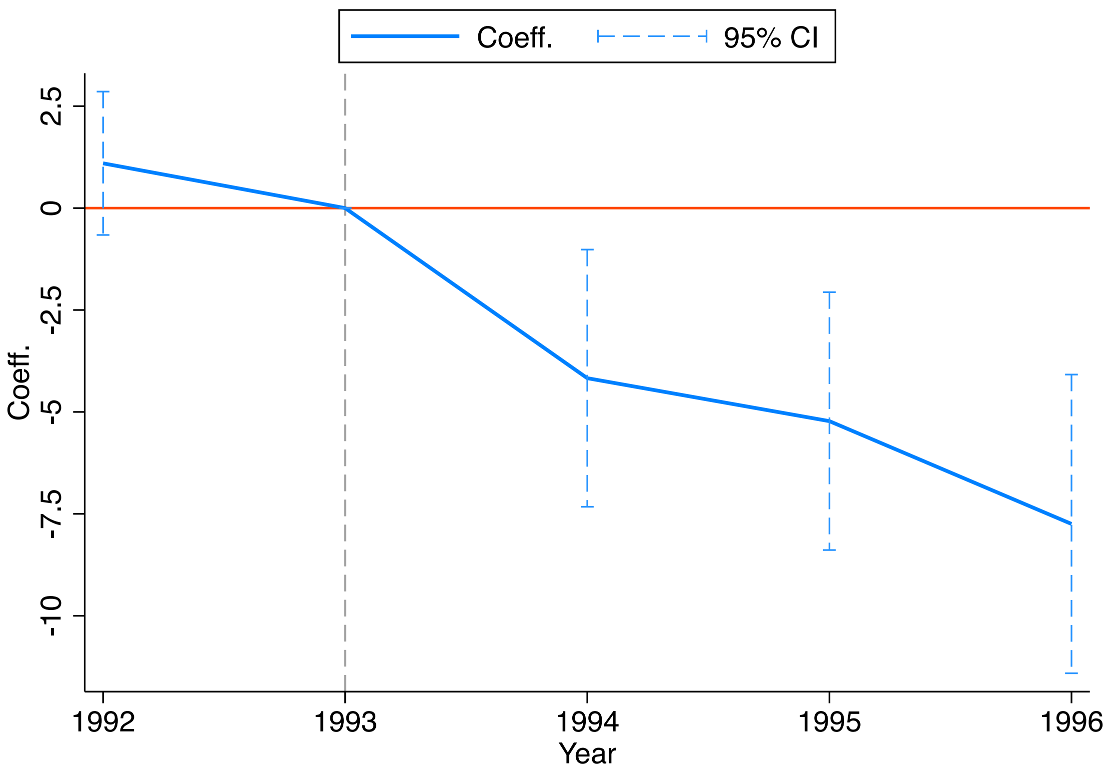


NOTES: This figure is constructed using 2018 HMDA home purchase originations. DTI ratios are top-coded at 60 per cent. Where DTI ratios are binned we assume the DTI is equal to the bin midpoint, and for > 60 we use 65%. DTI is reported by the lender and is the ratio of monthly mortgage payments, property insurance, property taxes, debt payments and certain other financial obligations to gross monthly income. Sources: HMDA.

