CONSUMER FINANCE OUTCOMES OF BANKING WITH CREDIT UNIONS*

Andrés Shahidinejad†

December 2024

Abstract

This paper studies the economic effects of banking with U.S. credit unions (CUs). Using credit reports and administrative data, causal effects are estimated with a novel instrument for banking with a CU: the distance-weighted density of nearby CUs. Banking with CUs causes borrowers to have fewer delinquencies, higher credit scores, and lower bankruptcy rates. I find support for several mediating mechanisms: CUs charge lower interest rates, price in less risk-sensitive ways, resell fewer mortgages, and are more accommodating of past-due borrowers. Results suggest that CUs behave differently than for-profit banks and are inconsistent with CUs behaving as "for-profits in disguise." JEL codes: L30, G21, G50, H25

*I am grateful to Matthew Notowidigdo, Neale Mahoney, Robert Gertner, and Constantine Yannelis for their advice and comments. I am also thankful to Olivier Darmouni, Benedict Guttman-Kenney, Agustín Hurtado, Scott Nelson, Yingjie Qi, Raghuram Rajan, Jordan van Rijn, Amir Sufi, and seminar participants at Chicago Booth, European Finance Association, FDIC, Federal Reserve, ITAM, NBER Summer Institute, Midwestern Finance Association, Rand Corporation, and University of Wisconsin-Madison for their comments and suggestions. The results in this paper are calculated (or derived) based on credit data provided by TransUnion, a global information solutions company, through a relationship with the Kilts Center for Marketing at the University of Chicago Booth School of Business. TransUnion (the data provider) has the right to review the research before dissemination to ensure it accurately describes TransUnion data, does not disclose confidential information, and does not contain material it deems to be misleading or false regarding TransUnion, TransUnion's partners, affiliates or customer base, or the consumer lending industry.

†Northeastern University, a.shahidinejad@northeastern.edu, Hayden Hall, 370 Huntington Ave., Boston, MA 02215

1 Introduction

In the U.S., nonprofit banks account for a significant share of the supply of consumer finance. Better known as Credit Unions (CUs), their market shares of various consumer lending products have grown since the early 2000s. They hold about 8% of the banking industry's assets under management and more sizeable shares of various lending markets: 26% in personal loans, 13% in mortgage originations, and 28% of auto loans (Experian, 2017). Furthermore, CUs had over 124.3 million memberships, indicating that an even larger fraction of the U.S. adult population are customers of nonprofit banks (National Credit Union Administration, 2020).

Two fundamental institutional features distinguish CUs from for-profit banks. First, they as member-only cooperative organizations. Furthermore, only individuals who belong to a CU's field(s) of membership can become its members, owners, and customers. Fields of membership are defined in terms of an occupation, association, or geography. Second, they are nonprofit organizations that seek to provide better prices and banking services to their customers.

CUs provide a fresh, sizeably important, and understudied setting to research the role of relationships and private information in consumer finance (Petersen and Rajan, 1994; Hauswald and Marquez, 2006), and the costs and benefits of alternative banking models. Their constitution as member-only cooperatives is a potential channel by which individuals with more advantageous and hard-to-measure creditworthiness may be differentially sorting across bank types. Similarly, CUs are a rich setting to explore long-standing questions regarding the nature of nonprofit firms (Hansmann, 1980; Newhouse, 1970; Pauly and Redisch, 1973; Rose-Ackerman, 1996). Unlike with nonprofits in general, CUs have relatively measurable –if putative– objectives and researchers can access high-quality data on them and the banking sector.

This paper contributes to the literatures on banking, credit unions, and nonprofit organizations with three main findings. The first is documenting that the populations of borrowers that originate loans with CUs and competing banks are surprisingly similar on observable dimensions. This is striking considering that the institutional features of CUs are designed around selective membership. In particular, CU membership requirements are often based on occupation. Second, CUs offer lower interest rates and yield better outcomes post-origination than competing for-profit banks. I provide evidence that this is true in both the correlational and the causal sense. Third, CUs approach banking differently: their pricing is less risk-sensitive, they are less likely to resell mortgages on their balance sheet, and they are more accommodating of past-due borrowers. Altogether, these mechanisms contribute to the differential borrower outcomes and appear consistent with the cooperative nature of CUs and inconsistent with the view that CUs are pure for-profits in disguise (Weisbrod, 1975; Hirth, 1999; Duggan, 2000).

My primary data sources are the Home Mortgage Disclosure Act (HMDA) data and an anonimyzed 10% sample of credit records provided by TransUnion, one of the major national credit rating agencies. The HMDA data are a quasi-universe of the mortgages originated in the U.S., and therefore provide a good way to evaluate differences in the borrower segments that CUs and banks focus on. Overall, the distributions of loans both lenders originate are much more alike than they are different. On average, CUs originate mortgages that are slightly smaller and to individuals with slightly lower incomes, but the distributions of loan size and borrower income are strikingly similar. CU and bank loans are distributed across similar geographic measures of creditworthiness, income, and urbanity. The characteristics of the populations of auto loan borrowers at CUs and banks are also strikingly similar. These facts are consistent with two interesting conclusions about differential selection into CUs. First, CU membership requirements do not seem to affect the overall pool of CU mortgage borrowers on observable dimensions. That is, any potential differential selection into CUs is not strong enough to manifest itself in the aggregate data. Second, because CUs seem to be competing with banks for similar customers, increasing access to credit is not a central feature of CU participation in the mortgage market.

To study interest rates, I develop a novel method to infer rates from credit report data on mortgages. Then, using consumer credit data from TransUnion linked with mortgage data from HMDA, I show evidence that individuals with CU-originated mortgages get interest rates that are 18.9 basis points lower on average and experience better credit-profile outcomes up to four years after origination. Three years after origination, bank mortgages are three times as likely as CU mortgages to be 90+ days past due, and they owe three times as much. CU mortgages appear less likely to be charged off or sent into collections than small bank mortgages, but not than big bank mortgages. The better performance by CU-originated loans spills over into the individual's overall credit profile. The credit scores of individuals with a CU mortgage improve by one percentile point more than the credit scores of individuals with bank mortgages. Most notably, the likelihood of having a bankruptcy record is lower for individuals with CU mortgages.

In principle, however, these differences are due to both mutual selection effects on unobservable dimensions and bank treatment post-origination. Credit record data allow me to evaluate whether CUs charge lower interest rates and provide better credit outcomes even after accounting for individual risk, an impediment in the literature and policy debate on CUs thus far. To allay concerns about selection on characteristics that are unobservable in my data, I instrument for whether the individual originates with a CU versus a bank. The instrumental variable (IV) is a distance-weighted density measure of CU branches within 10 kilometers of the address of the new mortgage. The instrument correlates with and is predictive of CU choice. The argument for the instrument's validity relies on the assumption that individuals do not choose the property they buy based on factors correlated with the density of CU branches around it, and that CU density only affects the individual credit outcomes via the choice of mortgaging with a CU. Based on the logic that borrowers choose where to live based on considerations that do not include CU branch locations per se, I argue for and show evidence that lends credence to these assumptions.

Accounting for risk and selection, LATE estimates of the CU effect on interest rates suggest that individuals who borrow from CUs as a result of quasi-exogenously being closer to CU branches get interest rates that are 29.5 basis points than if they had borrowed from a small bank. Under reasonable assumptions, this difference implies cost savings of approximately 0.5% of the mortgage loan's principal amount. For the broader population, I estimate that CUs charge 10.7 basis points less than comparable banks. From a bank's balance sheet perspective, these differences in interest rates are similarly economically meaningful. I apply the same technique of controlling for observable risk and selection to credit outcomes, focusing on the comparison of CUs and small banks. For the overall credit profile outcomes, this specification amounts to an IV-difference-in-differences research design. IV estimates generally confirm the conclusions from the descriptive event studies.

These results also hold in the market for auto loans. By going to a CU, auto loan borrowers pay interest rates that are on average 46 basis points lower, are less likely to become past due and have bankruptcy files, and have better credit scores on average. These differences take into account and control for borrower creditworthiness, loan terms, and geographic factors —including the CU density instrumental variable. The caveat is, however, that I cannot identify the lender beyond whether they are a CU or a bank, and therefore the set of variables I can analyze is more limited.

If CUs offer better interest rates and credit outcomes, why do they not dominate the consumer lending market? First, substantial search frictions sustain price dispersion across consumer lending markets over time (Woodward and Hall, 2012; Argyle et al., 2023; Stango and Zinman, 2016). Second, although loosely enforced, CU membership requirements probably make it seem harder than it is to borrow from a CU, thus limiting the reach of CUs. Third, given their nonprofit, cooperative mission, it is unclear that CUs would aim to dominate the lending market. In the last section of the paper I show evidence for three mechanisms by which CU and bank behavior differ and can lead to these differential prices and outcomes.

The first is that the *slope* of the CU pricing function is less sensitive to variation in the observable risk characteristics of the borrower and the lending environment. CU interest rates show a flatter gradient on credit score. Contrary to banks, CU interest rates are lower in areas with high applicant rejection rates and higher in areas with high creditworthiness. Using an estimated pricing function, I also show that the expected *level* of counterfactual CU interest rates is lower across the entire support in my sample. I interpret these results to mean that CU pricing seems to be less optimized around profit seeking than for-profit bank pricing and is also consistent with a cooperative dimension of CUs.

Second, CUs are drastically less likely to resell mortgages onto the secondary market, consistent with them following a different mortgage banking model. I correlate all lenders' mortgage resell rates to credit outcomes and find that loans from lenders with low resell rates experience better outcomes, although these improvements do not spill over onto the borrower's credit profile. This finding is consistent with the idea that banks that keep mortgages on their balance sheet better internalize incentives to originate good risk *and* improve outcomes post origination. In particular, maintaining the loan on balance sheet creates incentives for the lender to originate good risk and to resolve repayment issues post origination.

Third, among mortgages that become past due, CUs are more likely to accommodate the borrower's circumstance by granting the loan a forbearance, deferment, or declare it affected by a natural disaster. Similarly, CUs are less likely to foreclose, repossess, charge-off or send a loan to collections when it is past due. This is consistent with the causal estimates showing that CUs do something after mortgage origination to yield better borrower credit outcomes. These findings are consistent with CUs yielding better borrower outcomes by being *relatively* accommodating and forgiving when their loans are behind on payments.

Motivated by the finding that CUs were less likely to charge high interest rates prior to 2010, I explore heterogeneity in CU effects across time. I segment the data by whether a mortgage was originated in $2004-2008$, $2009-2011$, and $2012-2017$. The CU effect on credit outcomes was greatest between 2004 and 2008. One interpretation of these results —consistent with a growing body of evidence on poor underwriting standards prior to the Great Recession— is that the quality of bank-originated mortgages was then at its lowest relative to CUs. That is, when demand and ex-ante profit opportunities seemed high, CU prices were at their relative lowest and their quality was its relative highest.

The remainder of the paper proceeds as follows. Section 1.1 reviews the literature on nonprofit organizations, its main application being to the hospital industry, and then the brief literature on CUs. Section 2 provides institutional background on CUs in the U.S. Section 3 describes the main data sources and the results of merging HMDA data with credit records data. Section 4 presents aggregate facts about CUs' market focus and descriptive evidence that CUs offer better interest rates and quality in terms of credit outcomes. Section 5 presents the causal research design and its results. Section 6 presents evidence on mechanisms and heterogeneity behind the results. Section 7 concludes.

1.1 Literature Review

Cororaton (2019) and van Rijn et al. (2019) provide evidence that CUs have different objective functions than banks. Cororaton (2019) finds CUs are able to grow lending by forgoing profits. van Rijn et al. (2019) argue that, in alignment with their mission, CUs offer their executives less powerful incentives to pursue profits. Consistent with my empirical results in the mortgage setting, van Rijn et al. (2021) use the Survey of Consumer Finances to show CUs charge lower auto loan rates. The literature has taken many approaches to modeling the specific nonprofit form of CUs. Smith (1981) models CUs as size maximizers, Flannery (1974) and Leggett and Stewart (1999) model CUs as balancing the interests of savers and borrowers, and Cororaton (2019) and van Rijn and Li (2019) model CUs as maximizing a combination of returns and member utility.

Because CUs are commercial nonprofits, I focus on the existence and behavior of *commercial* nonprofit firms in this review (donation-based nonprofits generally do not compete with for-profits

(Weisbrod, 1975) and I consider to be outside the scope of this review). Malani et al. (2003) categorize theories about commercial nonprofits according to the key substantive assumptions that explain their existence: altruism, cooperatives, and noncontractible quality. Since Newhouse (1970), a common way to specify altruism has been to model the entrepreneur as maximizing a combination of product quantity and quality or both altruism and profits/consumption (Lakdawalla and Philipson, 2006). Pauly and Redisch (1973) model the nonprofit as a cooperative effort by physicians to eliminate an outside investor's residual claims on hospital revenue. The cooperative form is then a way for key employees to gain control of the organization. Hansmann (1980) argues nonprofits exist to solve a market failure arising from noncontractible quality. When customers cannot observe the quality of a product, the supplier has opportunities to profiteer. But because customers anticipate this incentive, they are less likely to purchase, and the market for the product potentially unravels. The nonprofit form, defined by the "non-distribution constraint," curbs the incentive to profiteer and signals to the customer that the firm can be trusted not to shirk on quality.

Due to the empirical importance of nonprofits within the sector, U.S. hospitals have been a focal setting for the study of nonprofits (Urban Institute, 2020). The evidence on whether nonprofit hospitals provide better health outcomes is mixed (Malani et al., 2003). Duggan (2000) studies a policy change in California during the 1990s that created opportunities for hospitals to profit by treating indigent patients. Both for-profit and private nonprofit hospitals were equally likely to pursue financial incentives without improving quality of care for the poor. More generally, the literature seems to support the view that nonprofit hospitals do not provide better prices or quality for patients than for-profit hospitals. A more recent strand in the literature argues theoretically (Philipson and Posner, 2009) and empirically (Capps et al., 2020) that nonprofit hospitals should not receive leniency in antitrust scrutiny. However, recent evidence in Garthwaite et al. (2018) finds nonprofit hospitals disproportionately provide the bulk of uncompensated care to the uninsured, which is generally covered by forgoing profits.

2 Institutional Background

Five characteristics distinguish CUs from other financial institutions (U.S. Treasury, 2001). First, CUs are member-owned cooperatives and each member is entitled to one vote in electing the board of directors. Second, CUs do not issue capital stock; instead they create capital via retained earnings. Third, CUs rely almost exclusively on volunteer boards of directors whom the members elect from their ranks. Fourth, CUs are nonprofit instead of shareholder-owned institutions. That is, earnings can be retained as capital or returned to members in the form of interest on share accounts, lower interest on loans, or via other products and services (e.g., financial education or micro-loans). Finally, CUs may only accept as members individuals identified in the CU's articulated field of membership. According to the National Credit Union Association (NCUA), the CU regulator and provider of deposit insurance, CU membership is limited to individuals who share a common bond of occupation, association, or community (i.e., a geography).

Within these defining features, CUs are a heterogeneous class of banks. The table below provides one example of a CU for each field of membership category. CUs are very heterogeneous in asset size, as they range from US $149 billion in assets under management to one with a balance sheet of US $20,000, resembling a household in size. As membership rules have been relaxed, many CUs now have dual fields of membership, and fields of membership are stretched to include the family members of those directly included in the field. Other CUs have converted to geographic-based fields, presumably when doing so is conducive to membership growth. As the example below shows, community-based CUs can be liberal with respect to what it means to be located in a geography.

Examples of Credit Unions

Name	Category	Assets Under Management	Membership	Field of Membership
Navy Federal	Occupational	US $149 billion	10 million	Armed forces & their families
Adirondack Regional	Community	US $50 million	7,000	Individuals who live, work, own a business, worship, or attend school in Clinton, Essex, Franklin, or St. Lawrence County of NY state, & their families
Holy Trinity Baptist	Associational	US $20 thousand	100	Church members

The NCUA states that the "Federal Credit Union Act expects this national system [of Credit Unions] to meet 'the credit and savings needs of consumers, especially persons of modest means.'" Thus, credit unions have an explicit social objective in their mandate, which is not the case for for-profit banks. Historically, some have interpreted the rationale that CUs would receive special tax treatment on the basis that they would serve a restricted consumer base, would focus on low- to middle-income customers, and would offer a small set of products (Internal Revenue Service, 1979). For-profit banks complain that the tax exemption is an unfair competitive advantage that allows CUs to charge below-market interest rates on loans and offer above-market deposit rates (American Bankers Association, 2021). They argue that, after taking into account size and portfolio concentration, CUs are not substantively different from commercial banks. Because this argument is generally made on behalf of community banks or similarly-sized banks, I principally compare CUs with "small" banks, which I define as banks with assets in the same range of assets as CUs.

Since the early 2000s, the National Credit Union Administration has relaxed the rules on how fields of membership can be defined. It has also simplified the process by which CUs can change their field of membership, hold multiple fields at once, and convert to geographic-based fields of membership. In the last two decades, close to one thousand CUs converted to a community-based charter; from 1997 to 2017, community charters grew from 6.5% to 30.3% of all federally charted CUs (van Rijn, 2018).

3 Data

3.1 Credit Records Data

TransUnion is one of the three nationwide consumer credit-reporting agencies in the U.S. The data analyzed in this paper comes from a 10% sample of individual monthly records gathered by TransUnion from 2001 to 2020. If an individual is observed in the sample, they are in the sample for as many months as TransUnion observes them. TransUnion gathers data on an individual's credit lines from lenders, servicers, and public sources to construct individual credit reports for approximately 90% of the U.S. adult population (Consumer Financial Protection Bureau, 2016). Thus, credit records report the evolution of balances and repayment behavior for each line of credit. In this paper I focus on whether the loan is 90+ days past due, the past due balance, and whether the loan is charged-off or sent into collections. Credit records also report aggregated data about the borrower across credit lines. Here, I focus on the number of credit lines 90+ days past due, the total past due balance, credit scores, and bankruptcy records.

Whether payments are past due, and by how much, captures the most immediate negative outcome after origination. A more severely negative outcome is the mortgage being charged off and/or being sent into collections. Other variables aggregate across an individual's credit lines and reveal whether loan outcomes spill onto the overall credit profile. The number of trades and total amount past due are analogous to the mortgage-specific outcomes. The credit score represents an individual's creditworthiness and is an important factor that lenders use in deciding the quantity and price of credit they will offer. Last, I study whether the adverse mortgage outcomes translate into increased public bankruptcy records. Additionally, although credit records data do not contain information on prices, I can infer and analyze the interest rate on a sizeable fraction of mortgages.

I also use a sample of auto loans derived from TransUnion data to investigate whether the main results from the mortgages carry over to the auto loans. I study the same set of outcome variables and can control for individual, loan, and geographic characteristics of the loan. However, because there is no equivalent to the HMDA data for auto loans, I cannot obtain the specific identity of the lenders; I can only identify whether the lender is a CU or a bank.

3.2 Home Mortgage Disclosure Act (HMDA) Data

Since 1975, the HMDA has required many lenders to report application-level mortgage data. HMDA data cover approximately 90% of mortgages originated in the U.S. (Board of Governors, 2017). Supplementing the credit records data with HMDA data greatly enriches the information I observe.

Although the reporting requirements have changed over time, I use the following variables, which are available in my sample time period: the loan amount, applicant's income, loan type (conventional, FHA, VA, or FSARHS), loan purpose (home purchase, improvement, or refinancing), lien status (first, subordinate, or non-secured), property type (one-to-four family, manufactured, or multifamily), whether the property will be owner occupied, the census tract of the property, the identity and principal regulator of the lender, the lender's assets under management, whether the application was approved, and whether the loan was resold in the secondary market in the same calendar year. HMDA data allow me to construct two measures of credit health at the level of geography: the fraction of mortgage applications that were accepted and the mean credit score at the tract-year level. I can also construct similar measures at the lender level: lender-year mortgage application acceptance rates and secondary market resell rates. Including HMDA data also improves the identification of the originating lender in cases where the loan is not being serviced by the originating lender. Although HDMA data include other variables (especially after 2017), I restrict myself to these variables because they are consistently available going back until 2004.

3.3 Merge of Credit Records and HMDA Data

I combine the mortgage credit records and HMDA data with a "fuzzy" merge based on three variables: the year of origination, the census tract, and the loan amount rounded to the nearest thousand. The merge requires equality of the three variables and uniqueness in both datasets. The requirement that records be unique combinations of the values of these three variable cells in each dataset provides a high degree of confidence that the resulting matches are accurate, because the HMDA data amounts to a near universe of mortgages. Furthermore, the requirement that mortgages be unique combinations of tract-year-amount excludes many observations from the merge, which means no attempt is made to find their corresponding credit records.

Appendix Figure A1 shows the fraction of originated mortgages in HMDA that I am able to match. Because the credit records data are a 10% sample, this is the maximum match rate. On a yearly basis, the match rate oscillates between 1% and 7% (between 10-70% of the maximum) and is notably higher since 2009 (this coincides with data improvements TransUnion introduced on that same year). The most obvious reason my match rate is not closer to 10% is the requirement that loans be unique combinations of amount, tract, and origination year. Because I would not be able to have a confident match on non-unique loans, I do not attempt to merge them.

The dataset resulting from the fuzzy merge contains approximately 5.1 million mortgages originated between 2004 and 2017. After restricting to loans which a) are conventional, b) are for one-to-four family properties, c) are owner occupied, and d) have standard terms of 10, 15, 20, or 30 years, the sample contains approximately 3.4 million loans. I further restrict the data to bank originated loans (i.e., mortgage company loans are excluded) and the final sample includes approximately 2.3 million mortgages.

Despite the clear limitations of a fuzzy merge, the merge quality is good. Appendix Figure A2 shows the fraction of CU-originated mortgages in the HMDA and fuzzy-merged sample are similar across all years. The first part of Table 1 compares the mean value of variables in my matched sample and in the HMDA universe. I show these variables by CU, small bank and large bank. The mean loan amount and applicant income are larger in the matched sample than in HMDA. That this difference is present for all lender types means my sample is skewed toward larger loans, but not in a differential way across lender types. Loan resell rates are also slightly higher in the matched sample for CUs and small banks. Overall, though, the remaining variables suggest my matched sample is sufficiently representative of the HMDA universe.

3.4 Other Data Sources

I primarily use FDIC and NCUA Call Reports and Summary of Deposits for banks' asset size and branch locations. To complement branch-location data of CUs, I use the geocodes provided in the "Your-economy Time Series" by the Business Dynamics Research Consortium. I use S&P Ratewatch data to compile descriptive statistics on CU and bank prices. Ratewatch data are branch-level surveys of interest rates and fees for various deposit and loan products. Finally, I use the 2010 Census tract data on the urban versus rural population and the 2012 five-year American Community Survey to gather data on a tract's median income.

4 Descriptive Evidence

This section documents various facts related to the selection of borrower types across CUs and banks, differences in CUs' approach to mortgage financing, and mortgage credit outcomes. I primarily contrast CUs to "small" banks, to contrast with a group of comparable in assets. Specifically, I define small banks as those with less than 105% the assets of the largest CU in each year.

4.1 Borrower Selection into CUs

Using the HMDA universe of data, Figure 1 plots the distributions of many features of borrowers and mortgages originated between 2004 and 2017. I separately plot the distribution for each bank type. For bank and geographic features, distributions are weighted by the number of mortgages.

Panels (a) and (b) show that the populations of CU and bank borrowers originate similarly sized loans and have similar income. Although CUs offer a mass of some smaller loans, Table 1 shows this is at least in part because a larger fraction of CU mortgages are for home improvement. CUs seem to more evenly distribute their originations across purposes (home purchase, improvement, or refinance) and are more likely to originate mortgages without a first lien on the property.

Panels (c)–(f) summarize characteristics of the mortgages based on geographic measures of creditworthiness, income, and urbanity. These geographic measures also show that the populations of CU and bank borrowers are very similar. Although their similarity is the main fact, it is worthwhile noting some of the relatively minor differences between them. Relative to large banks, CUs' geographic focus is slightly less urban (86% vs. 81%) and slightly lower income (US $73K vs. US $67K). Relative to small banks, however, CUs' geographic focus is slightly more urban (78% vs. 81%) but of an equal income level (US $67K vs. US $67K). To benchmark these means differences, note 80.7% of the U.S. population lives in an urban area, and tracts with median incomes of US $73K and US $67K are at the 79th and 73rd percentile of tracts, respectively. To emphasize the role of CUs and small banks at different levels of these distributions, Appendix Figure A3 displays the fraction of CU and small bank mortgages by the level of the distribution. Panels (e) and (f) show CUs disproportionately supply loans in tracts with high mortgage rejection rates and low credit scores. At the same time, panel (g) shows CUs provide relatively few of the loans in tracts with the lowest income. Last, Table 1 shows that, for the matched sample, the distributions of credit scores are remarkably similar across banks and CUs.

To summarize, the main economic finding is that the populations being served by CUs and banks are much more alike than they are different. Overall, CUs and banks, especially small ones, seem to be present in similar geographies, and their borrower focus is, on the whole, quite similar. The significant degree of overlap in CU and bank borrower profiles shows that CUs do not differentially select borrowers on these observable dimensions. If anything, these observable characteristics of borrowers suggest that CUs lend to individuals with slightly lower income on average, and therefore do not support the idea that CUs select an advantageous risk pool of borrowers.