FIGURE A.5

EFFECT OF EARLY LOAN PROSPECTOR ADOPTION ON MORTGAGE DENIAL PROBABILITY

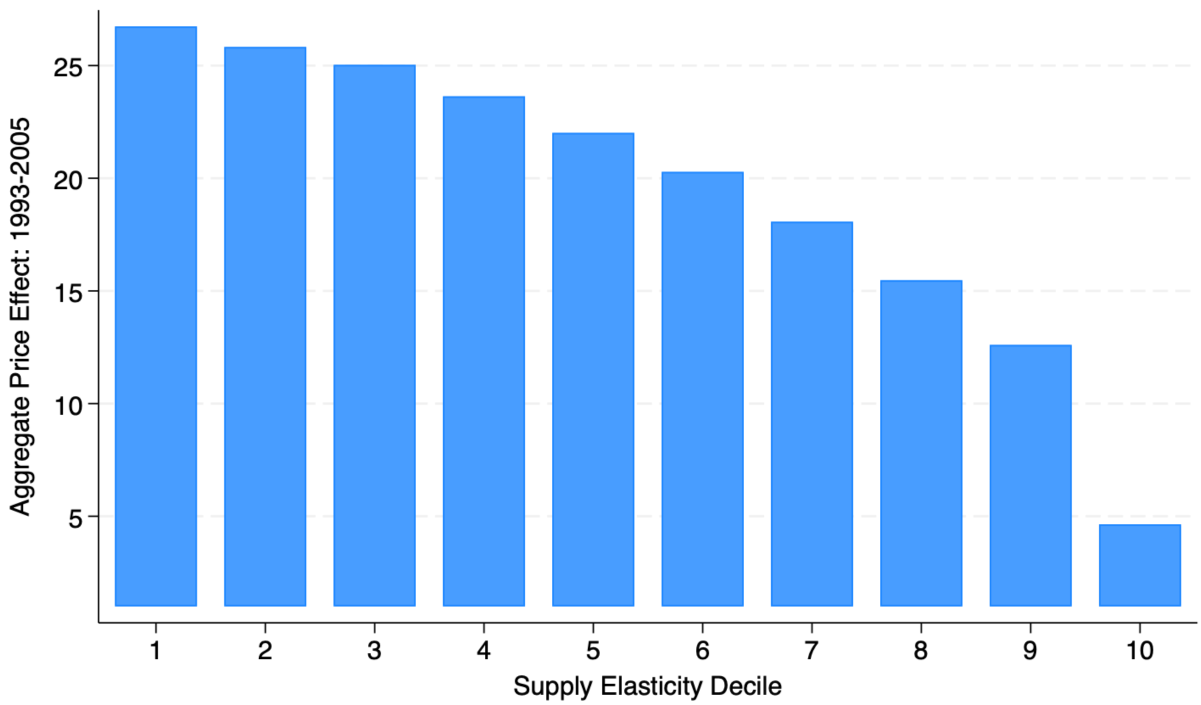
Panel A. Denial probability



NOTES: This figure plots estimates $\{\beta_k\}$ from a difference-in-differences specification that includes ZIP5 (z) $\times$ year (t) $\times$ income quartile (n) fixed effects and lender (l) $\times$ income quartile fixed effects: $Y_i = \alpha_{l,n} + \gamma_{z,n,t} + \sum_{k \neq 1993} \mathbb{1}_{t=k} \beta_k LP_l + \epsilon_i$. The dependent variable is an indicator equal to one if the application is denied and zero otherwise, so estimates are percentage point changes in denial probability relative to 1993. The sample is HMDA purchase applications from owner-occupiers reported by initial users of Loan Prospector or Desktop Underwriter. Standard errors are clustered by lender $\times$ income quartile. Sources: HMDA and authors' calculations.

FIGURE A.6

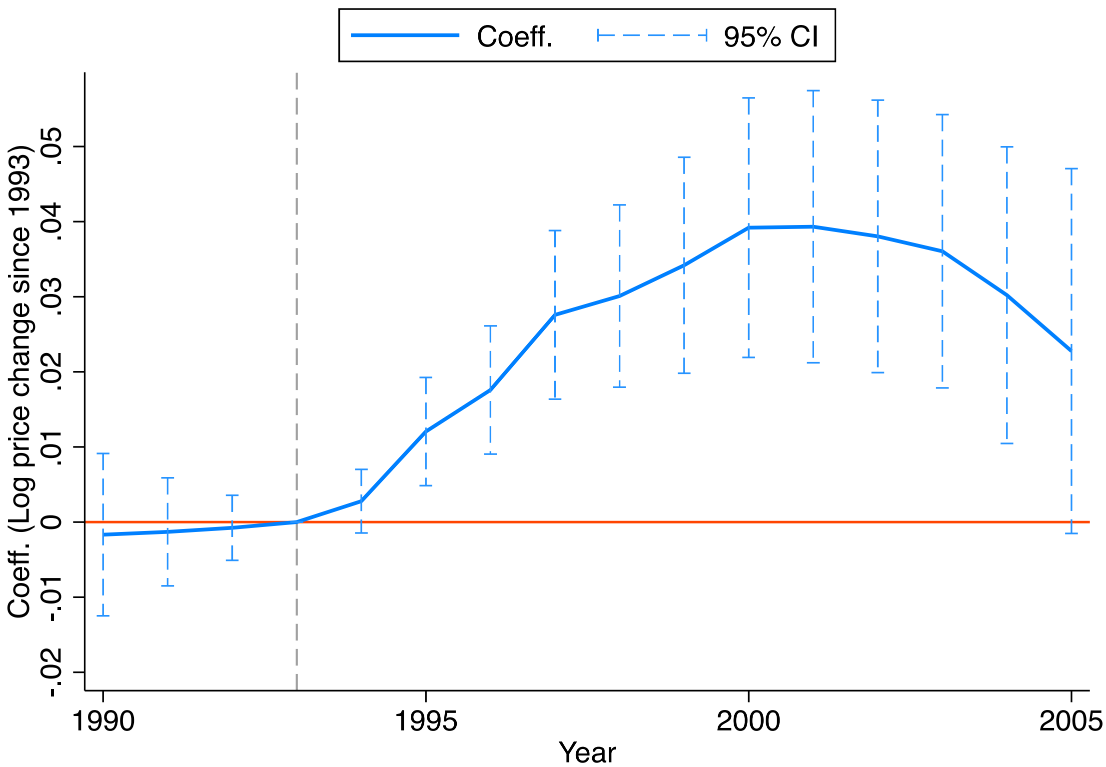
AGGREGATE PRICE EFFECT BY HOUSING SUPPLY ELASTICITY: 1993-2005



NOTES: We compute the aggregate price effect in supply elasticity decile i as follows: $\sum_{k=1994}^{2005} s_k \times ME^i_{1994+(2005-k)}$ where s_k is year k adoption implied from aggregate usage statistics, and ME^i is the marginal effect of early Loan Prospector adoption in elasticity bin i. We set $ME^i_t = ME^i_{1996}$ for all $t \geq 1996$. Before aggregating over the supply elasticity distribution, we assign locations with missing elasticity to the top bin.