Although this evidence is not conclusive, it shows that CUs largely do not extend access to credit along the extensive margins of new markets or segments of consumers. Yet, because CUs restrict their members to fields of membership, they may have an informational advantage on their members relative to banks and their borrowers, and that they may be able to extend access along the intensive margin. That is, conditional on income, credit score, and on receiving a loan, CUs may be able to offer higher loan sizes. However, as shown in Appendix Figure A4 CU and small bank conditional loans sizes are no different. I interpret this as inconsistent with CUs increasing access to credit on the intensive margin or with advantageous selection by CUs.

Whether extending credit to those who do not have access is the *raison d'etre* of CUs is unclear from the law and policy discourse. From a conceptual economic perspective, however, increasing access to credit would perhaps be the largest way in which nonprofit banks could be creating value. I offer three non-exclusive ways to interpret the results, concluding that CUs do not extend access to new credit in mortgage markets. One is that for-profit banks have the capacity to reach all market segments, and so no space is left for nonprofits to "open" new markets or segments. The active discourse on banking deserts implies this explanation is unlikely (Federal Reserve Bank of St. Louis, 2017). Another possibility is that CUs' mission is not centered on extending access, but rather on serving their current fields of membership, which may not include individuals without access to credit. Last, although HMDA cover approximately 90% of mortgages, a significant fraction of small CUs are excluded from the sample I analyze because of HMDA's reporting requirements. It is therefore possible that borrowers at the bottom of the income and credit distribution are disproportionately served by CUs that do not report to HMDA. Although the criteria are complex, usually only lenders with at least $28 million in assets under management report to HMDA. Because approximately one third of CUs do not meet the reporting requirements, CUs are disproportionately excluded from the HMDA data and they do not reveal the full extent of their activities.

4.2 Banking Models

Figure 2 shows CUs follow different mortgage banking models. Although CUs and banks have similar mean rejection rates, the mortgages originated by CUs show more variation among CUs in their likelihood of rejecting applicants —especially relative to large banks. Mortgages originated by CUs are drastically less likely to be resold in the secondary market and therefore more likely to be kept on the CUs balance sheet. This finding is interesting because of the argument that the secondary market incentivized loose underwriting standards, especially prior to the financial crisis of 2008 (Acharya et al., 2011). Section 6.4 empirically explores this idea further.

4.3 Credit Outcomes

Using the matched sample of mortgages, Figure 3 plots the evolution of four of the credit outcomes described in Section 3.1 by bank type. Panels (a) and (b) show the means for mortgage-specific outcomes every six months up until four years after origination. Panels (c) and (d) show the mean of quarterly credit profile outcomes, four years before and after origination.

Focusing on outcomes at the three-year mark, CU mortgages have a 0.7% chance of being 90+ days past due, whereas small bank mortgages have a 1.4% chance and large bank mortgages have a 1.6% chance of the same event. Similarly, the mean amount past due for individuals with CU mortgages is US $118, whereas the average amount past due for those with small and large bank mortgages is US $297 and US $356, respectively.[1] Panel (c) shows mean individual credit scores of individuals with a CU mortgage grow by more than individuals with bank mortgages. At three years, CU credit scores are 1.5 percentile points higher than small bank credit scores and 0.75 percentile points higher than large banks scores. This contrasts with three months prior to origination, when large bank scores were 0.3 percentile points greater than CUs and small bank scores were 0.6 percentile points lower. Overall, CU mean credit scores grew about 1 percentile point more than mean bank scores. Panel (d) shows the mean number of individual public bankruptcy records. At three years, 1 in 40 individuals in the CU group will have one bankruptcy record, whereas approximately 1 in 34 individuals in both bank groups will have one record.

Prior to origination, CU borrowers' credit scores and bankruptcy records are improving by more than bank borrowers'. This suggests caution in attributing these differences to CU behavior post origination and that there may be differential selection on unobservable borrower characteristics.

5 Causal Evidence

This section builds on the descriptive evidence in the previous section by controlling for individual risk characteristics observable in their credit profile and by adjusting for remaining endogenous sorting into CUs versus banks. Pretrends in Figure 3 panels (c) and (d) indicate the possibility of advantageous selection into CUs. The process of searching for a mortgage immediately prior to origination exposes individuals and lenders to information that might induce endogenous sorting across banks. Furthermore, fields of membership technically mean that not all individuals can join every CU. As factual matter, however, CU membership requirements are loose enough that any individual can likely find *a* CU that they can join —even if they cannot join every CU. Treating

[1] Appendix Figure A5 shows that the likelihood that any mortgage is charged-off or sent into collections at the three-year mark is roughly 0.2%. Small bank mortgages have a slightly higher probability of being charged-off or sent into collections, than both CUs and large banks, which show no difference between them. Appendix Figure A5 also shows the mean individual number of trades 90+ days past due and total amount past due, show the mortgage-level differences do spill over onto the credit profile.

CUs as a class of banks therefore ameliorates some of these concerns.

This identification challenge motivates a research design based on an IV that isolates quasi-random variation in the type of bank that individuals get a mortgage from. Section 5.1 describes the construction of the instrument. Section 5.2 presents the results from applying the instrument to the estimate of the "CU effect" on interest rates. Section 5.3 combines the event studies with the instrument to tease out the "CU treatment effect" on credit outcomes. For mortgage-specific outcomes, this approach amounts to a standard IV research design, but for credit profile outcomes the research design is an application of an IV to difference-in-differences (IV-DiD).

5.1 Instrumental Variable: CU Branch Density

Using geocoded bank and CU branch data, I construct the instrument which I refer to as "CU Density." For each mortgage i in my matched sample, I restrict attention to the CU and bank branches within a 10-kilometer radius of each address for which the mortgage was originated.[2] I compute the distance $d(i,l)$ of each branch of lender l to the individual's new address. The instrument is then the density of CU branches over all CU and bank branches, weighted by the inverse of their distance from the new address:

$$CuDensity_i = \frac{\sum_{l \in CU} \frac{1}{d(i,l)}}{\sum_{l \in CU} \frac{1}{d(i,l)} + \sum_{l \in Bank} \frac{1}{d(i,l)}}.$$

The validity of the instrument relies on the standard arguments of relevance and monotonicity, and exogeneity and independence. As a measure of robustness to the arbitrary nature of some of the choices in constructing the instrument, I construct various alternative CU branch density instruments. Each instrument differs in the specification of geographic catchment and distance weighting. I construct a total of nine instruments based on three different geographic catchment definitions —10-kilometer radii, nearest 20 branches, and either— and three different distance

[2] For the small set of mortgages whose address has no branches within a 10-kilometer radius, I use the 20 nearest branches to it within a 200-kilometer radius.

weighting measures —inverse distance, the exponential of negative distance multiplied by 0.1, and uniform weights regardless of distance.[3] Appendix Figure A6 plots the correlations between these nine instruments and shows all of them are highly correlated.

5.1.1 Relevance and Monotonicity

The relevance assumption formalizes the idea that the *CuDensity* instrument is positively associated with the probability that an individual chooses to bank with a CU. The monotonicity assumption formalizes the requirement that no individual chooses against banking with a CU as a result of having a higher *CuDensity* value. Let CU be an indicator function for whether an individual chooses to bank with a CU, and let X be a vector of individual controls. Both these assumptions can be formalized as follows:

$$CU\left(z', X\right) \geq CU\left(z, X\right) \;\; \forall z, z' \in CuDensity \text{ s.t. } z' > z.$$

Figure 4 plots the fraction of individuals who choose a CU by bin of *CuDensity* without conditioning on X. For the instrument values of the vast majority of the population, the fraction choosing a CU is monotonically increasing in the instrument. A notable exception is that the CU fraction of those with *CuDensity* values below 0.01 is lower than of those with values between 0.01 and 0.04. Overall, Figure 4 show the instrument is strongly associated with CU choice, consistent with the assumption above. The exceptions identified above are handled by controlling for a vector of controls X, which include various geographic measures. As reported later in the paper, the F-statistic of the first-stage regression of CU on instrument and controls is well above 8,000 in the sample comparing CUs with small banks and over 500 when comparing with large banks.[4]

I rationalize the relevance of the instrument through a model of bounded search for lenders

[3] The nearest 20 branches are restricted to branches within 200 kilometers. All instruments that restrict attention to 10 kilometers supplement the few properties with no branches within 10 kilometers using the 20 nearest branches.

[4] As another measure of the instrument's relevance, Appendix Figure A7 plots the histogram of the instrument value for the CU, small bank, and large bank mortgages in the matched sample. The CU distribution of *CuDensity* first-order stochastically dominates that of both other bank types.

within a proximate geography. Individuals sample quotes from a fixed number of lenders in their vicinity and then choose a lender from among those they quoted. The intuition is straightforward: if an individual samples a limited number of quotes from banks or CUs around them, the fraction of CU quotes they will sample will be a function of the fraction of total CUs around them, and thus, so will their choice. Appendix Section C describes the model in further detail.

Using the terminology of Angrist et al. (1996), the monotonicity assumption implies no instrument "defiers." That is, no individuals are less inclined to use a CU because they live in an area with higher CU density. Effectively this assumption means the data contain three types of individuals: (a) those who, regardless of CU density, always use a bank (never takers); (b) and those who, likewise, always use a CU (always takers); and (c) those who use a bank when in a low CU density area and use a CU when in a high CU density area (compliers). Assuming that heterogeneous treatment effects of banking with a CU, then the recovered IV estimates will reflect marginal treatment effects for those who are compliers on the relevant margin of $CuDensity$.

Although the individual compliers are not identifiable in the data, Appendix Table A1 describes some characteristics of them as a group. For ease of interpretation, I discretize the continuous instrument into a binary variable based on its median value and follow the method in Angrist and Pischke (2009). Overall, approximately 10% of the sample are compliers and 20% of the treated are compliers. For each control variable, the table shows the percentage of compliers among the observations with a "high" value of the variable. For binary variables, a high value is a 1, and for continuous variables, a high value is one that is above the median. With the exception of bank loan resell rates, compliers are never less than half or more than twice the fraction of overall compliers in the high category of a variable. Compliers are less likely to have higher loans, income, and resold loans; they are less likely to come from areas with high credit scores and income; and they are less likely to go to CUs with high loan resell rates and assets. On the whole, compliers are roughly balanced across the credit score distribution.

5.1.2 Independence and Exclusion

The independence and exclusion assumptions formalize the idea that, after controlling for X, $CuDensity$ is as good as randomly assigned and that it only affects outcomes Y by way of its influence on CU:

$$(CuDensity \perp\!\!\!\perp Y, CU(z) \ |X) \quad \forall z \in CuDensity.$$

Distance-based instruments do not generally satisfy the independence and exclusion restrictions unconditionally, and thus, many of their applications are conditional on geographic controls (Mountjoy, 2021; Card et al., 2020). Because CUs are more likely than small banks to be located in urban areas, controlling for factors related to geography and urbanization seems *a priori* important.

Although directly testing the instrument's conditional exogeneity is not possible, I present two analyses which support the assumption. The first is based on the logic of coefficient and R-squared movement when including additional controls ((Altonji et al., 2005; Oster, 2015) to test for omitted variable bias. The second analysis is similar in its underlying logic, and compares the predictiveness of controls to the instrument versus CU choice. The intuition of these analyses is that there should be little correlation between the instrument and the controls we do not expect to have a relationship. Then, to the extent that these observable controls are indicative of the relationship to unobservables, this helps support the assumption that the instrument only affects outcomes via CU.

The set of controls I use includes the following: the logarithm of the bank's total assets; a bank-year's mortgage resell and rejection rates; the census tract-year mortgage application rejection rate, its mean credit score, its fraction of urban residents, and its median income; a dummy for whether the individual mortgage was resold; dummies for the loan's purpose, lien status, and terms; the logarithm of the loan amount; dummies for whether the loan was originated before 2009, after 2011, or in between; dummies for whether the applicant was under 35, over 60, or in between; and dummies for whether the applicant's credit score prior to origination was in the top quartile, the third quartile, or below the median.

Appendix Table A2 reports the results of various "first-stage" regressions of CU choice on $CuDensity$. In the first column, with no controls, the intercept shows 17% of individuals choose a CU when $CuDensity = 0$, and changing $CuDensity$ from 0 to 1 would increase this percentage by 45 points. To give a more relevant interpretation of the magnitude, moving across the interquartile range of $CuDensity$, from .075 to .229, changes predicted CU choice from 20% to 27%. This is consistent with the evidence presented in the previous section that the relevance assumption holds. Columns (2) and (3) show the loan- and tract-characteristic controls help purge the instrument of potentially confounding relationships, as anticipated. Relative to column (3), however, columns (4) through (8) show that even as the coefficient on the instrument is relatively stable from (0.36 to 0.33), the R-squared increases when additional controls are included (from 0.12 to 0.18). Then assuming we can draw information from the observed controls added in columns (4) through (8) about the effect that adding unobserved controls would have, this analysis helps allay concerns about the validity of the independence and exclusion assumption.

Appendix Table A3 compares the predictiveness of the controls on CU choice relative to the instrument. The first column reports the results of an OLS regression of $CuDensity$ on X and the second regresses CU on X. This analysis is helpful because the second regression serves as a benchmark for the degree of correlation between the controls and the instrument. As shown by the R-squared of each regression, the controls are less predictive of the instrument than they are of CU choice. If this was the opposite, it would be a cause for concern about the instrument's validity.

A few other facts are consistent with the instrument's exogeneity. First, the results from Section 4.1 do not evidence different location choices by CUs and banks based on credit-related geographic variables. This helps allay concerns that CUs choose branch locations based on different criteria than banks. Second, the instrument varies at the individual level and so can capture individual variation in otherwise similar broader geographies. Third, although individuals sometimes consult with a lender before choosing which home to buy, it seems reasonable that the decision of which lender to go to is conditional on them extending loans in the area they are interested in purchasing a house in (or already live in), and not vice versa.

5.2 The CU Effect on Interest Rates

In this section I analyze whether CUs offer lower interest rates than competing banks. Previous research has shown descriptive evidence that CUs charge lower interest rates than competing banks do. Appendix Figure A8 uses survey data from S&P RateWatch and plots (a) the mean bank interest rates of various savings and loan products from 2001 to 2019 and (b) the interest rate differential between CUs and banks for each product. These data show that CUs have generally offered better list prices than for-profit banks for most of the last two decades. Appendix Figure A9 shows that CUs were less likely to charge high interest rates prior to the Great Recession than both types of banks. Although these data are suggestive evidence, there is a limit on what one can infer from list-price data because consumer financial products ultimately involve personalized pricing. Thus far, academic and policy analyses comparing bank and CU interest rates have not been able to go beyond correlational observations such as this one, because they are unable to control for individual risk characteristics and selection. The credit data along with the IV identification strategy allow me to causally explore whether a "CU effect" on interest rates exists. This section first explains how to infer interest rates from credit records data and then analyzes the CU effect on interest rates.

5.2.1 Backing Out Implied Interest Rates from Credit Records

Although credit data do not directly contain interest rates, they can be inferred. For a loan l that is observed over multiple months t, observing its term T_l and remaining balance $Balance_{lt}$ is sufficient to back out the interest rate $IntRate_l$ and monthly payment amount $Payment_l$.[5] The annuity present value equation of any loan l at time t links the two observed variables ($Balance_{lt}$, $T_l - t$) to the two unknown variables ($IntRate_l$, $Payment_l$). For any given month of data on its own, this equation is under-determined. But with data on two different months of loan records, say t and t', we can identify the two unknowns by setting up a system of two equations with two unknowns:

$$Balance_{lt}(1 + IntRate_l)^{T_l - t} IntRate_l = Payment_l((1 + IntRate_l)^{T_l - t} - 1)$$
$$Balance_{lt'}(1 + IntRate_l)^{T_l - t'} IntRate_l = Payment_l((1 + IntRate_l)^{T_l - t'} - 1)$$

I solve numerically for the pair $(Payment_l, IntRate_l)$ that satisfies the system above.[6] Two months of data are technically sufficient to identify both the interest rate and payment amount. Empirically, however, any one pair of observations may lead to an incorrect estimate for many reasons, including the following (a) recording errors, such as delays, in the variables; (b) missed payments; and (c) whether the interest rate is variable or fixed is unobserved. Having more than two months of observations is therefore useful because it allows me to estimate a single loan's interest rate multiple times, once for each pair of months, and apply quality controls to the estimates. Appendix B further details the algorithm that makes use of more than two months of data to ensure the accuracy of the inferred interest rates.

Given the quality restrictions imposed, the algorithm produces an estimate of the implied interest rate for 67% of the mortgages in the matched sample. Appendix Table A4 reports the results from a regression of an indicator variable for whether I am able to calculate an interest rate on the main set of controls X. I expect the main reason for my inability to calculate an interest

[5] The monthly payment amount reported to credit agencies usually includes the sum of payments for mortgage interest and principal, taxes, and other fees collected by the loan servicer (e.g., home owner association fees). Therefore, it cannot be used to infer interest rates directly.

[6] One can also solve for the interest rate without solving for the payment amount by using a root-solver on the following equation, which is be obtained by dividing the first and second equations above:

$$\frac{Balance_{lt}}{Balance_{lt'}} = \frac{(1 + IntRate_l)^{T_l - t} - 1}{(1 + IntRate_l)^{T_l - t} - (1 + IntRate_l)^{t' - t}}$$

rate to be due to some irregularity in the timely reporting of information to the credit bureau, which can lead to incorrectly inferred interest rates.

These results are informative of the external validity of the exercise of relating interest rates to CU versus bank prices. They speak to the question of how does the 67% of observations I am able to calculate a rate for differ from the total sample. Appendix Table A4 shows I am more likely to be able to calculate rates for mortgages with a CU, from tracts with higher credit scores, that were resold, for a home improvement, with a higher loan amount, from borrowers above 60 and below 35. The remaining variables are negatively correlated with having a calculated interest rate.

Figure 5 compares the monthly mean of calculated interest rates with the market-wide monthly average of 30-year fixed-rate mortgages. This comparison validates the approach to extracting interest rates, because nothing in the algorithm would mechanically make interest rates align over time. The fact that the plot shows a high degree of temporal correlation is very encouraging because it shows that, on average, this method accurately infers interest rates.

5.2.2 Interest Rates at CUs and Banks

First, I run an OLS regression of interest rates on a dummy variable CU, whether the originator is a CU. Second, I run the same OLS regression but with controls for individual risk characteristics and bank controls. Third, I run a two-stage-least-squares IV regression in which I instrument for CU with $CuDensity$, while maintaining the aforementioned controls. The equations below summarize the approach, where equations (3) and (4) are the second and first stages of the 2SLS regression. In all cases, I am interested in the coefficient β_1 and in interpreting it as the causal effect of CUs on interest rates, relative to banks.

$$InterestRate_i = \beta_0 + \beta_1 CU_i + \epsilon \tag{1}$$

$$InterestRate_i = \beta_0 + \beta_1 CU_i + \boldsymbol{\alpha} X_i + \epsilon \tag{2}$$

$$InterestRate_i = \beta_0 + \beta_1 \widehat{CU_i} + \boldsymbol{\alpha} X_i + \epsilon \tag{3}$$

$$CU_i = \gamma_0 + \gamma_1 CuDensity_i + \boldsymbol{\delta} X_i + \varepsilon \tag{4}$$

Table 2 shows the estimates for β_1 in all three regressions. The first panel shows results relative to small banks and the second relative to large banks. The unconditional difference between CUs and small banks in the matched sample is -18.9 basis points, controlling for observable risk characteristics changes the estimate to -10.7 basis points, and adjusting for risk and selection via the instrument changes the estimate to -29.5. These estimates are statistically significant, and Appendix Table A5 shows the IV estimates are robust to how I construct the instrument.

The change in coefficients from column (1) to column (2) in Table 2 indicates the importance of controlling for characteristics that are in a credit report and observable to the lender at origination. Column (1) shows that CU borrowers get lower interest rates, but column (2) also shows that some of this difference is due to the different loan and borrower characteristics between the two groups of borrowers. Column (3) shows that the local average treatment effect (LATE) estimate of CUs on interest rates for compliers is about three times larger in magnitude than the OLS estimate in column (2). Compliers are borrowers who "respond" to the instrument in the sense that they are more likely to get a loan from a CU because they live around a higher density of CUs. It stands to reason that, compared to the population of borrowers, compliers face more constraints in searching for mortgages. This interpretation rationalizes the larger magnitude of the LATE estimate compared to the OLS estimate, as borrowers who search less are those that stand to gain the most from quasi-exogeneoulsy being located closer to lower interest rate lenders.

To help interpret the magnitude of a 0.295 interest rate difference from the borrower perspective, consider the example of two 30-year mortgages held for seven years, one with an interest rate of 5% versus another with a rate of 4.705%. As a percentage of the principal, the implied present value cost of the second mortgage is 0.60% lower under a 20% personal discount rate (Warner and Pleeter, 2001). For an average mortgage loan of US $200,000, the present value of savings amounts to approximately US $1,190. From the perspective of a banks' balance sheet, this difference is a significant reduction in interest revenue. The estimates relative to large banks are qualitatively similar; although the IV estimates an effect of a larger magnitude, but with weak statistical significance. Relative to the correlational evidence, this analysis suggests that, conditional on credit score and selection, CUs charge meaningfully lower interest rates for an important segment of borrowers.

I estimate counterfactual interest rates at CUs and banks for all individuals, and compare the predicted interest rates individuals would get under the two types of lenders after controlling for all covariates. Appendix Figure A10 shows a scatter plot of the predicted interest rates for all individuals at banks versus a CU. The horizontal axis plots predicted rates for individuals, which are counterfactual for CU borrowers. The vertical axis plots the difference between predicted rates at a CU and at a bank, counterfactual for individuals who borrowed from the "other" lender. The black solid line plots the OLS fitted line, and the hollow circles show the mean value of the difference for the bin. The plot shows the mean predicted CU rate is lower across the entire empirical support of predicted rates, and the CU difference is larger for higher interest rates. Overall, CUs vary interest rates based on risk factors as banks also do, but they seem to do so at a lower degree of sensitivity. These results are consistent with CUs being less profit maximizing.

A limitation of my analysis of prices is that it only includes interest rates and does not account for origination fees. Appendix Figure A11 shows CUs and banks charge roughly the same amount in mortgage origination fees, which is true both across time and interest-rate levels. Although these are derived from list-price data, this allays some of the concern that CUs may use more origination fees to compensate for lower interest rates.

5.3 The CU Effect on Credit Outcomes

This section builds on Section 4.3, which showed correlational evidence that individuals who originated mortgages with a CU experienced better credit outcomes up to four years after origination. First, by controlling for observable risk characteristics with an OLS regression. Second, by accounting for potential endogenous selection with the instrument. In this section, I combine these controls and instrument with the event-study approach. In all cases, I am interested in the evolution of the differential effect of having a CU mortgage on outcomes post origination.

I generically denote credit outcome variables as Y_{it}. The OLS specification includes the same set of controls X as in Section 5.2 and is specified as follows:

$$Y_{it} = \boldsymbol{\beta} CU_i \psi_{\tau(t,i)} + \boldsymbol{\delta} \psi_{\tau(t,i)} + \boldsymbol{\alpha} X_i + \epsilon_{it}. \tag{5}$$

$\psi_{\tau(t,i)}$ are event-time dummy variables that capture the number of months after origination that the credit outcome is being observed. The three mortgage-specific credit outcomes are limited to $\psi_{\tau(t,i)} > 0$, as the credit-profile outcomes include both pre- and post-event time dummies. The IV specification is

$$Y_{it} = \boldsymbol{\beta} \widehat{CU_i \psi_{\tau(t,i)}} + \boldsymbol{\delta} \psi_{\tau(t,i)} + \boldsymbol{\alpha} X_i + \epsilon_{it} \tag{6}$$

$$CU_{it} \psi_{\tau(t,i)} = \boldsymbol{\gamma^{post}} CuDensity_i \psi_{\tau(t,i)>0} + \boldsymbol{\gamma^{pre}} CU_i \psi_{\tau(t,i)<0} + \boldsymbol{\delta} X_i + \epsilon_{it}, \tag{7}$$

where equations (6) and (7) are the second and first stages, respectively. Because of the event time coefficients, Equation (7) is a system of as many equations as there are event time dummies. Note the instrument is only applied to the CU variable post-origination, as is standard in IV-DiD specifications applied to event-study frameworks (Hudson et al., 2017).

Figure 6 plots the $\boldsymbol{\beta}$ coefficients of equations (5) and (6) by event time, comparing CUs with small banks. The estimates of ordinary least squares (OLS) with controls corroborate the results from Section 4.3. The IV estimates are consistent with those results but have wider confidence intervals. Appendix Figure A12 shows results for the complete set of variables. The results in this section lend credence to the idea that CUs are doing something post origination to affect credit score outcomes in an advantageous way for their customers. Appendix Figure A13 shows the same results comparing CUs to large banks. Relative to large banks, the causal evidence is more mixed, but altogether leans in the same direction and confirms the qualitative conclusion that individuals with a CU mortgage experience better credit outcomes as a result of a CU effect.

6 Extensions, Mechanisms, and Heterogeneity

This section begins with a study of the auto loans market, in which I apply the analyses in the previous section of the mortgage market to a sample of auto loans derived from the TransUnion data. I afterward deepen the analysis of the mortgage market by showing evidence of mechanisms that mediate CU effects on prices and credit outcomes. First, I analyze ways in which CUs differ from banks in setting interest rates. Second, I study the differential treatment of loans that become past due. Third, I investigate whether bank loan resell rates are related to credit outcomes. Last, I investigate whether the effects on credit outcomes vary by time period: prior to 2009, post 2011, or in between these two years.