FIGURE A.7

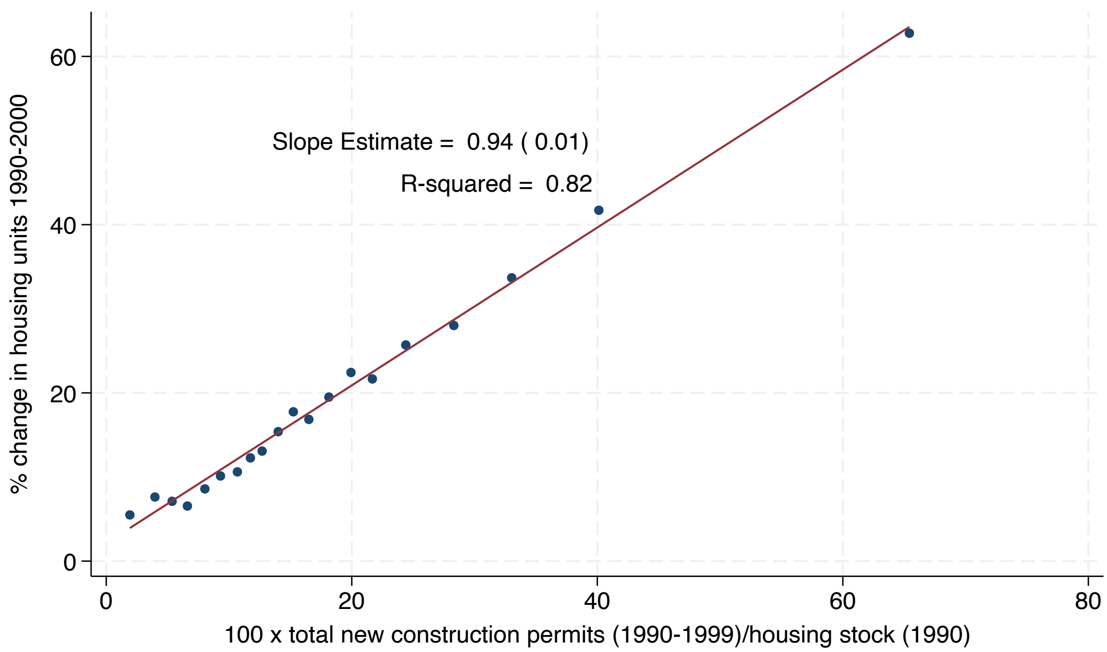
CUMULATIVE EFFECT OF LOAN PROSPECTOR ADOPTION ON HOUSE PRICES



NOTES: This figure plots estimates of $\{\beta_k\}$ from Equation 5. We include census division by year fixed effects and condition on the following variables interacted with year dummies, at the zip code level unless otherwise indicated: (county) coastal indicator, number of lenders, large lender market share, thrift market share, commercial bank market share, share of originations sold to either Fannie or Freddie, share of originations sold to Freddie, log median household income, Saiz (2010) housing supply elasticity, the combined zip code share of initial DU and LP users AUS_z. The sample includes zip codes in metropolitan counties with FHFA house price data available continuously from 1990 to 2005. Coefficients are interpreted as cumulative log price changes relative to 1993. Standard errors are clustered by CBSA. Sources: FHFA HPI; HMDA 1990 decennial census; BEA; NOAA list of coastal counties; and authors' calculations.

FIGURE A.8

USING THE BUILDING PERMITS SURVEY TO IMPUTE HOUSING UNITS



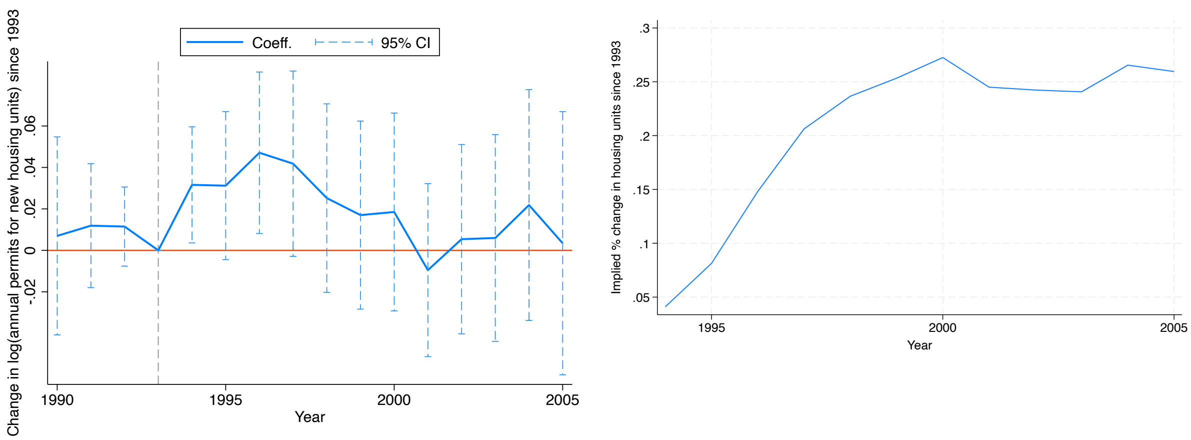
NOTES: Figure A.8 shows the relationship between actual and imputed growth in housing units from the 1990 to 2000 censuses. The vertical axis shows the actual percentage change in the number of housing units in a county. The horizontal axis shows the imputed growth as measured by total new construction permits in the BPS from 1990 to 1999.

FIGURE A.9

HOUSING SUPPLY RESPONSE MEASURED USING ANNUAL PERMITS ISSUED FOR NEW UNITS

Panel A. Cumulative log change in annual new construction permits

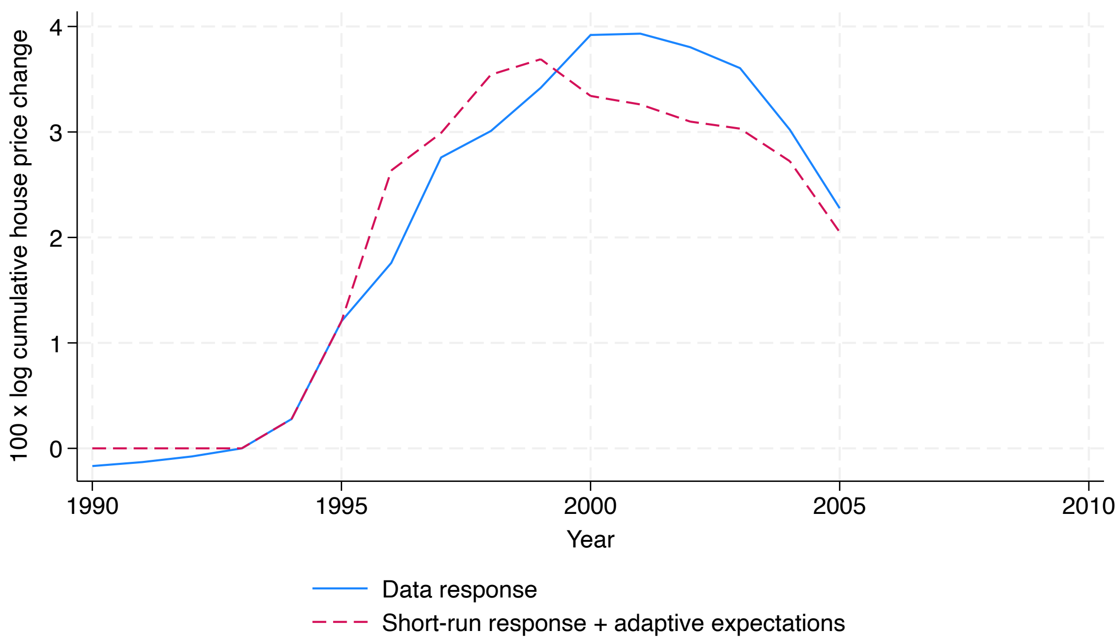
Panel B. Cumulative % change in total housing units



NOTES: Figure A.9 combines the estimates from $\log(Permits_{c,t}) = \delta_c + \gamma_{d,t} + \sum_{\substack{k=1990 \\ k\neq 1993}}^{2005} \mathbb{1}_{t=k}\left(\beta_k \frac{EarlyLP_c}{SD(EarlyLP_c)} + \alpha_k X_c\right) + \epsilon_{c,t}$ with 1990 data on the total county housing stock to plot the cumulative change in housing units.

FIGURE A.10

EXPLAINING THE LONG-RUN PRICE RESPONSE WITH ADAPTIVE EXPECTATIONS



NOTES: The blue line in Figure A.10 shows the data response from Figure A.7. The red line plots the model price response implied by combining (1) the short-run estimated price response (to 1995); (2) adaptive expectations applied to the short-run response from 1996 onwards; (3) the housing supply response estimated from data on building permits over the entire sample (4) eventual AUS adoption in the 'control' group.