6.1 Auto Loans

Auto loans are the second largest type of consumer loans primarily supplied by commercial lenders (Federal Reserve Bank of New York, 2023). Auto loans are important because a) CUs have an even larger market share of auto loans than mortgages and b) individuals of lesser means are potentially more represented in the auto loans market. I analyze the approximately 7.8 million auto loans originated by either CUs or commercial banks found in the TransUnion data sample.

Table 3 reports the results from the IV specification in equation (6) on the interest rates at origination and loan and credit outcomes three years after origination. The results imply that interest rates at CUs are 46 basis points lower than at banks, which is a similar magnitude found for mortgages. The results on differences in loan and credit outcomes are also consistent with the equivalent for mortgages. CU loans are less likely to become past due and charged off or sent into collections. CU borrowers have lower amounts past due, higher credit scores, and fewer bankruptcies. Although I do not detect statistically significant effects on the number of trades past due, these results are consistent with and lend credence to the idea of a CU effect.

In Appendix Figure A14 I show that, among the variables that I can analyze, the distribution of the characteristics of auto loans is similar for CUs and banks. Although these data are limited to the sample in TransUnion data (i.e., they are not close to the universe, like in HMDA) and I can analyze fewer variables, the similarity between banks and CUs is even stronger in auto loans than it is in mortgages. These results suggest that the takeaways about CUs from the mortgage market apply also to the auto loans market.

6.2 Mechanism: Pricing Function

This section builds on Section 5.2 and analyzes whether CUs differ in setting interest rates. Figure 7 plots predicted interest rates for CUs and banks by credit score ventile from an OLS regression with no controls. Although CUs charge lower interest across the entire score distribution, they charge notably less than banks to those with credit scores in the bottom 30% to 40%. Given the importance of credit scores in determining prices, this finding by itself suggests lower prices as well as less risk-sensitive pricing.

To better understand the differences between CUs' and banks' pricing functions, the following analysis investigates whether CUs price differently based on each covariate I use. For each control variable $x \in X$, I estimate the following 2SLS regressions:

$$InterestRate_i = \beta_0 + \beta_1 \widehat{CU_i} + \beta_2 \widehat{CU_i \times x} + \boldsymbol{\alpha} X_i + \epsilon \tag{8}$$

$$CU_i = \gamma_0 + \gamma_1 CuDensity_i + \gamma_2 CuDensity_i \times x + \boldsymbol{\delta} X_i + \varepsilon \tag{9}$$

$$CU_i \times x = \gamma_0 + \gamma_1 CuDensity_i + \gamma_2 CuDensity_i \times x + \boldsymbol{\delta} X_i + \varepsilon. \tag{10}$$

These equations are similar to equations (3) and (4), but they add an interaction between x and CU that reveals a potentially different way in which CUs load on those variables to set interest rates. Equations (9) and (10) are the two first stage equations, both of which rely on the *CuDensity* instrument, and equation (8) is the second stage. For each regression, I am interested in the coefficients β_2 and α^x, the banks' coefficient corresponding to x. The coefficient α^x captures how a bank's interest rates are affected by x, and β_2 captures how CUs potentially differ from banks on this specific dimension.

Figure A15 plots, for each regression, the α_x coefficients in black squares and the β_2 coefficients in blue diamonds. Although my data are insufficiently powered to make definitive claims, many variables also indicate that CUs price in less risk-sensitive ways than banks. Below, I highlight the variables that support this claim:

- **Tract characteristics:** Bank interest rates are higher in tracts with higher rejection rates and lower in tracts with high mean credit scores. These coefficients are consistent with pricing based on costs or risk. The CU coefficient on interest rates leans in the opposite direction.

- **Loan purpose:** Relative to mortgages for a home purchase, banks charge lower interest on home improvement and refinancing loans. By contrast, CUs do not differentiate interest rates for refinancings and charge higher interest on home improvement loans.

- **Credit Score:** Relative to banks, CUs charge lower interest rates to both below median and top quartile credit score individuals.

- **Loan Term:** Relative to banks, CUs charge higher interest rates on 10-year term mortgages, approximately the same on 15-year term mortgages, and less on 20-year term mortgages. Overall, CUs price more uniformly in relation to the length of the mortgage.
- **Time Period:** Bank interest rates were highest prior to 2009, and CUs offered lower prices at this time. Since 2011, when interest rates decreased, CU and bank rates have converged.

6.3 Mechanism: Accommodations and Credit Outcomes of Mortgages Past Due

The results in Section 5 suggest that CUs do something, other than selection, which contributes to the improved outcomes among its borrowers. One possible such action that the data allow me to explore is the treatment of mortgages that become past due. In this section I report estimates of a) the likelihood that a past-due mortgage receives an accommodation and b) the likelihood that a past-due mortgage experiences an even more negative outcome.

Accommodations include forbearance, deferment, and being reported as affected by a natural disaster. Negative credit outcomes include foreclosure, repossession, being charged-off, and being sent to collections. The first column of Table 4 reports that overall, CU loans are approximately 40% less likely to be past due than small bank originated loans (0.0114 versus 0.197). The second and third columns show that, conditional on being past due, CU loans are are approximately 20% more likely to receive an accommodation (0.0410 vs 0.0339) and 10% less likely to experience a negative credit outcome (.2815 vs. 0.3159). Although there are other possible explanations for these differences, these results are suggestive of a post-origination channel by which CUs "treat" their borrowers with better outcomes through more accommodating and forgiving actions.

6.4 Mechanism: Mortgage Reselling

A natural next step to better understand what may be behind the improved credit outcomes is to ask what bank-related variables correlate with the improved outcomes. Because of how strikingly different CU and banks are in the likelihood of reselling mortgages, I further investigate how resell rates are correlated with outcomes. To simplify the analysis, I focus on credit outcomes three years after origination and, in the case of credit profile outcomes, I compare them to three months prior. I segment loans based on the resell rates of their originating bank. I create 5 binary variables based on the percentage of loans they resell within the same calendar year: $Resell := \{< 10, 10-33, 34-66, 67-90, > 90\}$. For mortgage-specific outcomes, I run the OLS specification in equation (11):

$$Y_{it} = \boldsymbol{\beta} Resell_i + \boldsymbol{\alpha} X_i + \epsilon_{it}. \tag{11}$$

For credit-profile outcomes, I run the OLS specification in equation (12):

$$Y_{it} = \boldsymbol{\alpha} Resell_i + Post_{it} + \boldsymbol{\beta} Resell_i \times Post_{it} + \boldsymbol{\alpha} X_i + \epsilon_{it}. \tag{12}$$

I exclude the 0-10% binary variables and estimate one coefficient for each category. For both mortgage-specific and credit-profile outcomes, the coefficients β are of interest because they capture the potential heterogeneity among banks when classified by resell rates and their impact on credit outcomes. Appendix Figure A16 plots the estimated coefficients for the seven credit outcomes for small and banks.

The results show that for mortgage-specific outcomes lower resell rates are associated with better outcomes. But this does not extend to credit-profile outcomes. Considering that CUs are much less likely to resell their mortgages, the differential impact on mortgage-specific outcomes is a potential channel of the kinds of bank-level behavior that lead to better credit outcomes. This evidence is consistent with the idea that a banking model based on relatively little reselling of mortgages better internalizes the incentives to achieve good credit outcomes.

6.5 Heterogeneity over Time

As shown in Appendix Figure A9, CUs were less likely to charge high interest rates prior to 2010. The difference, though, is driven by bank behavior, because banks charged high interest rates on as many as one in four loans. The pre-crisis period displayed high demand for mortgages, and presumably the ex-ante opportunities for lenders to profit were higher than at other points in time. CUs did not follow the overall trend that banks did. One interpretation of this fact is that CUs were less prone to pursue the profit opportunities that appeared to exist prior to 2008. Although for a different time period, Cororaton (2019) argues CUs forwent profit opportunities in order to extend more loans than banks. This analysis is also interesting because it reveals the possibility that the "CU effect" was observed during the pre-crisis period and may not exist since then.

I segment my sample into three time periods, based on when mortgages were originated: pre-crisis 2004–2008, bust 2009–2011, and the post-Dodd-Frank period 2012–2017. Appendix Table A6 shows the result of estimating the 2SLS equations (3)-(4) by time period (equivalent to column (3) of Table 2). Indeed, the CU effect on interest rates has declined over time. Additionally, Appendix Figure A17 shows the results from estimating equation (5) on the three data subsamples of each period. The symbols plot the coefficient β corresponding to three years post origination. Although the estimates are noisy given the smaller sample sizes, the CU effect on credit outcomes relative to small banks seems to have been greatest from 2004 to 2008 – especially among the estimates with statistical significance. Altogether, this analysis provides some support for the idea that prior to the financial crisis, when underwriting standards were lowest, CUs differentiated themselves the most with improved credit outcomes.

7 Conclusion

This paper develops empirical evidence that, relative to comparably-sized banks, Credit Unions charge lower interest rates on mortgages and that their borrowers have fewer mortgage delinquencies, higher credit scores, and a lower risk of bankruptcy several years later. These results are consistent with the idea that CUs behave differently than banks. The empirical evidence on the mortgage market in this paper weighs against the hypothesis that CUs are pure for-profits in disguise. Important questions remain in need of better evidence, such as whether CUs transfer all the tax exemption to their members or if CUs expand access to credit. That this research finds at least some transfer of the tax exemption to customers stands in relatively optimistic contrast to the majority of the literature on non-profit hospitals, which, with important exceptions, suggests they behave like for-profit firms.

References

Acharya, V. V., M. Richardson, S. van Nieuwerburgh, and L. J. White (2011). *Guaranteed to Fail: Fannie Mae, Freddie Mac, and the Debacle of Mortgage Finance*. Princeton: Princeton University Press.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy 113*(1), 151–184.

American Bankers Association (2021). Comparative state law charts.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association 91*(434), 444–455.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Argyle, B., C. J. Palmer, and T. Nadauld (2023). Real effects of search frictions in consumer credit markets. *Review of Financial Studies 36*(7), 2685–2720.

Board of Governors, F. R. S. (2017). Residential mortgage lending in 2016: Evidence from the home mortgage disclosure act data.

Capps, C. S., D. W. Carlton, and G. David (2020). Antitrust treatment of nonprofits: Should hospitals receive special care. *Economic Inquiry 58*(3), 1183–1199.

Card, D., A. Fenizia, and D. Silver (2020). The health impacts of hospital delivery practices. *Working Paper*.

Consumer Financial Protection Bureau (2016). Who are the credit invisibles?

Cororaton, A. (2019). Banking on the firm objective. *Working Paper*.

Duggan, M. G. (2000). Hospital ownership and public medical spending. *The Quarterly Journal of Economics 115*, 1343–1373.

Experian (2017). The state of credit unions.

Federal Reserve Bank of New York (2023). Quarterly report on household debt and credit.

Federal Reserve Bank of St. Louis (2017). Banking deserts" become a concern as branches dry up.

Federal Reserve Board (2008a). Design and testing of effective truth in lending disclosures.

Federal Reserve Board (2008b). Does distance matter in banking?

Flannery, M. J. (1974). An economic evaluation of credit unions in the united states. *Federal Reserve Bank of Boston, Research Report 54*.

Garthwaite, C., T. Gross, and M. Notowidigdo (2018). Hospitals as insurers of last resort. *American Economic Journal: Applied Economics 10*(1), 1–39.

Hansmann, H. B. (1980). The role of the nonprofit enterprise. *The Yale Law Journal 89*(5), 835–901.

Hauswald, R. and R. Marquez (2006). Competition and strategic information acquisition in credit markets. *Review of Financial Studies 19*(3), 967–1000.

Hirth, R. A. (1999). Consumer information and competition between nonprofit and for-profit nursing homes. *Journal of Health Economics 18*(2), 219–240.

Hudson, S., P. Hull, and J. Liebersohn (2017). Interpreting instrumented difference-in-differences.

Internal Revenue Service (1979). M. state chartered credit unions under 501(c)(14)(a).

Lacko, J. M. and J. K. Pappalardo (2007). Improving consumer mortgage disclosures: An empirical assessment of current and prototype disclosure forms.

Lakdawalla, D. and T. Philipson (2006). Nonprofit production and industry performance. *Journal of Public Economics 90*(8-9), 1681–1698.

Leggett, K. and Y. Stewart (1999). Multiple common bond credit unions and the allocation of benefits. *Journal of Economics and Finance 23*, 235–245.

Malani, A., T. Philipson, and G. David (2003). Theories of firm behavior in the nonprofit sector. a synthesis and empirical evaluation. In E. L. Glaeser (Ed.), *The Governance of Not-for-Profit Organizations*, Chapter 6, pp. 181–215. Chicago: University of Chicago Press.

Mountjoy, J. (2021). Community colleges and upward mobility. *Working Paper*.

National Credit Union Administration (2020). Quarterly credit union data summary 2020 q4.

Newhouse, J. P. (1970). Toward a theory of nonprofit institutions: An economics model of a hospital. *American Economic Review 60*(1), 64–74.

Oster, E. (2015). Unobserved selection and coefficient stability: Theory and evidence. *Working Paper*.

Pauly, M. and M. Redisch (1973). The not-for-profit hospital as a physicians' cooperative. *American Economic Review 63*(1), 87–99.

Petersen, M. A. and R. G. Rajan (1994). The benefits of lending relationships: Evidence from small business data. *Journal of Finance 49*(1), 3–37.

Philipson, T. J. and R. A. Posner (2009). Antitrust in the not-for-profit sector. *The Journal of Law and Economics 52*(1), 1–18.

Rehbein, O. and S. Rother (2020). The role of social networks in bank lending. *Working Paper*.

Rose-Ackerman, S. (1996). Altruism, non-profits and economic theory. *Journal of Economic Literature 34*, 701–728.

Smith, D. J. (1981). A theoretic framework of the analysis of credit union decision making. *The Journal of Finance 39*, 1155–1168.

Stango, V. and J. Zinman (2016). Borrowing high versus borrowing higher: Price dispersion and shopping behavior in the u.s. credit card market. *Review of Financial Studies 29*(4), 979–1006.

Urban Institute, T. (2020). The nonprofit sector in brief 2019.

U.S. Treasury (2001). Comparing credit unions with other depository institutions.

van Rijn, J. (2018). The effect of membership expansion on credit union risk and returns.

van Rijn, J. and K. Li (2019). Credit union and bank subprime lending in the great recession. *Working Paper*.

van Rijn, J., S. Zeng, and P. Hellman (2021). Financial institution objectives and auto loan pricing: Evidence from the survey of consumer finances. *Journal of Consumer Affairs 55*(3), 995–1039.

van Rijn, J., S. Zeng, and B. Hueth (2019). Do credit unions have distinct objectives? evidence from executive compensation structures. *Working Paper*.

Warner, J. T. and S. Pleeter (2001). The personal discount rate: Evidence from military downsizing programs. *American Economic Review 91*(1), 33–53.

Weisbrod, B. A. (1975). Toward a theory of the voluntary nonprofit sector in a three-sector economy. In E. S. Phelps (Ed.), *In Altruism, Morality, and Economic Theory,*, pp. 171–195. New York: Russell Sage Foundation.

Woodward, S. E. and R. E. Hall (2012). Diagnosing consumer confusion and sub-optimal shopping effort: Theory and mortgage-market evidence. *American Economic Review 102*(7), 3249–3276.

8 Figures and Tables

Table 1.
Selection into CUs and Representativeness of Matched Sample

	CU		Small Bank		Large Bank	
	Sample	HMDA	Sample	HMDA	Sample	HMDA
Loan Characteristics						
– Loan Amount ($K)	184	139	231	191	273	224
– Applicant Income ($K)	105	96	121	103	135	113
Tract Characteristics						
– Credit Score	50	49	51	50	50	50
– Reject Rate	0.17	0.18	0.18	0.19	0.18	0.19
– Fraction Urban	0.80	0.81	0.75	0.78	0.84	0.86
– Median Income ($K)	70	67	71	67	78	73
Bank Characteristics						
– Loan Reject Rate	0.19	0.17	0.16	0.17	0.19	0.19
– Loan Resell Rate	0.36	0.30	0.65	0.60	0.70	0.68
– Assets ($M)	4,975	4,629	7,228	6,539	870,779	612,748
Lien Status						
– First	0.89	0.73	0.95	0.85	0.97	0.89
– Subordinate	0.11	0.21	0.04	0.12	0.03	0.10
– Not Secured	0.00	0.06	0.00	0.03	0.00	0.02
Loan Purpose						
– Home Purchase	0.24	0.24	0.35	0.38	0.24	0.33
– Home Improvement	0.10	0.22	0.05	0.11	0.04	0.07
– Refinance	0.66	0.55	0.60	0.52	0.72	0.60
Year						
– Pre-2009	0.10	0.47	0.19	0.30	0.20	0.25
– 2009 to 2011	0.21	0.33	0.24	0.52	0.28	0.56
– Post-2011	0.69	0.20	0.57	0.18	0.52	0.19
Borrower Age						
– Under 35	0.17		0.21		0.16	
– 35-60	0.65		0.64		0.66	
– Over 60	0.18		0.15		0.18	
Credit Score Percentile						
– Below 50	0.13		0.13		0.13	
– 50-75	0.40		0.40		0.38	
– Above 75	0.47		0.46		0.49	
Term (Years)						
– 10	0.14		0.06		0.05	
– 15	0.28		0.22		0.21	
– 20	0.07		0.06		0.07	
– 30	0.51		0.66		0.66	
N	284,696	7,639,394	868,671	27,608,599	1,163,459	33,234,647

Notes: Mortgage data are from HMDA, and credit records data are from TransUnion. The HMDA columns show means for the universe of conventional (i.e., non- FHA, VA, FSA, or RHS) mortgages with a first-lien on an owner-occupied, one-to-four family property in HMDA from 2004 to 2017. Means for tract- and bank-level characteristics are weighted by the number of originated mortgages.The Sample columns show means for the matched sample. Further details on the data match and variable constructions can be found in Section 3.

Table 2.
Differential Interest Rates at CUs vs. Banks

Relative to Small Banks

	Dependent variable: Interest Rate		
	(1)	(2)	(3)
CU	−0.189	−0.107	−0.295
	(0.060)	(0.029)	(0.087)
Controls	No	Yes	Yes
IV	No	No	Yes
Adj. R2	0.0026	0.4497	
1st Stage F-Stat.			8,482.48
N	767,991	767,991	767,991

Relative to Large Banks

	Dependent variable: Interest Rate		
	(1)	(2)	(3)
CU	−0.152	−0.046	−0.434
	(0.156)	(0.070)	(0.619)
Controls	No	Yes	Yes
IV	No	No	Yes
Adj. R2	0.0024	0.4171	
1st Stage F-Stat.			508.04
N	1,024,516	1,024,516	1,024,516

Notes: Data are from TransUnion merged with HMDA. This table reports β_1 estimates from the following three regressions investigating the CU effect on mortgage interest rates:

$$InterestRate_i = \beta_0 + \beta_1 CU_i + \epsilon$$
$$InterestRate_i = \beta_0 + \beta_1 CU_i + \boldsymbol{\alpha} X_i + \epsilon$$
$$InterestRate_i = \beta_0 + \beta_1 \widehat{CU}_i + \boldsymbol{\alpha} X_i + \epsilon$$
$$CU_i = \gamma_0 + \gamma_1 CuDensity_i + \boldsymbol{\delta} X_i + \varepsilon$$

X_i represents a vector of controls for borrower creditworthiness, loan terms, and geographic factors. Standard errors clustered at the bank level are reported in parentheses.

Table 3.
Auto Loans

Loan Outcomes

	Interest Rate	90+ Days Past Due	Amount Past Due	Charged-off or In Collections
	(1)	(2)	(3)	(4)
CU	-0.460	-0.002	-0.153	-0.011
	(0.198)	(0.002)	(0.036)	(0.006)
Controls	Yes	Yes	Yes	Yes
IV	Yes	Yes	Yes	Yes
1st Stage F-Stat.	2,700.7	2,700.7	2,700.7	2,700.7
N	7,848,838	7,848,838	7,848,838	7,848,838

Credit Outcomes

	Trades 90 Days Past Due	Amount Past Due	Score	Num. of Public Rec. Bankruptcies
	(5)	(6)	(7)	(8)
CU	0.004	-1.70	2.83	-0.055
	(0.017)	(0.431)	(1.03)	(0.017)
Controls	Yes	Yes	Yes	Yes
IV	Yes	Yes	Yes	Yes
1st Stage F-Stat.	2,656.0	2,656.0	2,656.0	2,656.0
N	7,775,323	7,775,323	7,775,323	7,775,323

Notes: Data are from TransUnion. This table reports on an analysis of the CU effect on auto loan interest rates, outcomes, and borrower credit outcomes, three years from loan origination. The table reports $\beta_{\tau=3} \in \boldsymbol{\beta}$ coefficient estimates from the following instrumented event study regressions (also detailed equations (6) and (7) of the text):

$$Y_{it} = \boldsymbol{\beta} \widehat{CU_i \psi_{\tau(t,i)}} + \boldsymbol{\delta} \psi_{\tau(t,i)} + \boldsymbol{\alpha} X_i + \epsilon_{it}$$

$$CU_{it} \psi_{\tau(t,i)} = \boldsymbol{\gamma}^{\boldsymbol{post}} CuDensity_i \psi_{\tau(t,i)>0} + \boldsymbol{\gamma}^{\boldsymbol{pre}} CU_i \psi_{\tau(t,i)<0} + \boldsymbol{\delta} X_i + \epsilon_{it},$$

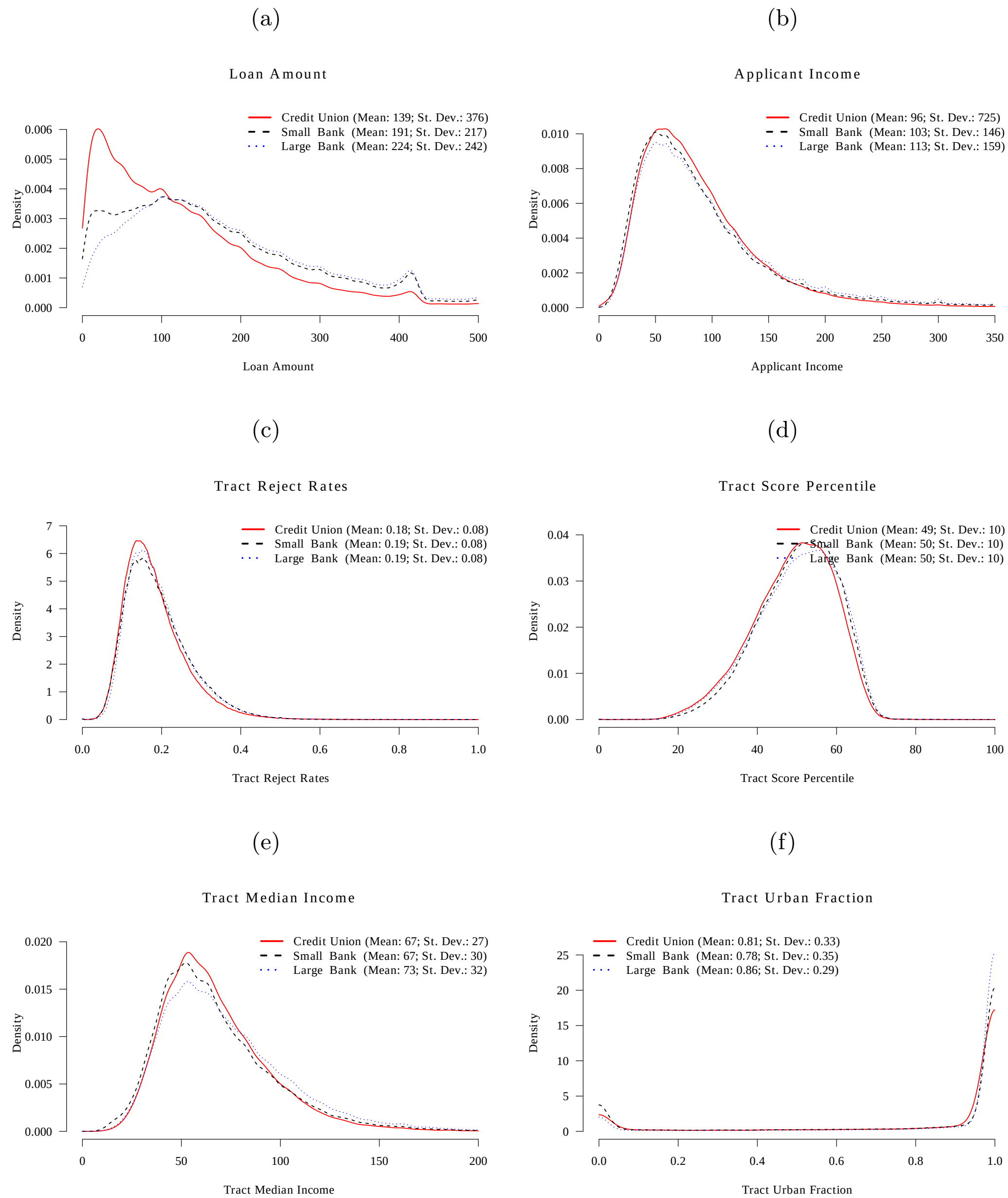
$\psi_{\tau(t,i)}$ are event-time dummy variables that capture the number of months after origination that the credit outcome is being observed. X_i represents a vector of controls for borrower creditworthiness, loan terms, and geographic factors. $d(i,l)$ is the distance from branch of lender l to the individual i's address. The instrument is then the density of CU branches over all CU and bank branches, weighted by the inverse of their distance from the new address: $CuDensity_i = \frac{\sum_{l \in CU} \frac{1}{d(i,l)}}{\sum_{l \in CU} \frac{1}{d(i,l)} + \sum_{l \in Bank} \frac{1}{d(i,l)}}$. Standard errors are reported in parentheses.