TABLE A.1

% OF FANNIE AND FREDDIE PURCHASES PROCESSED USING DESKTOP UNDERWRITER OR LOAN PROSPECTOR

	Fannie Mae		Freddie Mac	
	Report	Other	Report	Other
1995				
1996			25	
1997	9		22	54*
1998	22	26*	36	
1999	39		50	>75**
2000	56		56	
2001	59		62	
2002	60		60	
2003			64	
2004			61	

NOTES: This table shows the share of Fannie's purchases processed through DU and the share of Freddie's purchases processed through LP. The table includes numbers from the GSEs' annual reports and numbers reported by Fannie and Freddie representatives to trade journals. The discrepancies these two sources could reflect fluctuations in LP and DU usage within the calendar year, and differences between projected and realized usage. In particular, there is evidence that both Fannie and Freddie projected usage of 80-85% by 1999. These rates were apparently never realized, though during 1999 Freddie stated that over 75% of its purchases were processed through LP. Later annual reports suggest that DU and LP usage stabilized at a lower rate of around 60 per cent because both Fannie and Freddie made agreements with large lenders which allowed them to use alternative systems. Sources: Fannie Mae and Freddie Mac annual reports.

* Wilson, Caroline (1998). Automated Underwriting Goes Mainstream. America's Community Banker, 7(4):36; Gallaher, Douglas (1998). Getting a Payoff from Technology. Mortgage Banking, 58(6): 66–76.

** Murin, Joseph (1999). A Business Transformed by Technology. Mortgage Banking, 60(1): 152.

TABLE A.2

HOW IS SYSTEM CHOICE RELATED TO LENDER CHARACTERISTICS?

Dependent: 1 for initial adopters and 0 for matched control lenders.

	LP (1)	DU (2)
Share sold to Fannie or Freddie (1991-1993)	0.16	0.67**
	(0.12)	(0.26)
Average loan-to-income ratio (1991-1993)	0.12	-0.05
	(0.13)	(0.21)
Share LTI > 2.5 (1991-1993)	0.01	0.16
	(0.14)	(0.19)
Portfolio share (1991-1993)	0.11	0.14*
	(0.12)	(0.08)
Share bottom quartile income (1991-1993)	0.07	-0.01
	(0.11)	(0.07)
Conventional share of originations (1991-1993)	-0.07	-0.07
	(0.12)	(0.09)
Refinance share of originations (1991-1993)	0.03	-0.02
	(0.12)	(0.07)
Share of loans purchased (1991-1993)	-0.05	0.04
	(0.09)	(0.06)
Adjusted R-squared	-0.04	0.03
Number of Observations	36	68

NOTES: This table shows estimated coefficients from $Y_l = \alpha + \beta X_l + \epsilon_l$. In Column 1, Y_l is an indicator equal to 1 for initial Loan Prospector users and zero for matched control lenders (who were not initial users of either system). In Column 2, Y_l is an indicator equal to 1 for initial Desktop Underwriter users and zero for matched control lenders. Flagstar Bank is classified as a Loan Prospector user as it reported relying mainly on Loan Prospector during the period we analyze. Portfolio share is the share of loans originated by the institution which were not sold in the the calendar year of origination. 10%, 5% and 1% significance levels are denoted by *, ** and ***. Sources: HMDA.