Table 4.
Accommodations and Negative Outcomes on Loans Past Due

		Conditional on Ever Being 90+ Days Past Due	
	Ever 90+ Past Due	Accommodation	Negative Outcome
CU	0.0114	0.0410	0.2815
Small Bank	0.0197	0.0339	0.3159
Difference	-0.0083	0.0070	-0.0344
p-value:	0.0000	0.0275	0.0000

Notes: Data are from TransUnion merged with HMDA. The first column of this table reports the fraction of loans that were 90+ days past due at least once during the loan's lifetime. The second column reports the fraction of loans that, conditional on ever being 90+ days past due, received an accommodation. A loan is considered to have received an accommodation it was in deferment or in forbearance. The third column reports the fraction of loans that, conditional on ever being 90+ days past due, ultimately experienced a negative outcome. A loan is considered to have experienced a negative outcome if it was ever in foreclosure, repossession, charged off, or in collections. Reported p-values are for tests of difference in proportions, based on one-sided alternative hypotheses (the alternative hypothesis is that the CU proportion is smaller, larger, and smaller, for each column respectively).

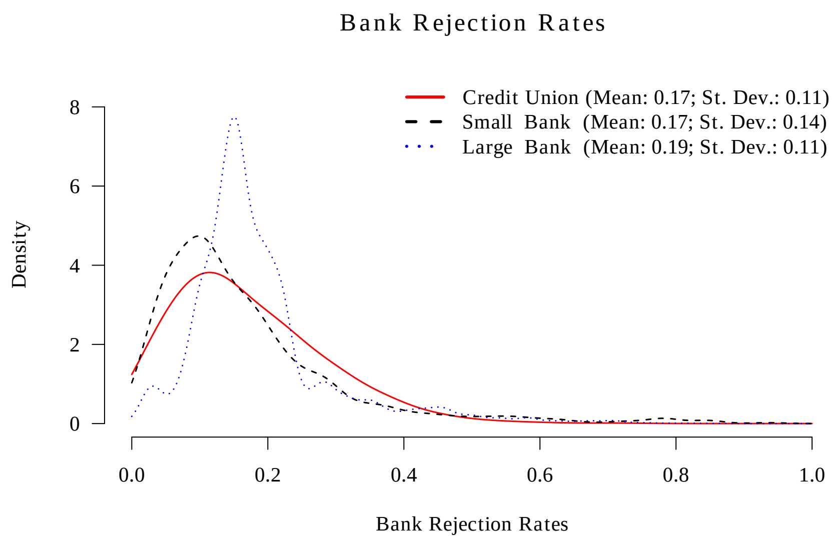
Figure 1.
Characteristics of Borrowers and Loans



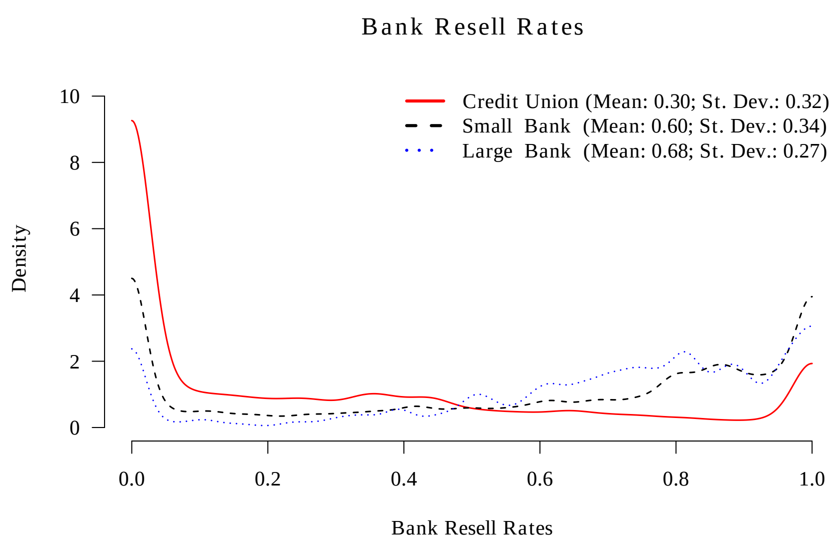
Notes: Data are from HMDA and credit records data are from TransUnion. Each subfigure plots kernel density estimates based on the universe of conventional (i.e., non- FHA, VA, FSA, or RHS) mortgages with a first-lien on an owner-occupied, one-to-four family property in HMDA from 2004 to 2017. Each plot shows the density of a given variable weighted by the number of originated mortgages. Each figure breaks out density estimates by the type of bank that originated the mortgage: Credit Unions, small banks, or large banks. A bank is defined as a small/large bank if its total assets reported to HMDA in the given year are less/more than 105% the total assets of the largest Credit Union in that year.
In Panels (a) and (b), amounts are presented in thousands of US Dollars. In Panel (c), tract reject rates are computed at the tract-year level, and the fraction of applications from a given tract that were reported to HMDA and were not originated due to a denial. In Panel (d), mean tract score percentiles are calculated at the tract-year level. Tract median income in Panel (e) is in thousands of constant 2012 US dollars and come from Census data. Panel (f) shows the distribution of tracts' urban fractions according to Census data. Densities are estimated using Gaussian kernels. Smoothing bandwidths are constant within each plot, although they differ across plots. To faithfully reflect the mass of data at the 0 and 1 bounds of the measure, Panel (f) "reflects" the smoothed data outside 0 and 1 boundaries and compute densities on the reflected vector limited to the [0,1] range. Let X be a vector such that its elements $x_i \in [0,1]$, the reflection R of vector X is defined as $R := \{-X, X, 2-X\}$. Densities are computed on R instead of X and multiplied by three to account for the added data.

Figure 2.
Mortgage Banking Models

(a)

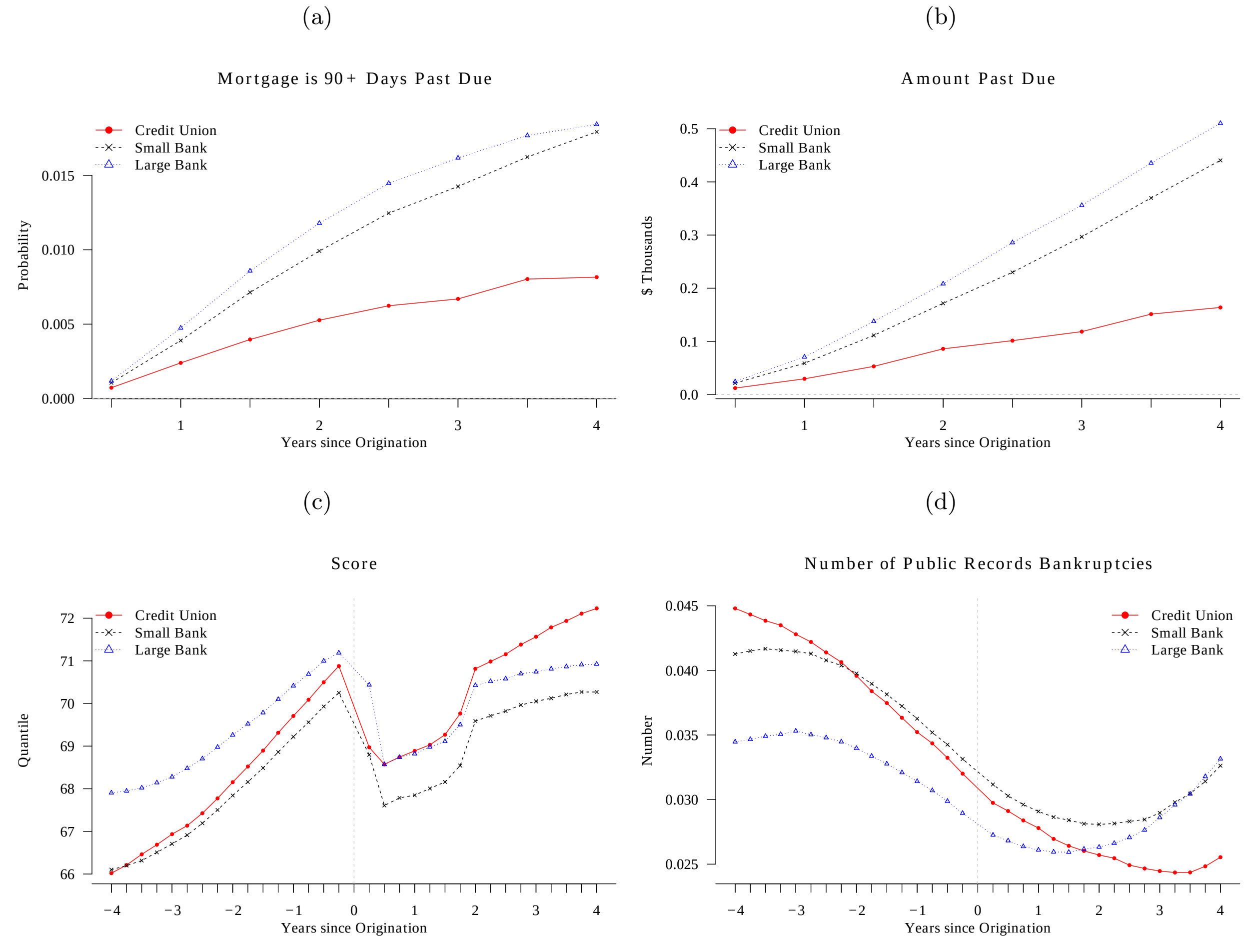


(b)



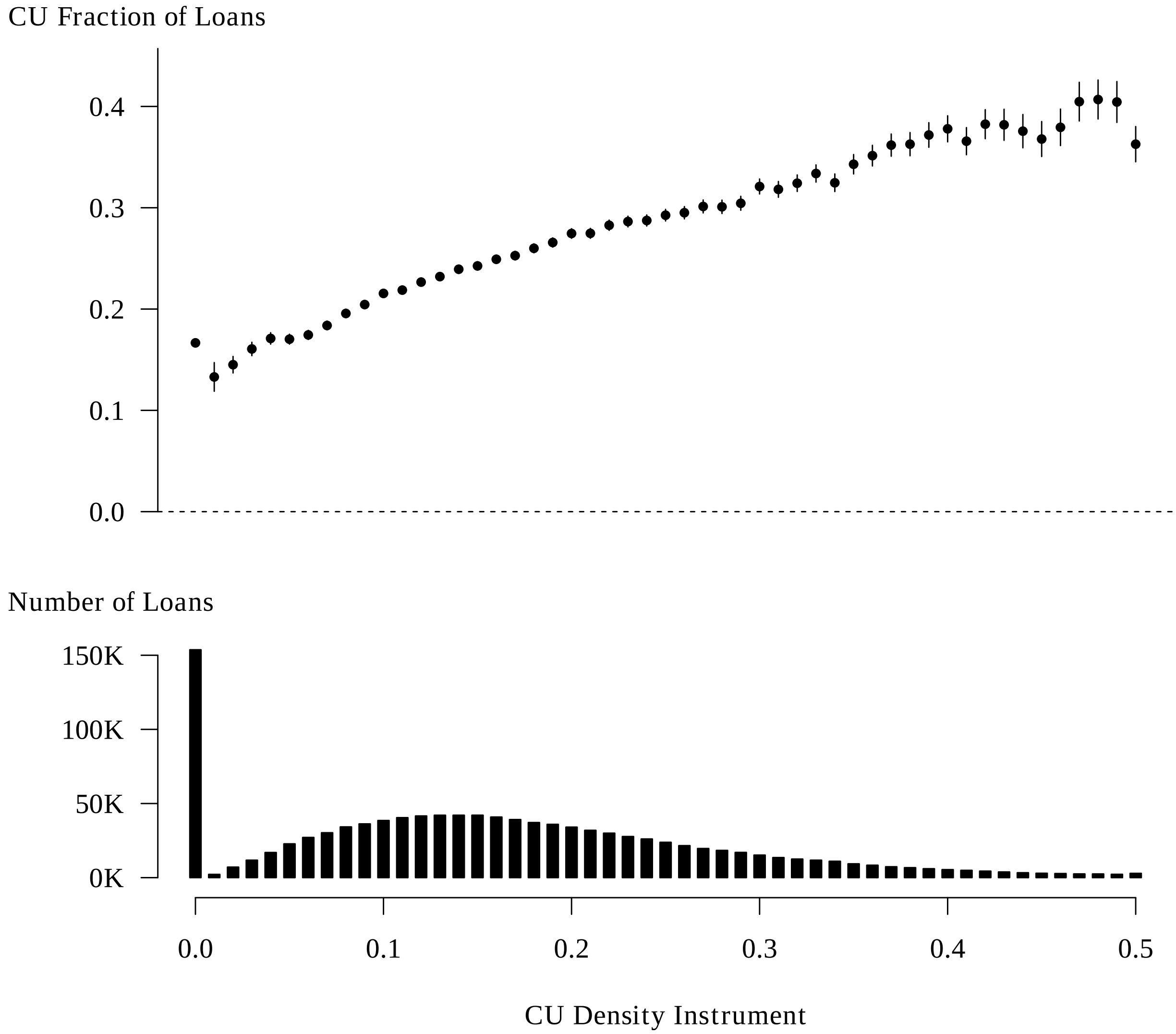
Notes: Data are from HMDA and credit records data are from TransUnion. Each subfigure plots kernel density estimates based on the universe of conventional (i.e., non- FHA, VA, FSA, or RHS) mortgages with a first-lien on an owner-occupied, one-to-four family property in HMDA from 2004 to 2017. Each plot shows the density of a given variable weighted by the number of originated mortgages. Each figure breaks out density estimates by the type of bank that originated the mortgage: Credit Unions, small banks, or large banks. A bank is defined as a small/large bank if its total assets reported to HMDA in the given year are less/more than 105% the total assets of the largest Credit Union in that year. In Panel (a), bank rejection rates are calculated at the bank-year level and represent the fraction of applications that a bank reported to HMDA and were not originated due to a denial. In Panel (b), bank resell rates are calculated at the bank-year level and represent the fraction of loans that were originated by a lender and sold within the same calendar year. Panel (b) “reflects” the smoothed data outside 0 and 1 boundaries and compute densities on the reflected vector limited to the [0,1] range. Let X be a vector such that its elements $x_i \in [0, 1]$, the reflection R of vector X is defined as $R := \{-X, X, 2 - X\}$. Densities are computed on R instead of X and multiplied by three to account for the added data.

Figure 3.
Credit Outcomes



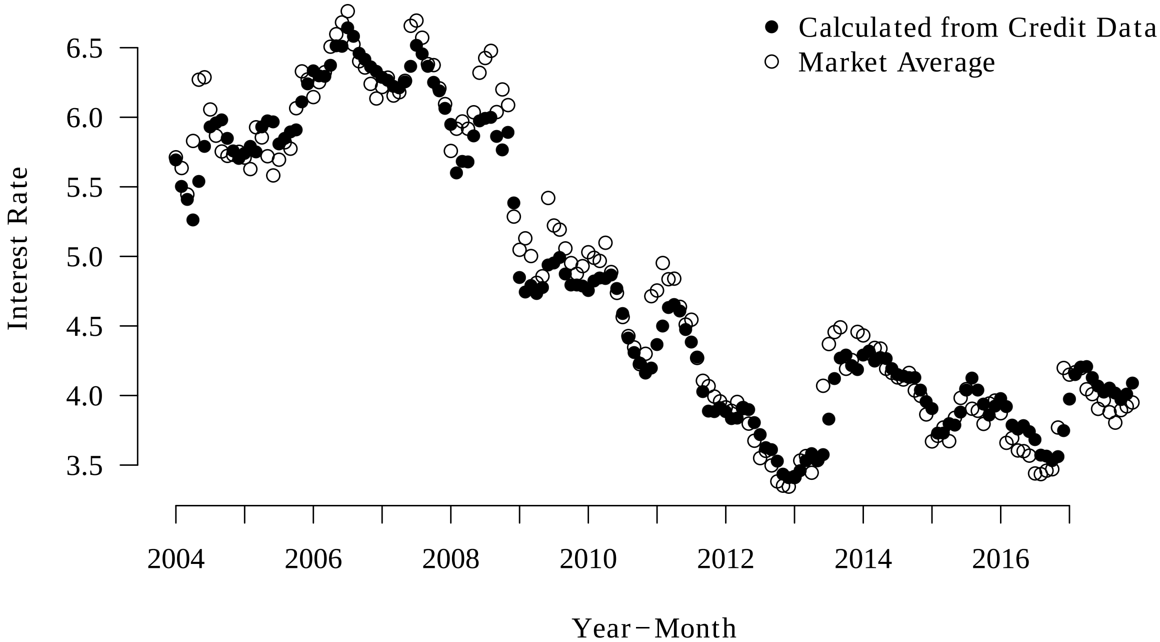
Notes: Data are from TransUnion merged with HMDA and restricted to 30-year term mortgages. Plots show the mean value by bank type of each variable according to event time, defined from the month of mortgage origination.

Figure 4.
Unconditional Instrument First Stage



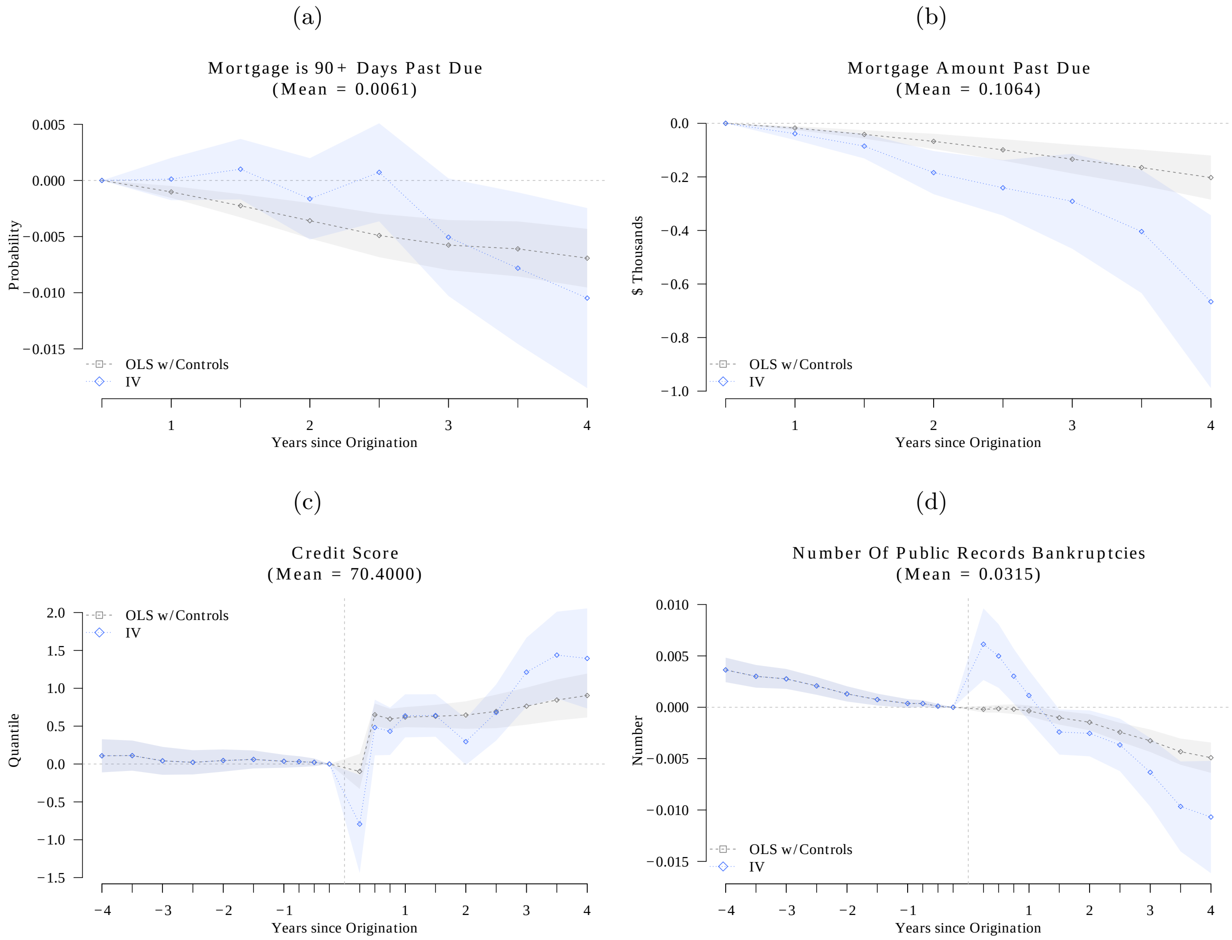
Notes: Data are from TransUnion merged with HMDA and restricted to CUs and small banks. The top panel plots the fraction of individuals using a CU by 0.01 instrument point bins. The vertical lines represent confidence intervals of the fraction CU. The CU density instrument is defined as follows: let $d(i, l)$ be the distance from each branch of a lender l to the individual i's address. The instrument is then the density of CU branches over all CU and bank branches, weighted by the inverse of their distance from the new address: $CuDensity_i = \frac{\sum_{l \in CU} \frac{1}{d(i,l)}}{\sum_{l \in CU} \frac{1}{d(i,l)} + \sum_{l \in Bank} \frac{1}{d(i,l)}}$. The bottom panel plots a histogram of observations per bin. Figure excludes the 1.7% of observations with CU density above 0.5.

Figure 5.
Derived vs. Actual Interest Rates



Notes: Credit data are from TransUnion merged with HMDA and restricted to 30-year term mortgages. Market Average data are the of 30-year fixed rate mortgage average estimates from FRED. Plot shows the mean of originations and market averages by month. Section 5.2.1 details the algorithm used to back out implied interest rates. Interest rates are calculated based on the evolution of outstanding loan balances throughout the maturity of the loan.

Figure 6.
CU Treatment Effect on Mortgage Outcomes



Notes: Data are from TransUnion merged with HMDA. Plots show estimates of the differential effect that originating a mortgage with a CU has relative to originating a mortgage with a **small** bank. The shaded areas represent 95% confidence intervals based on standard errors clustered at the bank level. Appendix Figure A13 contains the equivalent results when the reference group is large banks. The OLS w/controls estimates report $\boldsymbol{\beta}$ estimates from the following regression:

$$Y_{it} = \boldsymbol{\beta} CU_i \psi_{\tau(t,i)} + \boldsymbol{\delta}\psi_{\tau(t,i)} + \boldsymbol{\alpha} X_i + \epsilon_{it}$$

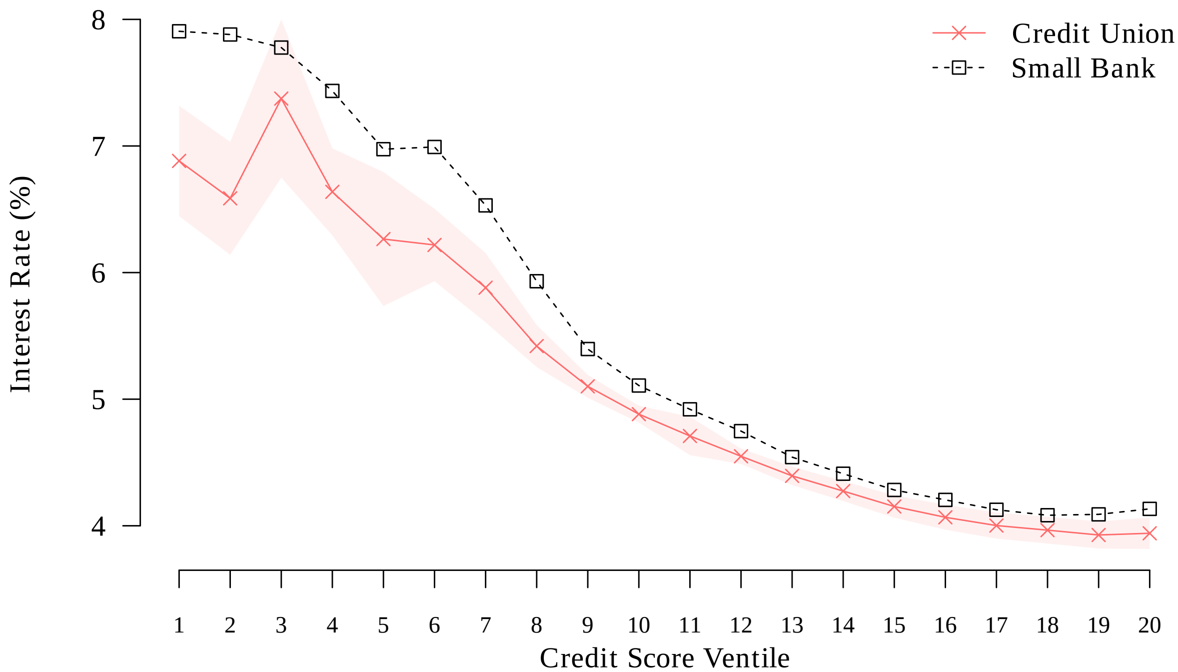
which is also detailed equation (5) of the text. The IV estimates in blue diamonds report $\boldsymbol{\beta}$ coefficient estimates from the following instrumented event study regression:

$$Y_{it} = \boldsymbol{\beta} \widehat{CU_i \psi_{\tau(t,i)}} + \boldsymbol{\delta}\psi_{\tau(t,i)} + \boldsymbol{\alpha} X_i + \epsilon_{it}$$

$$CU_{it}\psi_{\tau(t,i)} = \boldsymbol{\gamma^{post}} CuDensity_i \psi_{\tau(t,i)>0} + \boldsymbol{\gamma^{pre}} CU_i \psi_{\tau(t,i)<0} + \boldsymbol{\delta} X_i + \epsilon_{it},$$

which is also detailed equations (6) and (7) of the text. $\psi_{\tau(t,i)}$ are event-time dummy variables that capture the number of months after origination that the credit outcome is being observed. X_i represents a vector of controls for borrower creditworthiness, loan terms, and geographic factors. $d(i,l)$ is the distance from branch of lender l to the individual i's address. The instrument is then the density of CU branches over all CU and bank branches, weighted by the inverse of their distance from the new address: $CuDensity_i = \frac{\sum_{l \in CU} \frac{1}{d(i,l)}}{\sum_{l \in CU} \frac{1}{d(i,l)} + \sum_{l \in Bank} \frac{1}{d(i,l)}}$. Standard errors are reported in parentheses.

Figure 7.
Interest Rates by Credit Score



Notes: Data are from TransUnion merged with HMDA and restricted to CUs and small banks. Plots coefficients from a regression of interest rates on credit score ventiles interacted with a CU dummy variable.

Internet Appendix

A Appendix Figures and Tables

Appendix Table A1.
Complier Characteristics

Variable	Percent Compliers
Loan Characteristics	
– log(Loan Amount)	6.0
– Loan was Resold	6.5
– log(Applicant Income)	7.9
Tract Characteristics	
– Credit Score	8.6
– Reject Rate	9.7
– Fraction Urban	9.4
– Median Income	8.2
Bank Characteristics	
– Loan Resell Rate	4.3
– Loan Reject Rate	10.6
– log(Assets)	5.9
Lien Status	
– First	9.7
– Subordinate	11.7
– Not Secured	19.5
Loan Purpose	
– Home Purchase	7.9
– Home Improvement	14.7
– Refinance	10.8
Years	
– Pre-2009	6.7
– 2009 to 2011	8.4
– Post-2011	11.1
Age	
– Under 35	9.6
– 35-60	9.8
– Over 60	11.8
Credit Score Percentile	
– Below 50	11.3
– 50-75	10.3
– Above 75	9.4
Term (Years)	
– 10	17.1
– 15	11.6
– 20	11.0
– 30	7.7
Overall Percent Compliers	10.0
Percent Compliers of Treated	20.2

Notes: Data are from TransUnion merged with HMDA, limited to CUs and small banks. Complier percentages are calculated using the methodology in Angrist and Pischke (2009) after discretizing the instrument into a binary variable based on the median value of the instrument. For binary control variables B, table reports the percent of compliers among those for whom $B = 1$. For continuous control variables C, table reports the percent of compliers among those for whom $C >= median(C)$.

Appendix Table A2.

Stability of Instrument First Stage to Variations in Controls

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Instrument	0.45	0.40	0.36	0.33	0.32	0.33	0.33	0.33
		(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Loan Characteristics	log(Loan Amount)		−0.04	−0.06	−0.04	−0.04	−0.00	−0.00	−0.00
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
	Loan was Resold		−0.24	−0.25	−0.06	−0.06	−0.05	−0.05	−0.05
			(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
	log(Applicant Income)		0.00	−0.01	−0.01	−0.02	−0.04	−0.04	−0.04
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Tract Characteristics	Credit Score			−0.03	−0.05	−0.03	−0.03	−0.03	−0.03
				(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
	Reject Rate			−0.05	−0.06	−0.03	−0.02	−0.02	−0.02
				(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
	Fraction Urban			0.05	0.07	0.08	0.08	0.08	0.08
				(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
	Median Income			0.00	0.00	0.00	0.00	0.00	0.00
				(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Bank Characteristics	Loan Resell Rate				−0.39	−0.39	−0.39	−0.39	−0.39
					(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
	Loan Reject Rate				0.09	0.10	0.07	0.07	0.07
					(0.16)	(0.15)	(0.15)	(0.15)	(0.15)
	log(Assets)				−0.03	−0.03	−0.03	−0.03	−0.03
					(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Years	Pre-2009					−0.12	−0.11	−0.11	−0.10
						(0.02)	(0.01)	(0.01)	(0.01)
	Post-2011					0.04	0.05	0.05	0.05
						(0.01)	(0.01)	(0.01)	(0.01)

Continued on next page

Appendix Table A2 – continued from previous page

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Loan Purpose	Subordinate						0.11	0.11	0.11
							(0.02)	(0.02)	(0.02)
	Not Secured						−0.02	−0.02	−0.02
							(0.08)	(0.08)	(0.08)
Lien Status	Home Improvement						0.05	0.04	0.04
							(0.02)	(0.02)	(0.02)
	Refinance						0.07	0.07	0.07
							(0.01)	(0.01)	(0.01)
Term (Years)	10						0.07	0.07	0.07
							(0.02)	(0.02)	(0.02)
	15						0.02	0.02	0.02
							(0.01)	(0.01)	(0.01)
	20						−0.01	−0.01	−0.01
							(0.01)	(0.01)	(0.01)
Borrower Age	Under 35							−0.01	−0.01
								(0.00)	(0.00)
	Over 60							−0.00	−0.00
								(0.00)	(0.00)
Credit Score	Below 50								−0.01
									(0.00)
	Above 75								0.00
									(0.00)
	Constant	0.17	0.53	0.82	1.48	1.42	1.27	1.28	1.29
		(0.02)	(0.04)	(0.06)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)
	N	1,153,367	1,153,367	1,153,367	1,153,367	1,153,367	1,153,367	1,153,367	1,153,367
	Adj. R2	0.02	0.11	0.12	0.18	0.19	0.2	0.2	0.2

Notes: Data are from TransUnion merged with HMDA, limited to CUs and small banks. The table reports first-stage OLS estimates of regressing a CU indicator dummy on *CuDensity* and a varying set of controls. Data are from TransUnion merged with HMDA and restricted to CUs and small banks.

Appendix Table A3.
Preditiveness of Controls on CU Indicator and Instrument

		(1)	(2)
Loan Characteristics	log(Loan Amount)	−1.106	−0.704
		(0.025)	(0.078)
	Loan was Resold	−0.048	−5.340
		(0.034)	(0.106)
	log(Applicant Income)	−0.299	−4.066
		(0.025)	(0.078)
Tract Characteristics	Credit Score	−0.655	−3.132
		(0.017)	(0.053)
	Reject Rate	−1.396	−2.838
		(0.021)	(0.064)
	Fraction Urban	6.361	9.911
		(0.035)	(0.110)
	Median Income	−0.058	0.077
		(0.001)	(0.002)
Bank Characteristics	Loan Resell Rate	−1.419	−39.561
		(0.049)	(0.151)
	Loan Reject Rate	0.035	6.873
		(0.105)	(0.327)
	log(Assets)	−0.166	−3.135
		(0.007)	(0.021)
Years	Pre-2009	−0.177	−10.494
		(0.041)	(0.127)
	Post-2011	0.287	5.519
		(0.029)	(0.091)
Loan Purpose	Home Improvement	−0.790	10.804
		(0.063)	(0.195)
	Refinance	−1.428	−2.540
		(0.237)	(0.735)
Lien Status	Subordinate	−0.556	4.246
		(0.055)	(0.171)
	Not Secured	−0.388	6.550
		(0.027)	(0.085)
Term (Years)	10	0.604	6.884
		(0.051)	(0.157)
	15	0.036	1.810
		(0.031)	(0.097)
	20	−0.304	−1.461
		(0.050)	(0.155)
Borrower Age	Under 35	0.178	−1.281
		(0.031)	(0.096)
	Over 60	−0.481	−0.285
		(0.034)	(0.104)
Credit Score	Below 50	0.006	−1.412
		(0.038)	(0.118)
	Above 75	−0.223	0.065
		(0.026)	(0.080)
	Constant	33.038	139.461
		(0.201)	(0.623)
	Observations	1,153,367	1,153,367
	Adjusted R^2	0.066	0.195
	F Statistic (df = 23; 1153343)	3,561.916	12,184.350

Notes: Data are from TransUnion merged with HMDA, limited to CUs and small banks. The table reports first-stage OLS estimates of regressing a CU indicator dummy on $CuDensity$ and a varying set of controls. Data are from TransUnion merged with HMDA and restricted to CUs and small banks.

Appendix Table A4.
Differential Likelihood of Identifying Interest Rates

Category	Variable	Coefficient	St. Error
	CU	0.034	(0.008)
Loan Characteristics	log(Loan Amount)	0.020	(0.003)
	Loan was Resold	0.017	(0.004)
	log(Applicant Income)	−0.029	(0.002)
Tract Characteristics	Reject Rate	0.010	(0.002)
	Credit Score	0.001	(0.002)
	Fraction Urban	0.001	(0.003)
	Median Income	−0.000	(0.000)
Bank Characteristics	Loan Reject Rate	−0.059	(0.013)
	Loan Resell Rate	0.106	(0.031)
	log(Assets)	−0.002	(0.002)
Years	Pre-2009	−0.166	(0.009)
	Post-2011	−0.031	(0.008)
Loan Purpose	Home Improvement	−0.001	(0.008)
	Refinance	−0.136	(0.015)
Lien Status	Subordinate	0.015	(0.005)
	Not Secured	−0.004	(0.003)
Term (Years)	10	−0.130	(0.007)
	15	−0.124	(0.003)
	20	−0.040	(0.003)
Borrower Age	Under 35	0.003	(0.001)
	Over 60	0.014	(0.002)
Credit Score	Below 50	−0.004	(0.003)
	Above 75	−0.023	(0.002)
	Constant	0.798	(0.057)
	N	1,153,367	
	Adj. R2	0.03	

Notes: Data are from TransUnion merged with HMDA, limited to CUs and small banks. The table reports results from regressing an indicator variable for whether the observation has a successfully estimated interest rate or not. Standard errors calculated at the lender-level are reported in parentheses. Section 5.2.1 details the calculation of implied interest rates

Appendix Table A5.
Sensitivity of CU Effect on Interest Rate to Instrument Choice

	Within 10km		
	Inverse Distance	Negative Exponential	Branch Count
Credit Union	−0.295	−0.305	−0.305
	(0.087)	(0.090)	(0.092)
N	767,991	767,991	767,991
First-Stage F-Statistic	8,482.48	8,786.85	8,454.17

	Nearest 20 Branches		
	Inverse Distance	Negative Exponential	Branch Count
Credit Union	−0.291	−0.303	−0.294
	(0.081)	(0.083)	(0.082)
N	767,991	767,991	767,991
First-Stage F-Statistic	9,370.63	10,536.06	10,631.99

	Within 10km or Nearest 20 Branches		
	Inverse Distance	Negative Exponential	Branch Count
Credit Union	−0.308	−0.319	−0.311
	(0.085)	(0.089)	(0.089)
N	767,991	767,991	767,991
First-Stage F-Statistic	11,356.87	12,189.42	12,148.86

Notes: Data are from TransUnion merged with HMDA, limited to CUs and small banks. The table reports estimates of the β_1 coefficient from equation (3) for nine different instrumental variable specifications. The instrument and its alternative specifications are defined in Section 5.1. The coefficient for "Within 10km" and "Inverse Distance" corresponds to the estimate reported in column (3) of the top panel of Table 2.

Appendix Table A6.
Differential Interest Rates at CUs vs. Banks by Time Period

	Dependent variable: Interest Rate		
	'04-'08	'09-'11	'12-'17
CU	−0.973	−0.414	−0.185
	(0.271)	(0.151)	(0.090)
Controls	Yes	Yes	Yes
IV	Yes	Yes	Yes
1st Stage F-Stat.	801.6	1,331.73	6,635.85
N	111,297	187,225	469,469

Notes: Data are from TransUnion merged with HMDA, limited to CUs and small banks. This table reports β_1 estimates from the following three regressions investigating the CU effect on mortgage interest rates:

$$InterestRate_i = \beta_0 + \beta_1 CU_i + \epsilon$$

$$InterestRate_i = \beta_0 + \beta_1 CU_i + \boldsymbol{\alpha} X_i + \epsilon$$

$$InterestRate_i = \beta_0 + \beta_1 \widehat{CU}_i + \boldsymbol{\alpha} X_i + \epsilon$$

$$CU_i = \gamma_0 + \gamma_1 CuDensity_i + \boldsymbol{\delta} X_i + \varepsilon$$

which are also detailed in equations (1)–(3) of the main text. X_i represents a vector of controls for borrower creditworthiness, loan terms, and geographic factors. Standard errors clustered at the bank level are reported in parentheses.

Appendix Table A7.
Differential Interest Rates at CUs vs Small Banks: Complete Results

		OLS	IV	1st Stage	CU	Control	CU×Control
	Credit Union	−0.107	−0.295				
		(0.029)	(0.087)				
	Instrument CU Density			0.003			
				(0.000)			
Loan	log(Loan Amount)	−0.152	−0.152	−0.003	0.019	−0.134	−0.067
		(0.010)	(0.010)	(0.007)	(0.415)	(0.024)	(0.081)
	Resold	−0.120	−0.129	−0.053	−0.225	−0.081	−0.159
		(0.016)	(0.017)	(0.012)	(0.118)	(0.042)	(0.129)
	log(Applicant Income)	−0.030	−0.038	−0.040	−0.250	−0.036	−0.011
		(0.007)	(0.007)	(0.005)	(0.272)	(0.015)	(0.061)
Tract	Mean Credit Score	−0.049	−0.056	−0.029	−0.755	−0.081	0.095
		(0.006)	(0.006)	(0.008)	(0.228)	(0.013)	(0.041)
	Reject Rate	0.038	0.033	−0.024	0.118	0.093	−0.237
		(0.007)	(0.006)	(0.010)	(0.110)	(0.016)	(0.060)
	Fraction Urban	−0.008	0.010	0.078	−0.435	−0.037	0.201
		(0.012)	(0.016)	(0.014)	(0.130)	(0.039)	(0.137)
	Median Income	0.001	0.001	0.001	−0.551	0.000	0.004
		(0.000)	(0.000)	(0.000)	(0.141)	(0.000)	(0.002)
Bank	Resell Rate	−0.003	−0.080	−0.391	−0.163	−0.002	−0.338
		(0.038)	(0.055)	(0.033)	(0.115)	(0.074)	(0.217)
	Reject Rate	0.750	0.757	0.069	−0.350	0.676	0.354
		(0.194)	(0.180)	(0.147)	(0.145)	(0.346)	(0.949)
	log(Bank Assets)	−0.017	−0.022	−0.031	0.648	−0.014	−0.046
		(0.007)	(0.008)	(0.010)	(1.256)	(0.010)	(0.064)
Lien	Subordinate	1.393	1.419	0.111	−0.315	1.288	0.276
		(0.056)	(0.057)	(0.025)	(0.083)	(0.192)	(0.348)
	Not Secured	0.749	0.753	−0.021	−0.315	0.351	0.991
		(0.081)	(0.087)	(0.076)	(0.083)	(0.433)	(0.991)

Continued on next page

Appendix Table A7 – continued from previous page

		OLS	IV	1st Stage	CU	Control	CU×Control
Purpose	Home Improvement	−0.010	−0.003	0.044	−0.488	−0.203	0.635
		(0.028)	(0.026)	(0.016)	(0.119)	(0.069)	(0.185)
	Refinance	−0.200	−0.188	0.067	−0.488	−0.236	0.230
		(0.012)	(0.015)	(0.008)	(0.119)	(0.024)	(0.101)
Year	Pre-2009	1.465	1.445	−0.104	−0.365	1.461	−0.147
		(0.030)	(0.030)	(0.014)	(0.136)	(0.054)	(0.214)
	Post-2011	−0.841	−0.829	0.054	−0.365	−0.860	0.119
		(0.017)	(0.018)	(0.011)	(0.136)	(0.035)	(0.119)
Age	Under 35	−0.049	−0.052	−0.013	−0.308	−0.043	−0.044
		(0.006)	(0.006)	(0.003)	(0.093)	(0.019)	(0.082)
	Over 60	0.035	0.034	−0.001	−0.308	−0.002	0.123
		(0.006)	(0.006)	(0.004)	(0.093)	(0.023)	(0.074)
Score	Below Median	0.630	0.628	−0.014	−0.225	0.676	−0.184
		(0.025)	(0.025)	(0.005)	(0.091)	(0.052)	(0.142)
	Top Quartile	−0.223	−0.223	0.001	−0.225	−0.197	−0.102
		(0.006)	(0.006)	(0.003)	(0.091)	(0.017)	(0.063)
Term	10 years	−0.288	−0.276	0.067	−0.386	−0.505	0.551
		(0.028)	(0.030)	(0.018)	(0.085)	(0.081)	(0.199)
	15 years	−0.470	−0.467	0.018	−0.386	−0.485	0.080
		(0.016)	(0.016)	(0.005)	(0.085)	(0.036)	(0.110)
	20 years	0.016	0.012	−0.014	−0.386	0.049	−0.120
		(0.017)	(0.018)	(0.009)	(0.085)	(0.040)	(0.142)
	N	767,991	767,991	1,153,367			
	Adj. R2	0.45		0.204			
	F-stat.		8,482.48				

Notes: Data are from TransUnion merged with HMDA, limited to CUs and small banks. The first three columns of this table report coefficient estimates from the following regressions investigating the CU effect on mortgage interest rates:

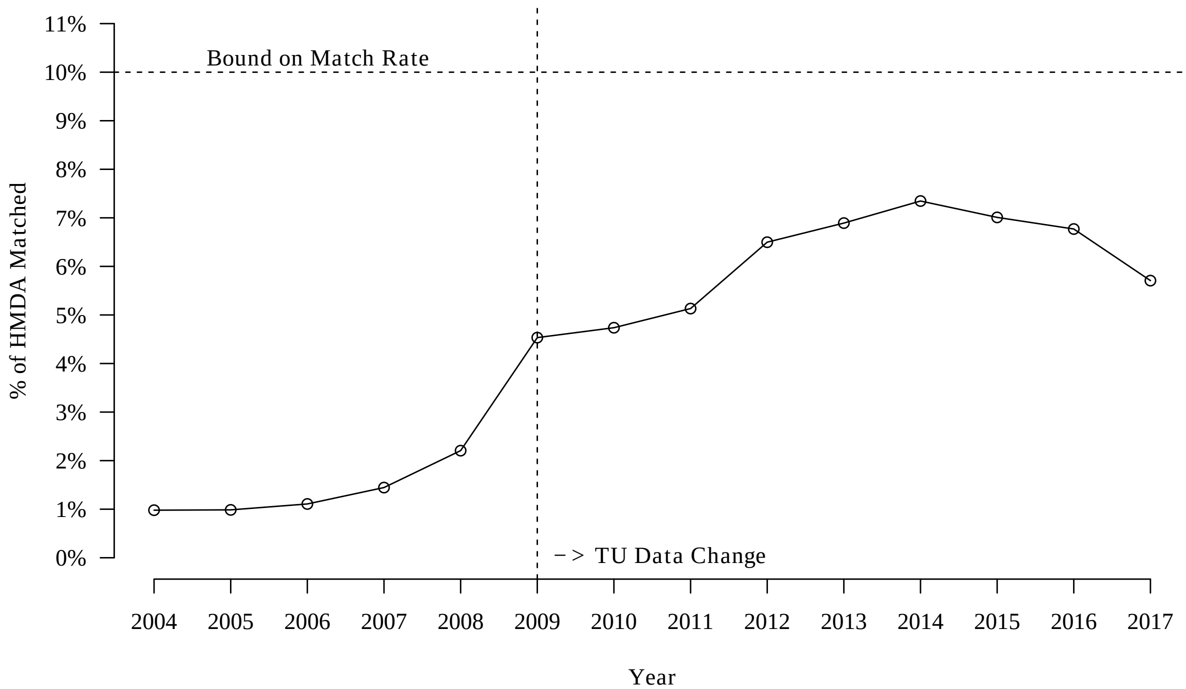
$$InterestRate_i = \beta_0 + \beta_1 CU_i + \boldsymbol{\alpha} X_i + \epsilon$$

$$InterestRate_i = \beta_0 + \beta_1 \widehat{CU_i} + \boldsymbol{\alpha} X_i + \epsilon$$

$$CU_i = \gamma_0 + \gamma_1 CuDensity_i + \boldsymbol{\delta} X_i + \varepsilon$$

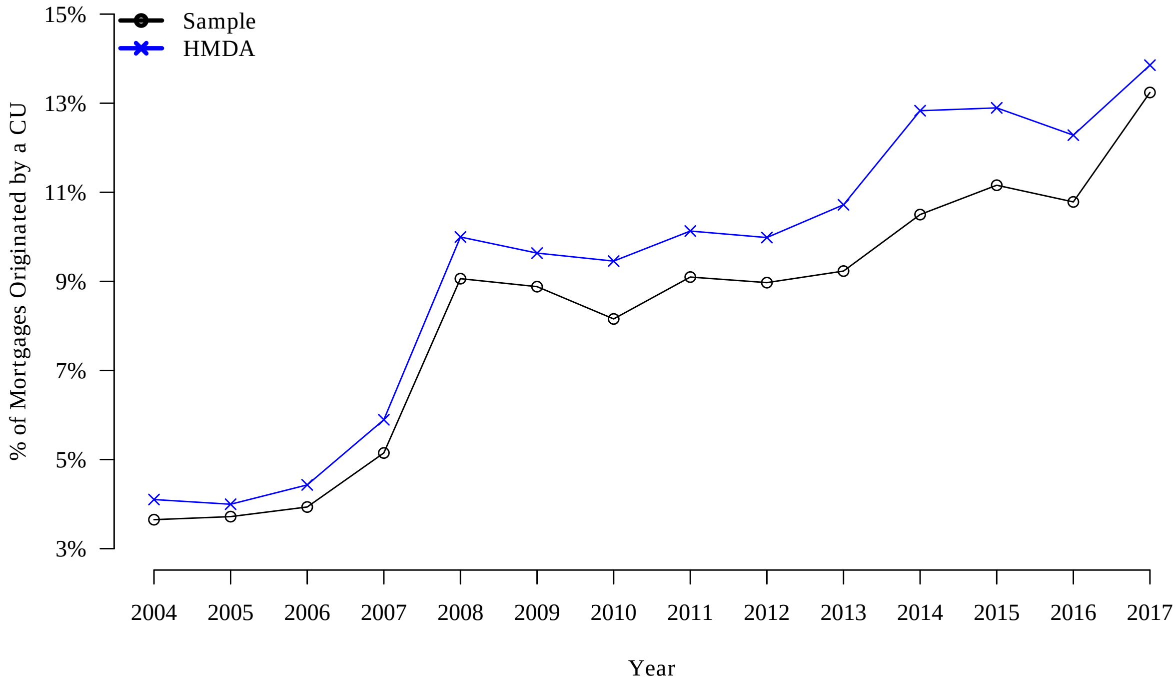
which are also detailed in equations (2)–(3) of the main text. X_i represents a vector of controls for borrower creditworthiness, loan terms, and geographic factors. The last set of columns reports regression coefficients for the equation: $InterestRate_i = \beta_0 + \beta_1 \widehat{CU_i} + \beta_2 \widehat{CU_i \times x} + \boldsymbol{\alpha} X_i + \epsilon$, also equation (8) of the main text. Each triple of coefficients in the last three columns reports the estimates β_1, β_2, and α^x for each of the regressions. For loan, tract, and bank variables, each coefficient triple is derived from a separate regression. For the rest of the groups, the coefficients come from one regression per group as they are a variable being discretized into dummies or categorical variables treated as dummies. The excluded category for each group, in descending order, is: first lien, home purchase, between 5–60, third quartile, 30 years, and 2009–2011.

Appendix Figure A1.
TU-HMDA Match Rate



Notes: Figure plots the number of mortgages in the sample of merged TransUnion and HMDA data (including non-bank originations) as a fraction of total originations in the HMDA data. The bound on the match rate is set at 10% because the credit records data are a 10% sample.

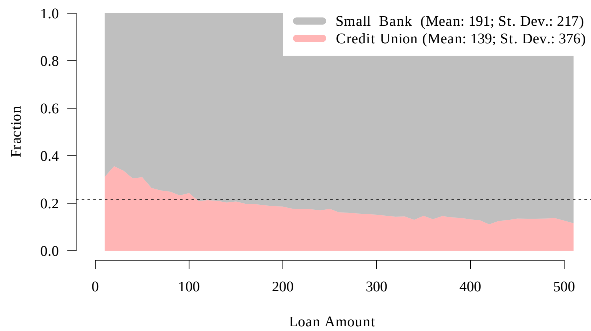
Appendix Figure A2.
CU Representation in Population vs. Matched Sample



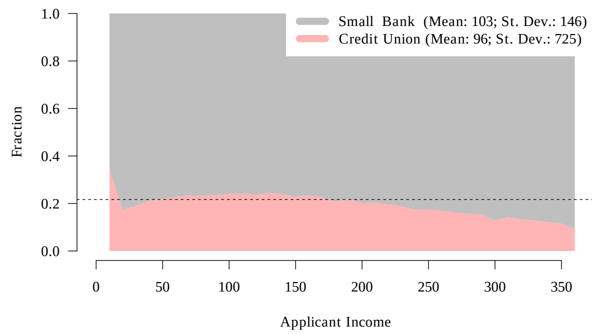
Notes: Figure plots the number of CU-originated mortgages as a fraction of total originations in two samples: the HMDA universe and in the matched sample of TransUnion and HMDA.