TABLE A.3

CREDIT OUTCOMES ARE DRIVEN BY AFFECTED LOAN TYPES

	Excl. Fannie, FHA, VA		Fannie, FHA, VA	
	High LTI	LTI	High LTI	LTI
	(1)	(2)	(3)	(4)
LP adopter × Post	5.1***	8.9***	-0.1	2.3**
	(1.1)	(2.0)	(0.6)	(1.1)
Estimation	OLS	OLS	OLS	OLS
Lender × Income Q × Purchaser FE	X	X	X	X
ZIP5 × Income Q × Yr × Purchaser FE	X	X	X	X
N	277,938	277,938	380,135	380,135

NOTES: Columns 1 and 3 show estimates from $HighLTI_i = \alpha_{l,n,p} + \gamma_{z,n,p,t} + \beta LP_l \times Post_t + \epsilon_i$, where l is the lender, n is the income quartile, p is the purchaser type and t is the year. Columns 2 and 4 show estimates from $LTI_i = \alpha_{l,n,p} + \gamma_{z,n,p,t} + \beta LP_l \times Post_t + \epsilon_i$. Columns 1 and 2 exclude loans for which the purchaser is indicated to be Fannie Mae or for which the loan is FHA-insured or VA-guaranteed. Columns 3 and 4 restrict the sample only to loans for which the purchaser iis indicated to be Fannie Mae or for which the loan is FHA-insured or VA-guaranteed. We exclude loans insured by the Farmer Home Administration. Standard errors are clustered by lender × income quartile.

TABLE A.4

CORRELATION BETWEEN AUS USAGE AND PROCESSING TIME

Dependent variable: Days from application to closing/denial.

	Originated		Denied		All	
	(1)	(2)	(3)	(4)	(5)	(6)
AUS Used	-2.91***		11.25***		4.32***	
	(0.03)		(0.10)		(0.03)	
AUS & Approved		-3.58***		15.04***		4.12***
		(0.03)		(0.11)		(0.03)
AUS & Issue		6.12***		6.81***		6.52***
		(0.05)		(0.12)		(0.05)
Ln(Loan Amount)	2.24***	1.97***	2.53***	2.72***	2.21***	2.15***
	(0.02)	(0.02)	(0.05)	(0.05)	(0.02)	(0.02)
Ln(Income)	0.05**	-0.34***	2.72***	2.48***	2.08***	2.04***
	(0.02)	(0.02)	(0.06)	(0.06)	(0.02)	(0.02)
Number of Observations	5,648,571	5,648,571	871,779	871,779	6,520,350	6,520,350

NOTES: Columns 1, 3 and 5 show estimates of β from: $Time_i = \alpha + \beta AUSused_i + \gamma X_i + \epsilon_i$, where $AUSused_i$ is an indicator equal to one if an automated underwriting system was used for that application and zero otherwise. Columns 2, 4, and 6 show estimates from: $Time_i = \alpha + \beta_1 AUSapproved_i + \beta_2 AUSissue_i + \gamma X_i + \epsilon_i$, where $AUSapproved_i$ is an indicator equal to one if an automated underwriting system was used and the AUS gave an approve decision (this does not imply the loan was ultimately originated). $AUSissue_i$ is an indicator equal to one if an automated underwriting system was used and the AUS did not give an approve decision (this does not imply the loan was ultimately denied). The sample in Columns 1 and 2 is 2018-2019 originations, Columns 3 and 4 show estimates for applications that are ultimately denied, and Columns 5 and 6 show the relationship for all denied applications and originated loans. 10%, 5% and 1% significance levels are denoted by *, ** and ***. Sources: Confidential HMDA.

TABLE A.5

HOUSE PRICE RESPONSE: 1993-1996

	Δ log Price	
	(1)	(2)
Early LP Market Share	1.49***	
	(0.44)	
Early DU Market Share		-0.64
		(0.41)
OLS/2SLS Stage	OLS	OLS
Division FE	Y	Y
N	6,426	6,426

NOTES: This table shows alternative house price responses estimated relative to matched control lender exposure. *Early LP Market Share* (LP_z) is the zip code exposure measure defined in Equation 1. *Early DU Market Share* (DU_z) is an analogous zip code exposure measure computed for early DU users. We include census division by year fixed effects and condition on the following variables interacted with year dummies, at the ZIP code level unless otherwise indicated: (county) coastal indicator, number of lenders, large lender market share, thrift market share, commercial bank market share, share of originations sold to either Fannie or Freddie, share of originations sold to Freddie, log median household income, Saiz (2010) housing supply elasticity. Column 1 conditions on the combined share of early LP users and their matched control lenders. Column 2 conditions on the combined share of early DU users and their matched control lenders. The sample includes ZIP codes in metropolitan counties with FHFA house price data available continuously from 1990. Standard errors are clustered by CBSA. Sources: FHFA HPI; HMDA 1990 decennial census; BEA; NOAA list of coastal counties; and authors' calculations.