Appendix Figure A3.
CU Fraction of Loans Across Distribution of Loan Characteristics

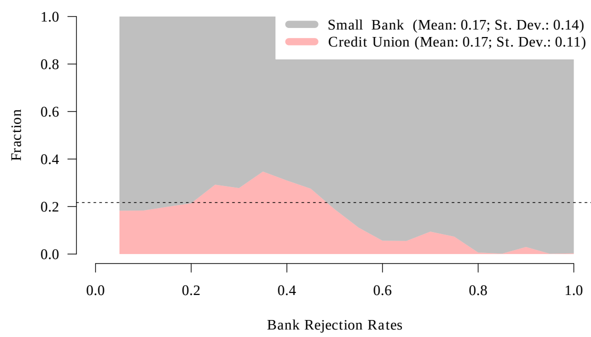
(a)



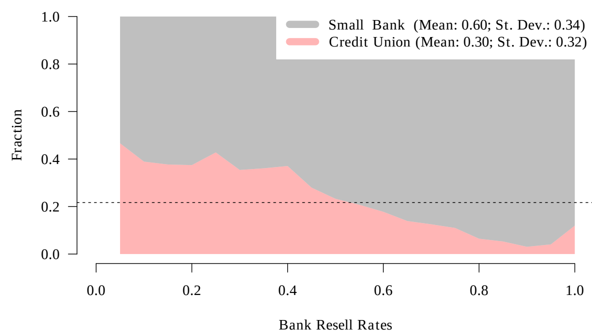
(b)



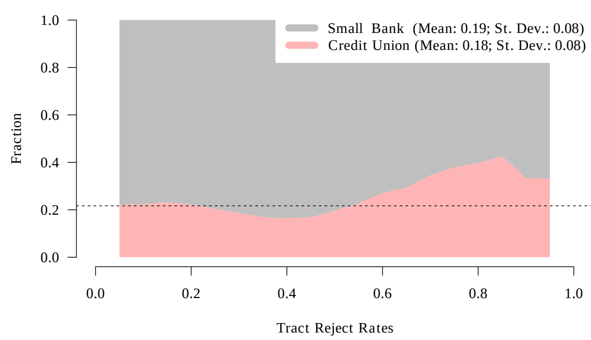
(c)



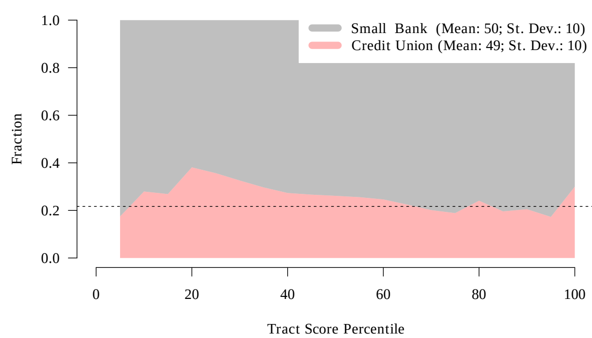
(d)



(e)

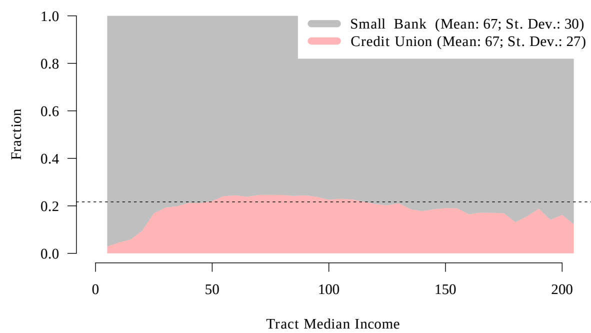


(f)

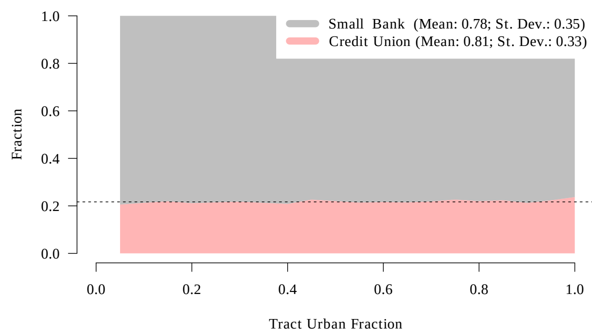


Appendix Figure A3. (continued)

(g)



(h)

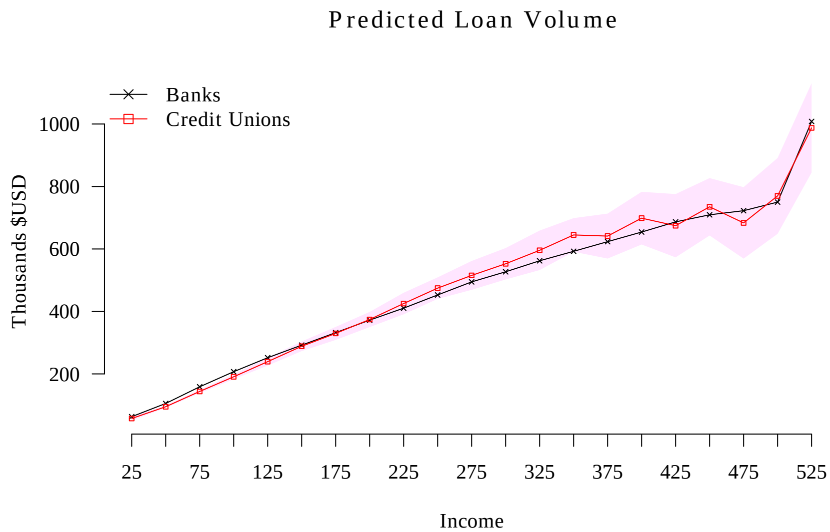


Notes: Mortgage data are from HMDA and credit records data are from TransUnion. Each plot is based on the universe of conventional (i.e., non-FHA, VA, FSA, or RHS) mortgages with a first-lien on an owner-occupied, one-to-four family property in HMDA from 2004 to 2017. A bank is defined as a small bank if its total assets reported to HMDA on the given year are less/more than 105% the total assets of the largest Credit Union in that year.

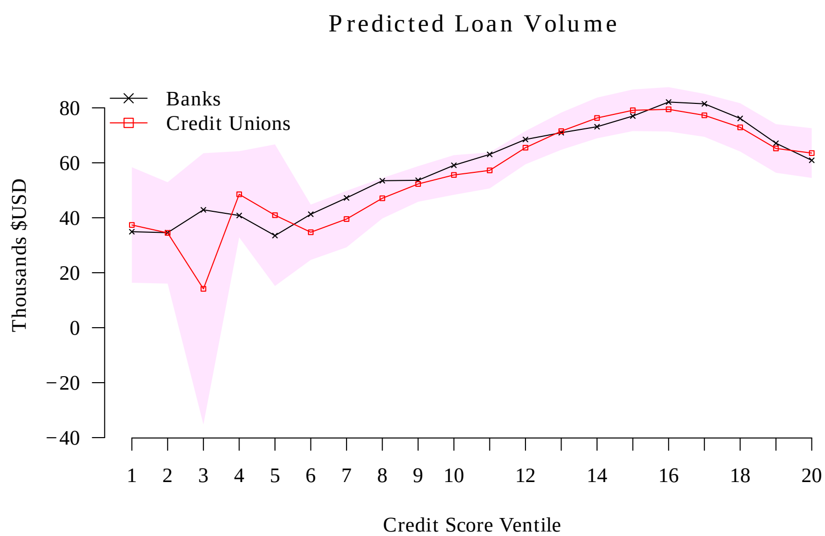
In Panels (a) and (b), amounts are presented in thousands of US Dollars. In Panel (c) bank rejection rates are calculated at the bank-year level and represent the fraction of applications that a bank reported to HMDA and were not originated due to a denial. In Panel (d) bank resell rates are calculated at the bank-year level and represent the fraction of loans that were originated by a lender and sold within the same calendar year. In Panel (e) tract reject rates are computed at the tract-year level, and fraction of applications from a given tract that were reported to HMDA and were not originated due to a denial. In Panel (f) mean tract score percentiles are calculated at the tract-year level. Tract median income in Panel (g) is in thousands of constant 2012 US Dollars and comes from Census data. Panel (f) shows the distribution of tracts' urban fractions according to Census data.

Appendix Figure A4

(a)



(b)



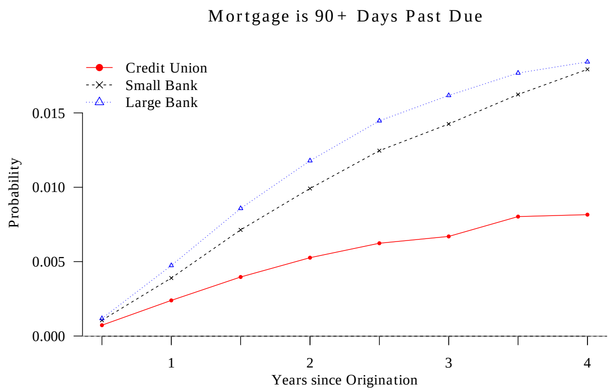
Notes: Data are from TransUnion merged with HMDA, restricted to loans for a home purchase with a first lien, comparing CUs to small banks. The figures plot the predicted loan amount based on the following regression

$$LoanAmount = CU + IncomeGroup \times CU + ScoreVentile \times CU.$$

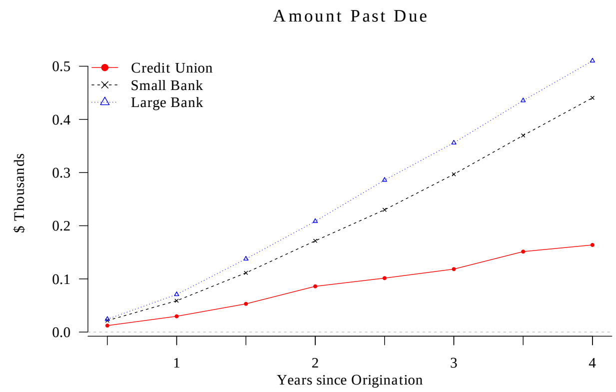
Panel (a) uses the 11 credit score ventile to set the intercept level and panel (b) uses the $75K to $100K group to set the intercept level.

Appendix Figure A5.
Mortgage Outcomes

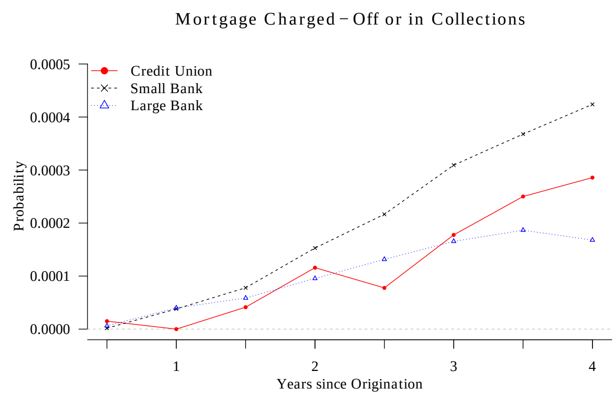
(a)



(b)

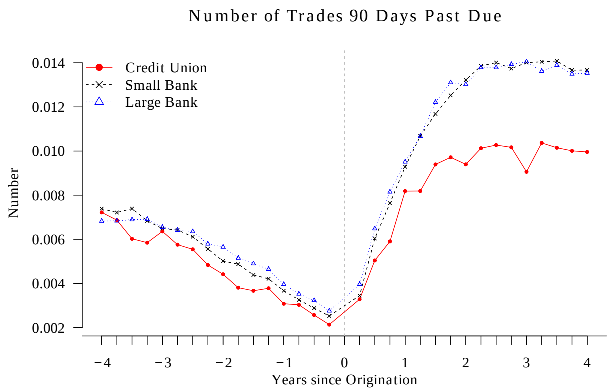


(c)

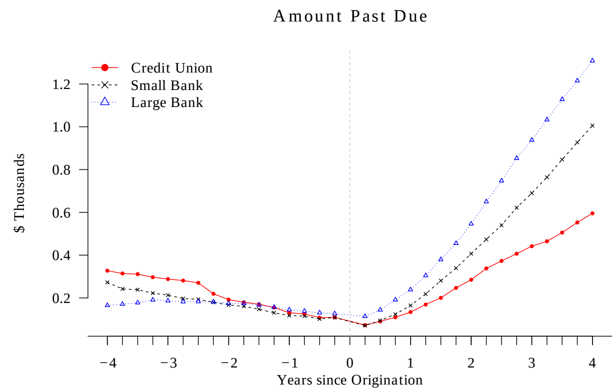


Appendix Figure A5. (continued)
Credit Profile Outcomes

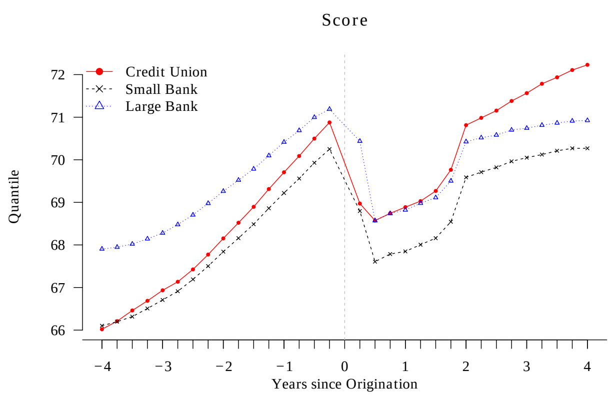
(d)



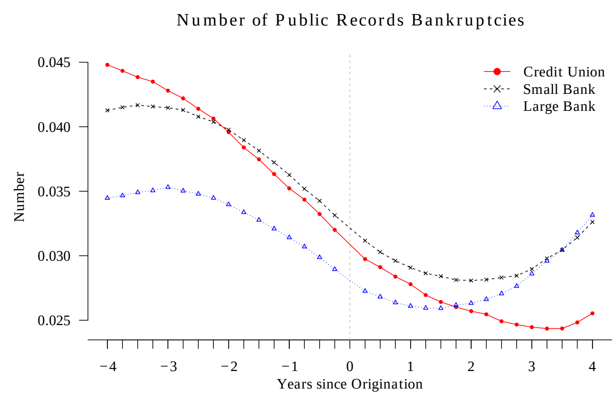
(e)



(f)

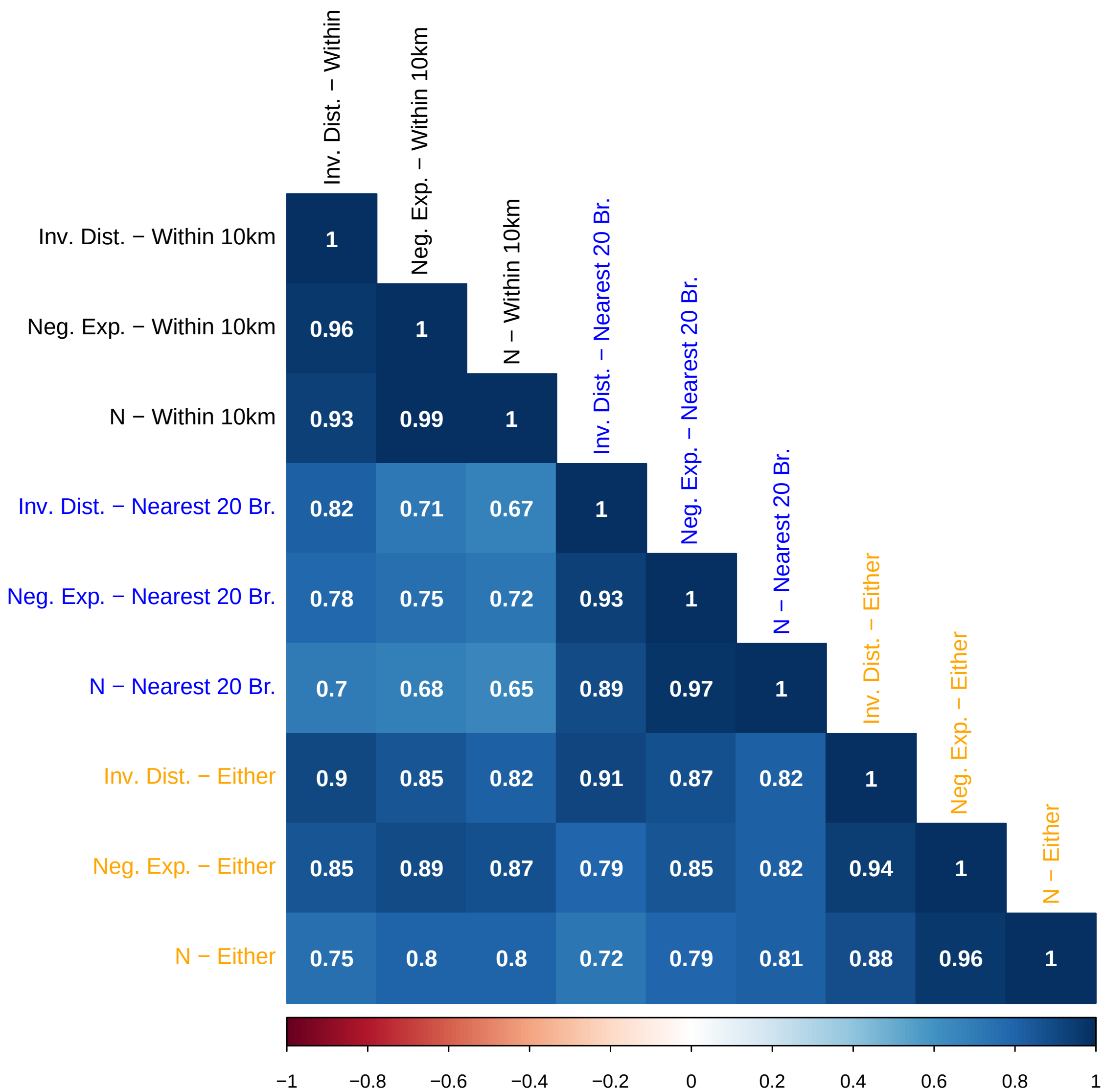


(g)



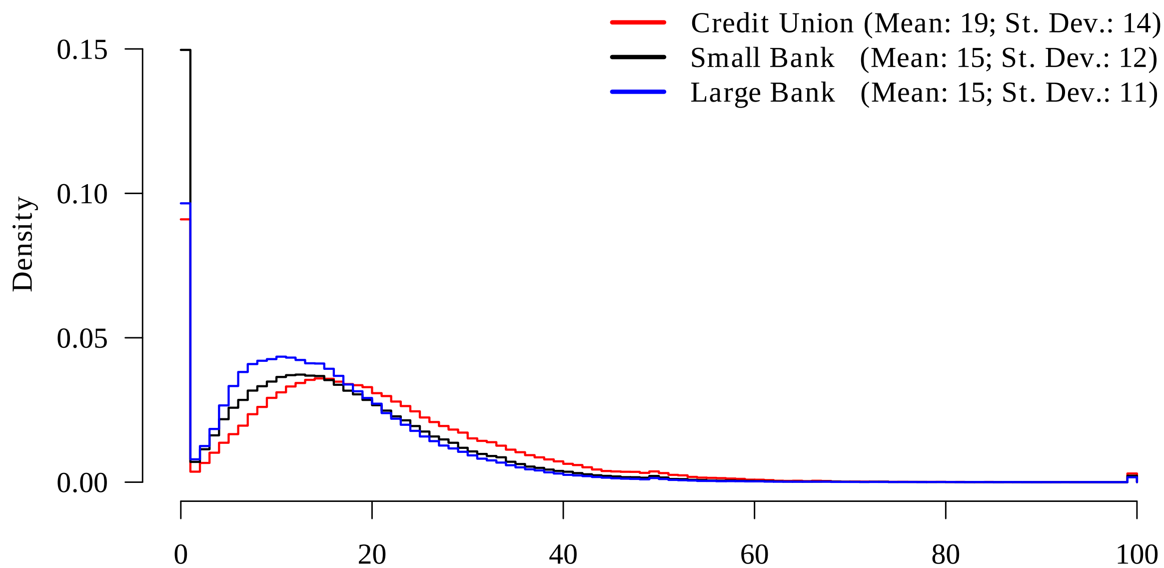
Notes: This is a more comprehensive version of Figure 3 in the paper. Data are from TransUnion merged with HMDA, restricted to 30-year term mortgages. Plots show the mean value by bank type of each variable according to event time, defined from the month of mortgage origination.

Appendix Figure A6.
Correlation Between Instrument Specifications



Notes: Data are from TransUnion merged with HMDA. The figure presents a correlation matrix of the nine instrument specifications considered. The instrument and its alternative specifications are defined in Section 5.1 of the main text.

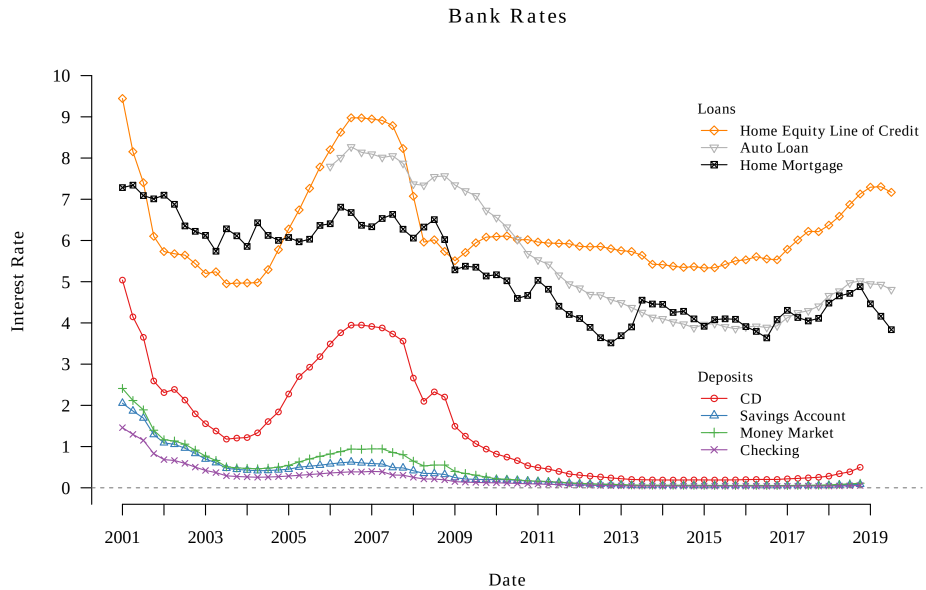
Appendix Figure A7.
Distribution of Instrument Values by Lender



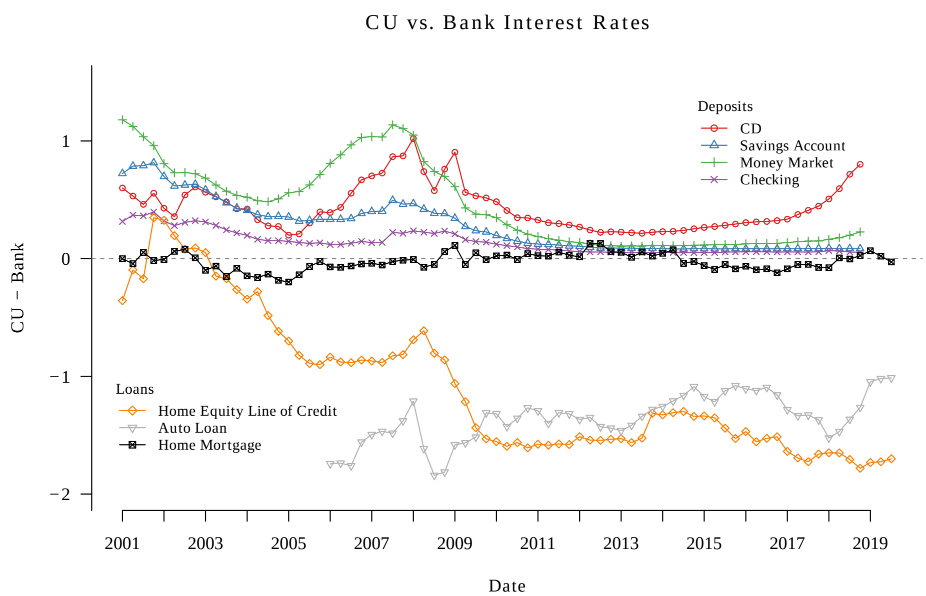
Notes: Data are from TransUnion merged with HMDA. The figure plots a histogram of the instrument values for mortgages by bank type. The instrument is defined as follows: let $d(i,l)$ be the distance from a branch of lender l to the individual i's address. The instrument is then the density of CU branches over all CU and bank branches, weighted by the inverse of their distance from the new address: $CuDensity_i = \frac{\sum_{l \in CU} \frac{1}{d(i,l)}}{\sum_{l \in CU} \frac{1}{d(i,l)} + \sum_{l \in Bank} \frac{1}{d(i,l)}}$. All three histograms use a bin width of one.

Appendix Figure A8.
List Prices at CUs and Banks

(a)

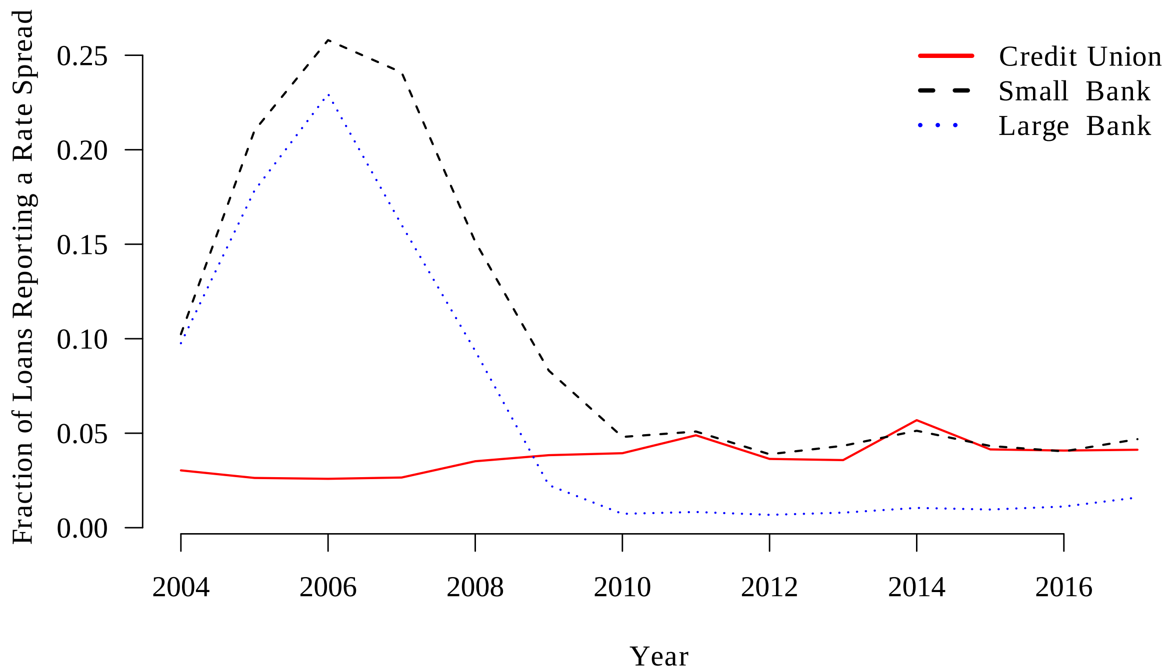


(b)



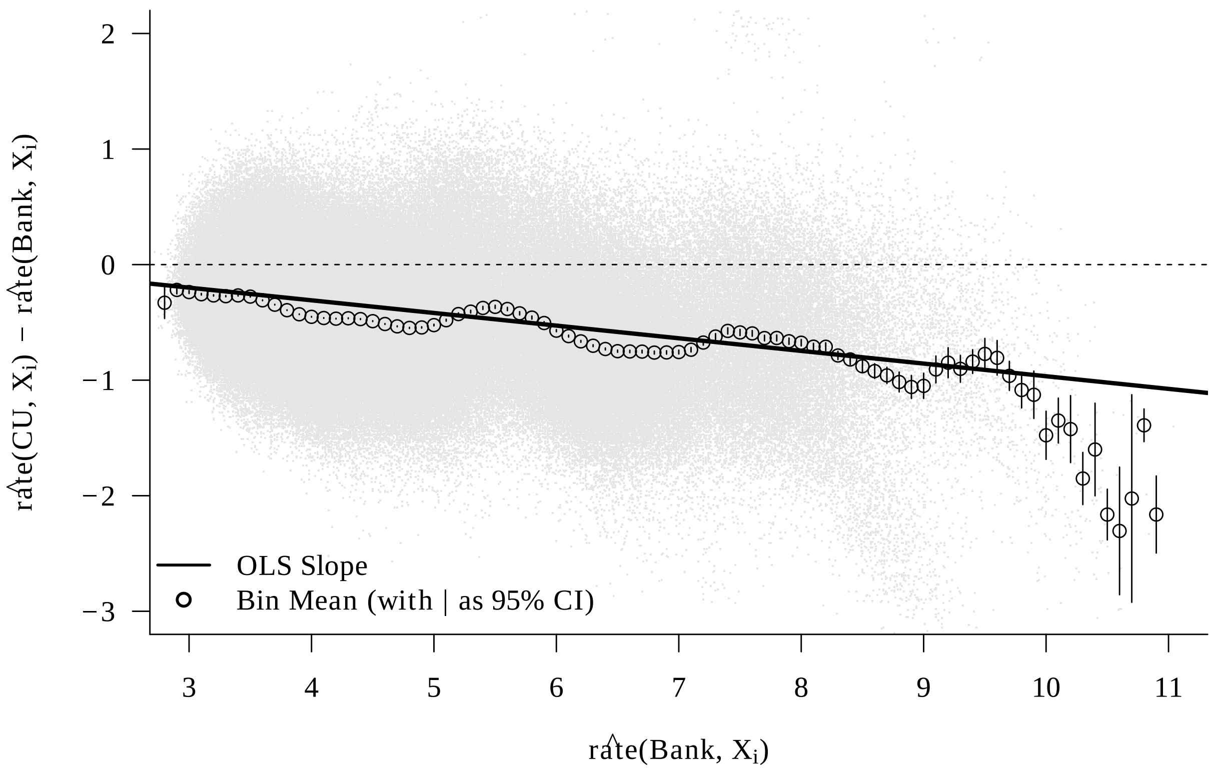
Notes: Data are from S&P RateWatch, which surveys bank and CU branches for loan and deposit prices. Panel (a) shows mean bank rates by month for various loan and deposit products. Panel (b) shows the difference in monthly mean prices for each loan or deposit product. Home Equity Line of Credit loans are tier 1 risk for loan-to-value ratio between 0-80 percent; Auto Loans are 48-month loan for a 2-year used automobile; Home Mortgage loans are 30-year fixed rate mortgage of US $175,000; CDs are 12-month Certificates of Deposit of US $10,000; Savings Accounts, Money Market accounts, and Checking Accounts that require a minimum balance of US $2,500.

Appendix Figure A9.
Loans Reporting a “High” Interest Rate in HMDA



Notes: Data are from the HMDA. The HMDA requires lenders to report a ”rate spread” on loans with an interest rate that is 3% (or 1.5%, depending on the year) points above the Treasury yield. Plot shows the fraction of loans by lender type and year that report a rate spread.

Appendix Figure A10.
Counterfactual Individual Prices at CUs vs. Banks



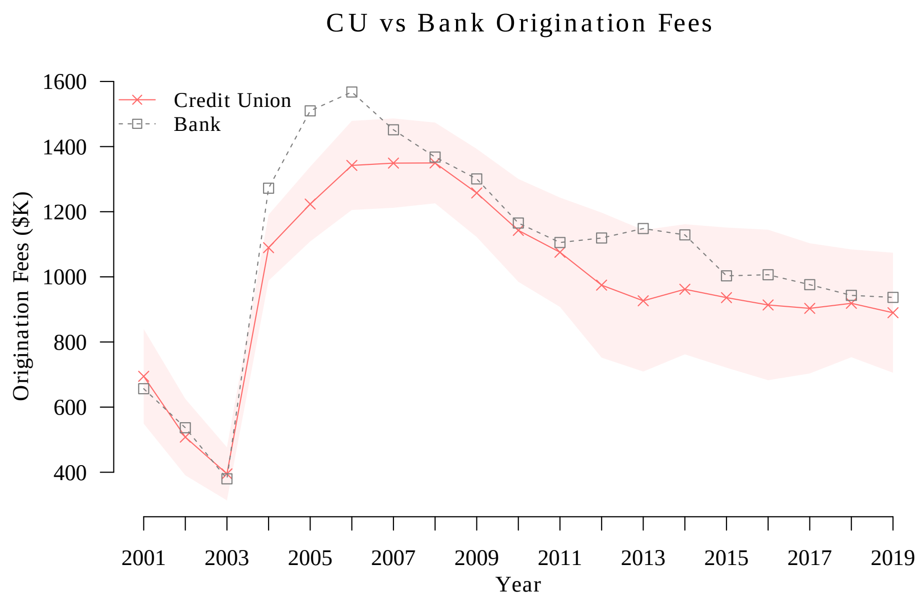
Notes: Data are from TransUnion merged with HMDA, restricted to CUs and small banks. Plots counterfactual CU and bank interest rates for all individuals, based on the model derived from the following instrumented regression:

$$InterestRate_i = \beta_0 + \beta_1 \widehat{CU}_i + \boldsymbol{\alpha} X_i + \epsilon$$
$$CU_i = \gamma_0 + \gamma_1 CuDensity_i + \boldsymbol{\delta} X_i + \varepsilon.$$

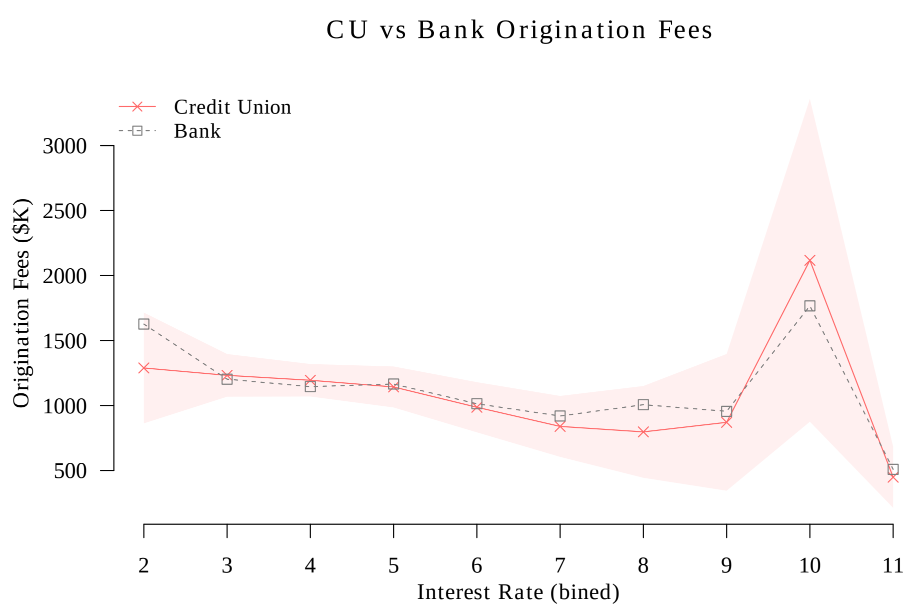
The $CUDensity$ instrument is defined as follows: let $d(i,l)$ be the distance from a branch of lender l to the individual i's address. The instrument is then the density of CU branches over all CU and bank branches, weighted by the inverse of their distance from the new address: $CuDensity_i = \frac{\sum_{l \in CU} \frac{1}{d(i,l)}}{\sum_{l \in CU} \frac{1}{d(i,l)} + \sum_{l \in Bank} \frac{1}{d(i,l)}}$.

Appendix Figure A11

(a)



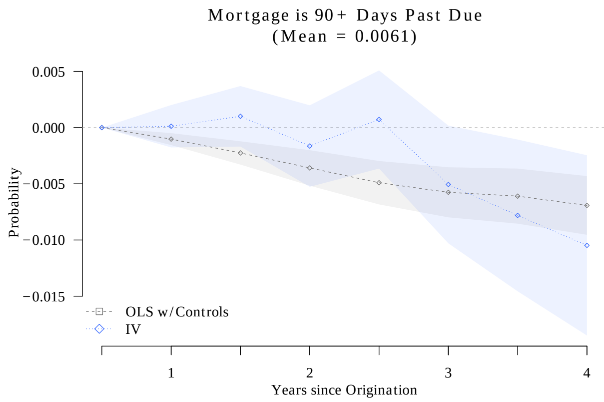
(b)



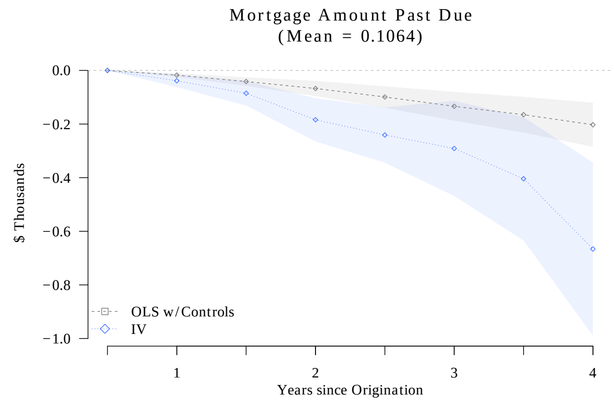
Notes: Data are from S&P Ratewatch. Panel (a) and (b) plot the coefficients from the following regression $OriginationFees = CU + Year + InterestRate + CU \times Year + CU \times InterestRate$, where CU is a dummy for whether the bank branch is a Credit Union, as opposed to a bank. Interest rates are grouped by rounding down to their nearest integer. Panel (a) shows the predicted origination fees for banks and CUs by year, and panel (b) shows the predicted origination fees by banks and CUs by interest rate level.

Appendix Figure A12.
CU Treatment Effect on Mortgage Outcomes

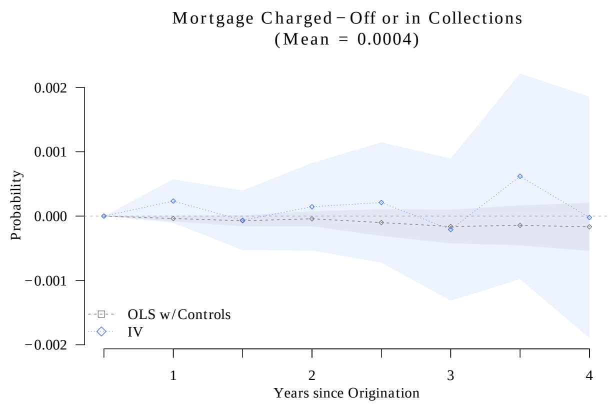
(a)



(b)

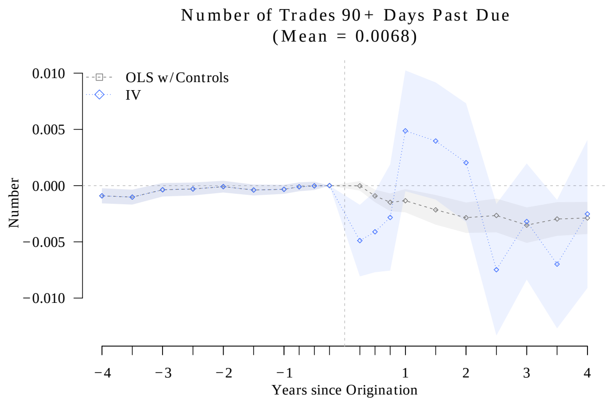


(c)

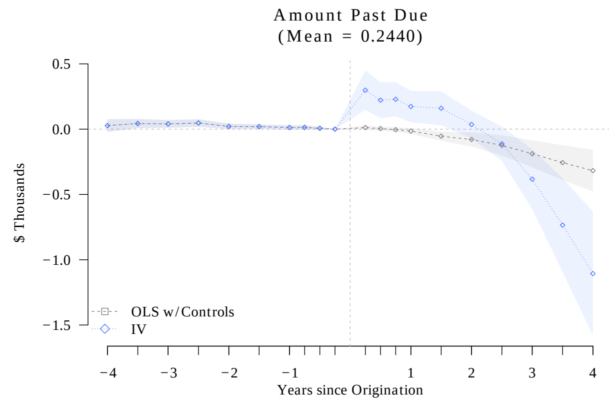


Appendix Figure A12. (continued)
CU Treatment Effect on Credit Profile Outcomes

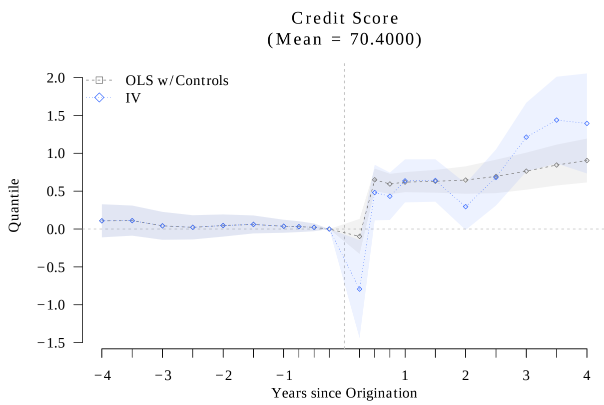
(d)



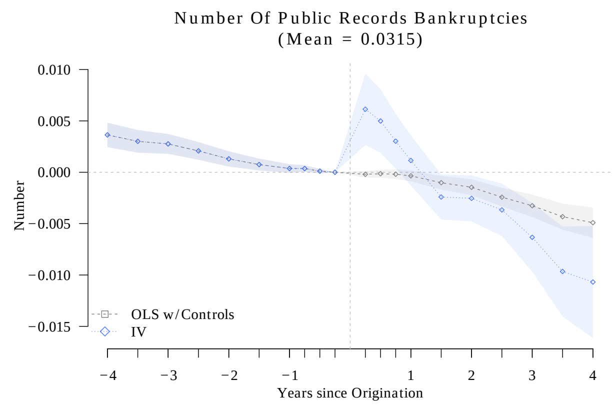
(e)



(f)



(g)



Notes: This is a more comprehensive version of Figure 6 in the main paper. Data are from TransUnion merged with HMDA. Plots show estimates of the differential effect that originating a mortgage with a CU has relative to originating a mortgage with a **small** bank. The shaded areas represent 95% confidence intervals based on standard errors clustered at the bank level. Appendix Figure A13 contains the equivalent results when the reference group is large banks. The OLS w/controls estimates report $\boldsymbol{\beta}$ estimates from the following regression:

$$Y_{it} = \boldsymbol{\beta} CU_i \psi_{\tau(t,i)} + \boldsymbol{\delta} \psi_{\tau(t,i)} + \boldsymbol{\alpha} X_i + \epsilon_{it}$$

which is also detailed equation (5) of the text. The IV estimates in blue diamonds report $\boldsymbol{\beta}$ coefficient estimates from the following instrumented event study regression:

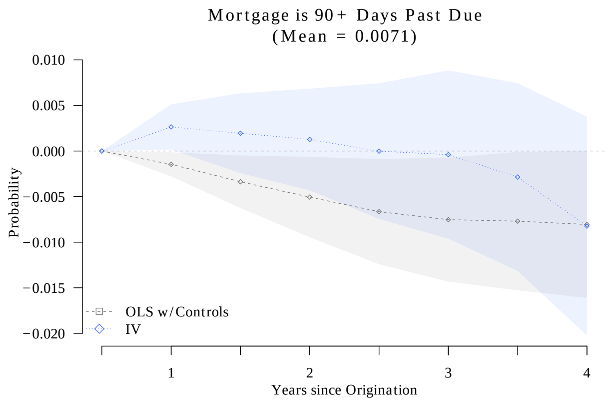
$$Y_{it} = \boldsymbol{\beta} \widehat{CU_i \psi_{\tau(t,i)}} + \boldsymbol{\delta} \psi_{\tau(t,i)} + \boldsymbol{\alpha} X_i + \epsilon_{it}$$

$$CU_{it} \psi_{\tau(t,i)} = \boldsymbol{\gamma^{post}} CuDensity_i \psi_{\tau(t,i)>0} + \boldsymbol{\gamma^{pre}} CU_i \psi_{\tau(t,i)<0} + \boldsymbol{\delta} X_i + \epsilon_{it},$$

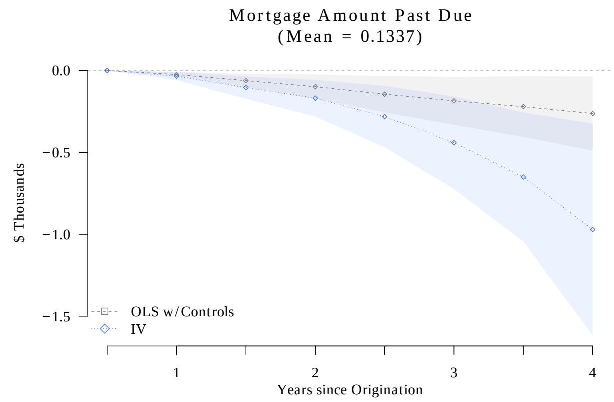
which is also detailed equations (6) and (7) of the text. $\psi_{\tau(t,i)}$ are event-time dummy variables that capture the number of months after origination that the credit outcome is being observed. X_i represents a vector of controls for borrower creditworthiness, loan terms, and geographic factors. $d(i,l)$ is the distance from branch of lender l to the individual i's address. The instrument is then the density of CU branches over all CU and bank branches, weighted by the inverse of their distance from the new address: $CuDensity_i = \frac{\sum_{l \in CU} \frac{1}{d(i,l)}}{\sum_{l \in CU} \frac{1}{d(i,l)} + \sum_{l \in Bank} \frac{1}{d(i,l)}}$. Standard errors are reported in parentheses.

Appendix Figure A13.
CU Treatment Effect on Mortgage Outcomes, to Large Banks

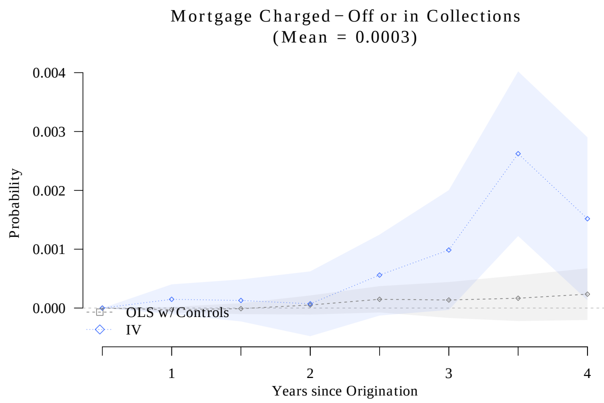
(a)



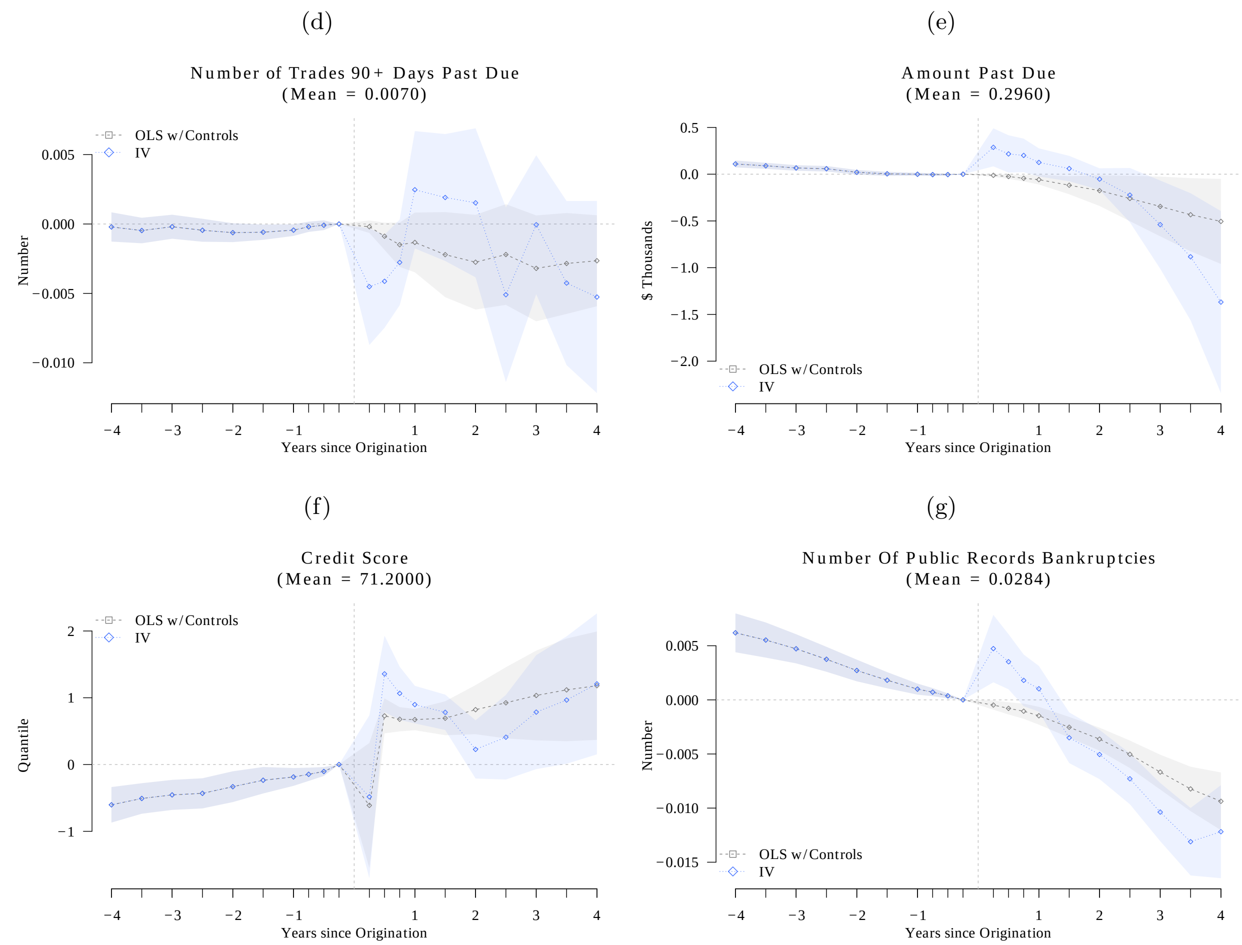
(b)