TABLE A.6

EFFECT OF THE EARLY LP MARKET SHARE ON HOUSE PRICES 1990-2000

	Δ log FHFA Index	Δ log Median Home Value		
	(1)	(2)	(3)	(4)
Early LP Market Share (1993)	3.82***	4.16***	3.18***	2.37***
	(1.12)	(1.13)	(0.66)	(0.66)
Division FE	X	X	X	X
Number of ZIP5	6,423	6,423	10,313	13,347
Number of Counties	587	587	724	1,224
Number of States	49	49	49	51
Number of Observations	6,423	6,423	10,313	13,347

NOTES: This table explores the effect of broadening the sample of ZIP codes beyond that allowed by the FHFA repeat sales price index. Column 1 shows the price response measured between 1990 and 2000 using our main dependent variable. Column 2 shows the change in log median home value from the Decennial Census over the same period holding the sample of ZIP codes fixed. Column 3 drops the requirement that the FHFA repeat sales price index be non-missing. Column 4 does not condition on housing supply elasticity as this increases the sample further. We condition on the following variables, at the zip code level unless otherwise indicated: (county) coastal indicator, number of lenders, large lender market share, thrift market share, commercial bank market share, share of originations sold to either Fannie or Freddie, share of originations sold to Freddie, log median household income, housing supply elasticity. Standard errors are clustered by CBSA. 10%, 5% and 1% significance levels are denoted by *, ** and ***. Sources: HMDA; FHFA HPI; 1990 decennial census and authors' calculations.

TABLE A.7
COVARIATE BALANCE TABLE

	Treated	Control		Std Diff.	
		Raw	Balanced	Raw	Balanced
Coastal	71.9	56.8	71.8	0.33	0.00
Elasticity	1.7	1.7	1.7	-0.08	-0.00
# HMDA Respondents	103.1	105.4	103.1	-0.05	0.00
Large HMDA Respondent Share	27.5	24.5	27.5	0.32	0.00
Thrift Share	26.0	26.3	26.0	-0.03	0.00
Commercial Bank Share	37.6	39.0	37.6	-0.11	-0.00
Log Median Income	10.5	10.5	10.5	0.20	0.00
Freddie Share	15.8	13.2	15.8	0.40	0.00
Fannie or Freddie Share	37.2	32.7	37.2	0.42	0.00
Early DU or LP Share	11.5	6.4	11.5	0.65	-0.00
New England	10.3	9.9	10.3	0.02	0.00
Middle Atlantic	21.3	12.4	21.3	0.22	0.00
East North Central	24.3	16.1	24.3	0.19	0.00
West North Central	5.5	7.2	5.5	-0.08	-0.00
South Atlantic	15.5	18.7	15.5	-0.09	0.00
East South Central	4.0	3.3	4.0	0.04	0.00
West South Central	2.9	9.9	3.0	-0.42	-0.00
Mountain	11.6	17.4	11.6	-0.18	0.00

NOTES: This table compares covariate means before and after reweighting. The treated group is defined as ZIP codes in the top quartile of the exposure distribution. The control group is all other ZIP codes in the sample. The standardized difference in Columns 4 and 5 is defined as $\frac{\bar{x}_T - \bar{x}_C}{s_{x,T}}$, where $\bar{x}_T$ is the mean in the treated group, $\bar{x}_C$ is the mean in the control group and $s_{x,T}$ is the standard deviation of x in the treated group.