(c)



Appendix Figure A13. (continued)
CU Treatment Effect on Credit Profile, Relative to Large Banks



Notes: Data are from TransUnion merged with HMDA. Plots show estimates of the differential effect that originating a mortgage with a CU has relative to originating a mortgage with a **large** bank. The shaded areas represent 95% confidence intervals based on standard errors clustered at the bank level. Appendix Figure A12 contains the equivalent results when the reference group is small banks. The OLS w/controls estimates report $\boldsymbol{\beta}$ estimates from the following regression:

$$Y_{it} = \boldsymbol{\beta} CU_i \psi_{\tau(t,i)} + \boldsymbol{\delta} \psi_{\tau(t,i)} + \boldsymbol{\alpha} X_i + \epsilon_{it}$$

which is also detailed equation (5) of the text. The IV estimates in blue diamonds report $\boldsymbol{\beta}$ coefficient estimates from the following instrumented event study regression:

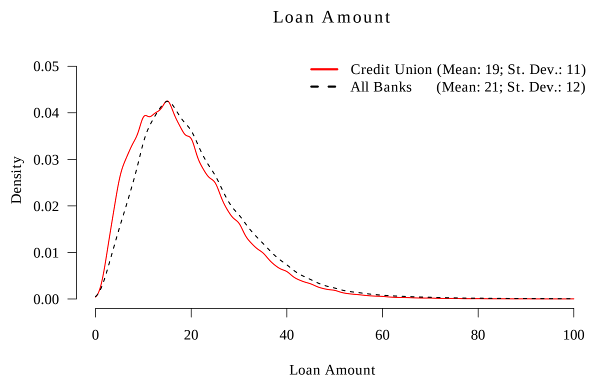
$$Y_{it} = \boldsymbol{\beta} \widehat{CU_i \psi_{\tau(t,i)}} + \boldsymbol{\delta} \psi_{\tau(t,i)} + \boldsymbol{\alpha} X_i + \epsilon_{it}$$

$$CU_{it} \psi_{\tau(t,i)} = \boldsymbol{\gamma^{post}} CuDensity_i \psi_{\tau(t,i)>0} + \boldsymbol{\gamma^{pre}} CU_i \psi_{\tau(t,i)<0} + \boldsymbol{\delta} X_i + \epsilon_{it},$$

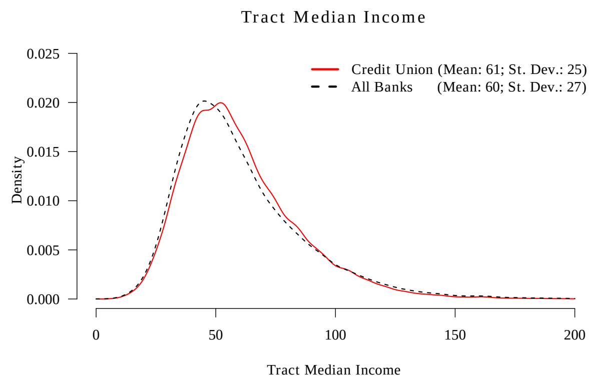
which is also detailed equations (6) and (7) of the text. $\psi_{\tau(t,i)}$ are event-time dummy variables that capture the number of months after origination that the credit outcome is being observed. X_i represents a vector of controls for borrower creditworthiness, loan terms, and geographic factors. $d(i,l)$ is the distance from branch of lender l to the individual i's address. The instrument is then the density of CU branches over all CU and bank branches, weighted by the inverse of their distance from the new address: $CuDensity_i = \frac{\sum_{l \in CU} \frac{1}{d(i,l)}}{\sum_{l \in CU} \frac{1}{d(i,l)} + \sum_{l \in Bank} \frac{1}{d(i,l)}}$. Standard errors are reported in parentheses.

Appendix Figure A14.
Characteristics of Auto Loans

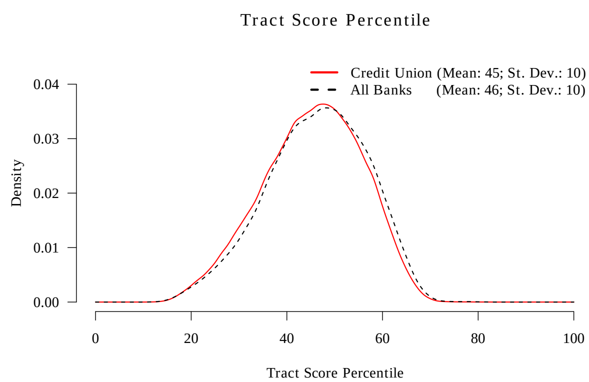
(a)



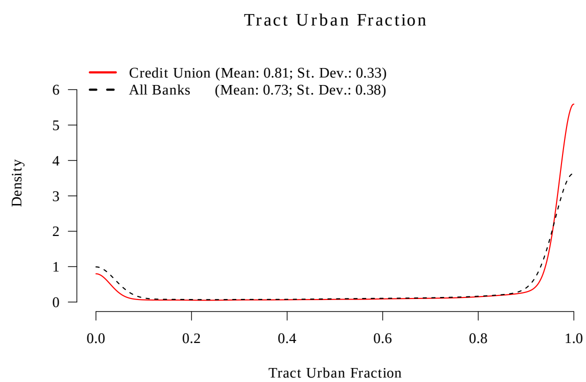
(b)



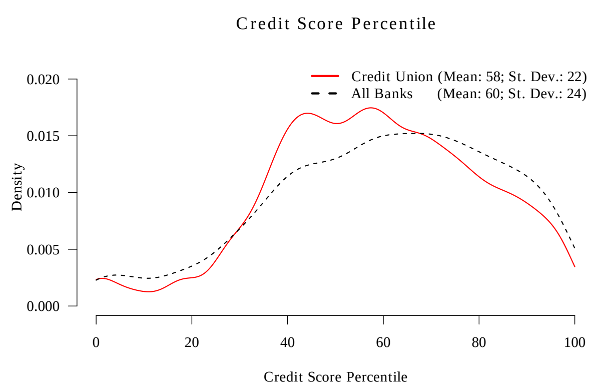
(c)



(d)

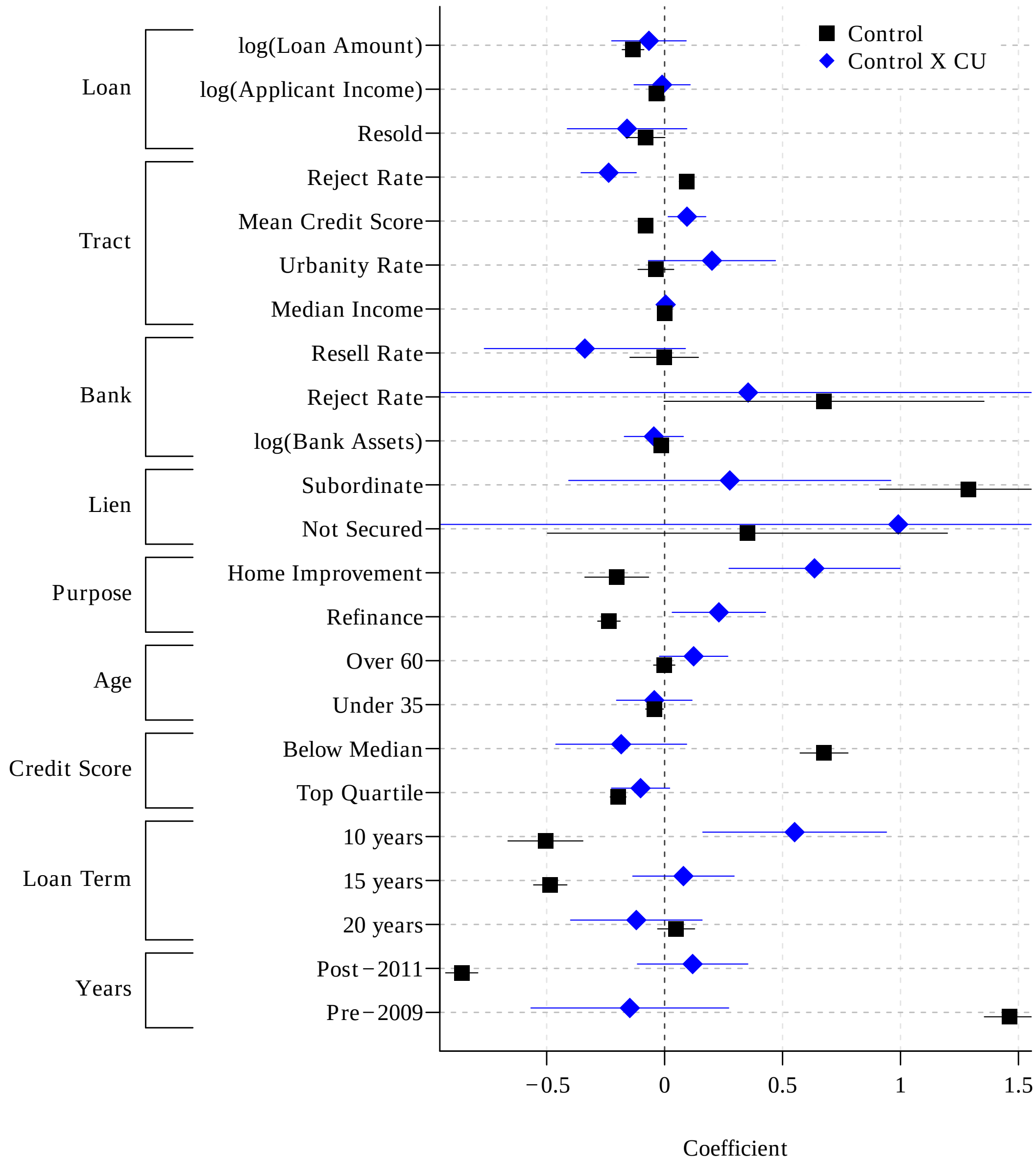


(e)



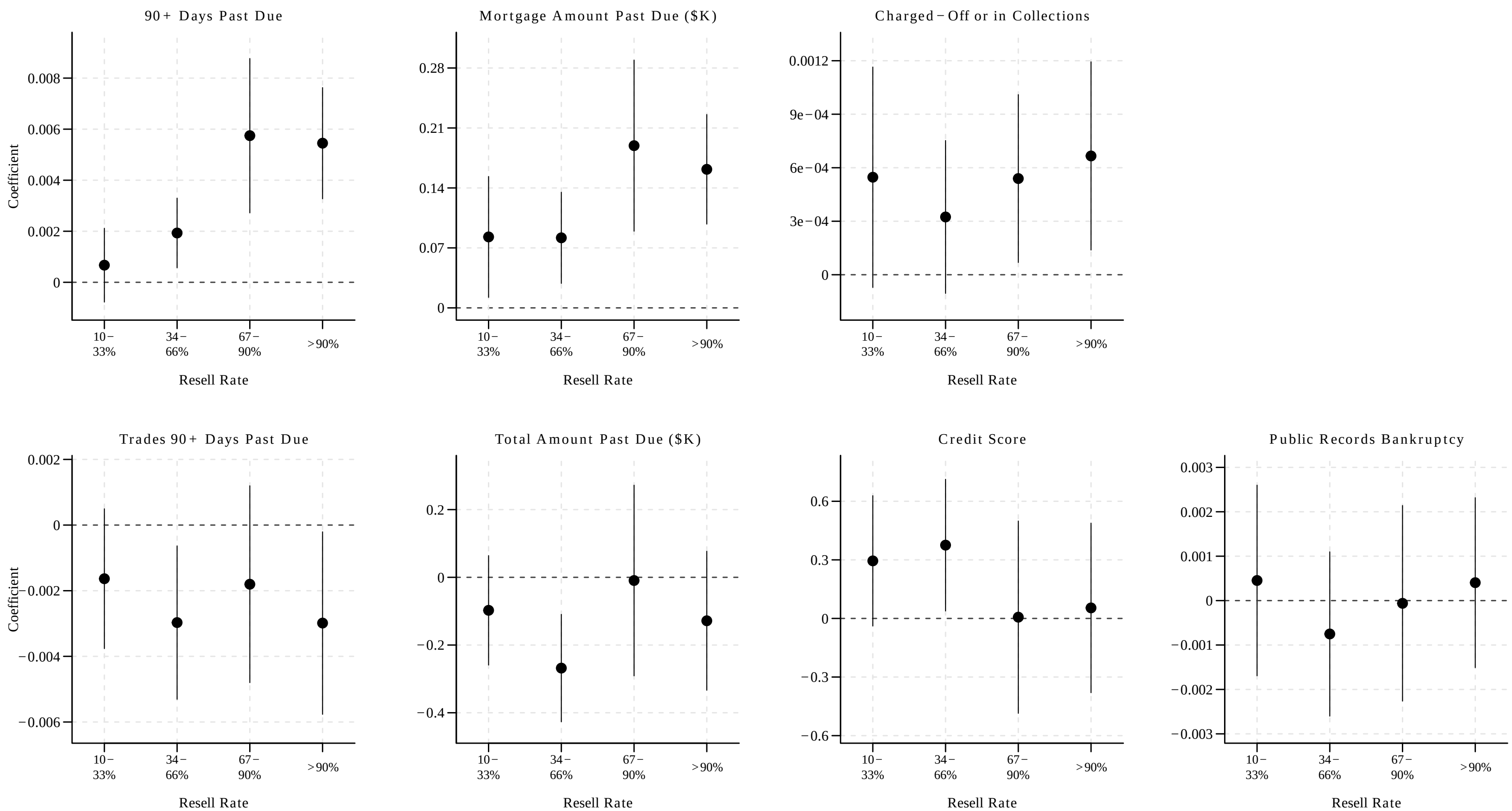
Notes: Data are from TransUnion and Census. Each subfigure plots kernel density estimates based on the originated by banks and CUs 2004 to 2017 in the TransUnion sample. Densities of a given variable are weighted by the number of originated loans. Panels (a) and (b) show amounts in thousands of US Dollars. Panel (b) is in constant 2012 US dollars and is derived from Census data. Panel (c)'s mean tract score percentiles are calculated at the tract-year level. Panel (d)'s tract urban fractions are derived from Census data. Panel (e) shows borrower credit score percentile at origination.

Appendix Figure A15.
Differentialy Priced Factors at CUs vs. Banks



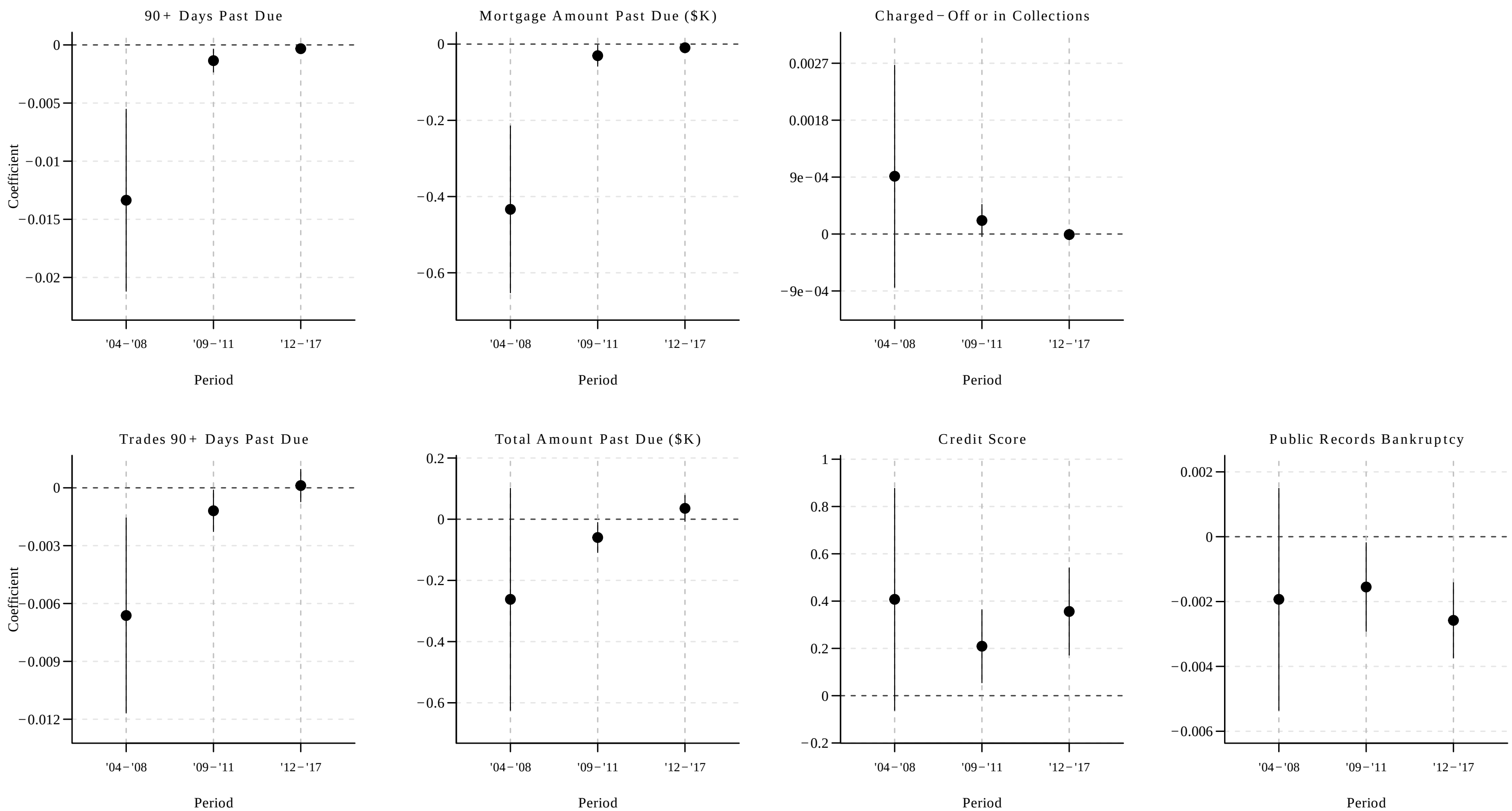
Notes: Data are from TransUnion merged with HMDA, restricted to CUs and small banks. Plots show regression coefficients for the equation: $InterestRate_i = \beta_0 + \beta_1 \widehat{CU}_i + \beta_2 \widehat{CU_i \times x} + \boldsymbol{\alpha} X_i + \epsilon$, also equation (8) of the main text. Plots show estimated coefficients of β_1 and β_2 defined in equation (8), in black squares and blue diamonds, respectively. The β_1 coefficient captures the coefficient of a given control variable on interest rates, and the β_2 coefficient captures the additive effect of the given variable on interest rates for CUs. The horizontal lines around squares and diamonds represent 95% confidence intervals based on standard errors clustered at the bank level. For loan, tract, and bank variables, each coefficient pair is derived from a separate regression. For the rest of the groups, the coefficients come from one regression per group because they are a variable being discretized into dummies or categorical variables treated as dummies. The excluded category for each group, in descending order, is: first lien, home purchase, between 35 and 60, third quartile, 30 years, and 2009–2011. Appendix Table A7 reports the coefficients for these excluded base categories in column “CU.”

Appendix Figure A16.
Heterogeneity in Credit Outcomes by Secondary Market Resell Rates



Notes: Data are from TransUnion merged with HMDA, for all banks. Plots show estimated coefficients of β from the following regressions: $Y_{it} = \boldsymbol{\beta} Resell_i + \boldsymbol{\alpha} X_i + \epsilon_{it}$ for the top panels, and $Y_{it} = \boldsymbol{\alpha} Resell_i + Post_{it} + \boldsymbol{\beta} Resell_i \times Post_{it} + \boldsymbol{\alpha} X_i + \epsilon_{it}$ for the bottom panels. The resell variables is discretized into 5 binary variables based on the percentage of loans they resell within the same calendar year: $Resell := \{< 10, 10-33, 34-66, 67-90, > 90\}$. I exclude the first category from the regression. The coefficients capture the association with changes in credit outcomes for CUs that have high, mid, and low rates of mortgage reselling. The horizontal lines around the symbols represent 95% confidence intervals based on standard errors clustered at the bank level.

Appendix Figure A17.
Heterogeneity in Credit Outcome Effects by Time Period



Notes: Data are from TransUnion merged with HMDA, restricted to CUs and small banks. Black squares represent estimates of β three years after origination from equation (5). An independent regression is run on three mutually exclusive data subsamples, one for each time period, and extracting the coefficient β corresponding to three years post origination. The horizontal lines around the symbols represent 95% confidence intervals based on standard errors clustered at the bank level.

B Backing Out Implied Interest Rates from Credit Records

Section 5.2.1 of the main text of the paper explains the identificacion strategy for interest rates from mortgages payment data. In this Appendix section, I detail the empirical algorithm I use to ensure the accuracy of the inferred interest rates. For each mortgage l, the calculation proceeds as follows:

1. **Restrict the data to observations between months 2 and 7.**

   This restriction serves two purposes: it reduces the probability that the data are affected by delayed payments and it restricts attention to the period during which even adjustable rate mortgages often have fixed interest rates.

2. **Solve the system of two equations and two unknown values for each month pair.**

   Let $\tau := (t, t')$ be a pair of months such that $t, t' \in \{2, 3, 4, 5, 6, 7\}$ and $t \neq t'$. Let Ψ be the set of all 15 possible τ. For each $\tau \in \Psi$, estimate the solution to the system and denote it by

   $$(IntRate_l^{\tau}, Payment_l^{\tau}) .$$

   To estimate the solution, I used the Newton-Rhapson method and set an interest rate of 3.5% and payment of $600 as starting values and limited the search to strictly positive solutions.

3. **Exclude implied interest rates for month pairs with unlikely values.**

   From the set of interest rates $IntRate_l^{\tau} : \tau \in \Psi$, discard $IntRate_l^{\tau}$ if it is greater than 15% or less than 1%.

4. **Exclude implied interest rates for mortgages with inconsistent rates across month pairs.**

   To increase confidence in the estimation, I discard mortgages for which the coefficient of variation of implied interest rates is greater than 0.4; that is, I consider calculating an interest rate impossible if:

$$\frac{StDev_{\Psi}\left(IntRate_l^{\tau}\right)}{Mean_{\Psi}\left(IntRate_l^{\tau}\right)} < 0.4.$$

5. **Take the mean of implied interest rates across all month pairs.**

   For the l not discarded in step 4, the estimated implied interest rate is the mean of $IntRate_l^{\tau}$ across all remaining $\tau \in \Psi$.

C Model of Bounded Search and the Probability of CU Choice

This section builds upon Section 5.1.1 of the main paper and rationalizes the monotonicity assumption behind the *CUDensity* instrument. The model is driven by two assumptions: (1) The probability that an individual quotes a lender is inversely related to the distance between them, and (2) borrowers sample a limited number of lenders for quotes. Both of these assumptions have empirical validation. Approximately 60% of mortgage borrowers conduct business in person, and the median distance between lender and borrower was seven miles in 2003 (Federal Reserve Board, 2008b).[7] Mortgage borrowers typically only contact one or two lenders, and generally not more than four or five (Woodward and Hall, 2012; Federal Reserve Board, 2008a; Lacko and Pappalardo, 2007).

For any given individual i, assume a total of B bank lenders and a total of C of CU lenders that operate within an arbitrary geographic radius around i. I index either lender type by l and $L = B + C$.[8] An individual i chooses to gather an exogenous number $n < L$ of mortgage quotes from among L. Let $q(i,l)$ be an indicator function for whether i quotes from l, and let $d(i,l)$ denote the physical distance between them. The probability that i quotes a lender is proportional to the inverse distance between them:

$$\mathbf{Pr}\left[q(i,l) = 1\right] \propto d(i,l)^{-1}.$$

Treating CUs and banks each as groups, the probability that i quotes a CU is equal to the fraction of CU branches weighted by their inverse distance:

$$\rho := \mathbf{Pr}\left[l \in C \mid q(i,l) = 1\right] = \frac{\sum_{l \in C} \frac{1}{d(i,l)}}{\sum_{l \in C} \frac{1}{d(i,l)} + \sum_{l \in B} \frac{1}{d(i,l)}}.$$

Let $c \leq n$ denote the number of CUs that i quotes. The probability that i quotes any given number

[7]Although technological changes likely have decreased the magnitude of the effect of physical distance on lending relationships, research finds that even as late as 2017, physical proximity is positively associated with higher mortgage lending volumes (Rehbein and Rother, 2020).

[8]I abuse notation and use B, C, and L to denote the set of lenders as well as the size of each set.

c of CUs is given by the binomial distribution:

$$\mathbf{Pr}\left[c|n\right] = \frac{n!}{c!(n-c)!}\rho^c(1-\rho)^{n-c}.$$

Using the mean of the binomial distribution, the expected number of CUs that get quoted as a fraction of all quotes is

$$\frac{\mathbb{E}\left[c|n\right]}{n} = \rho.$$

This equation clarifies the economic meaning of the *CUDensity* instrument in the context of this model. The instrument can be interpreted as the fraction of i's quotes that are from a CU, in expectation. To microfound its empirical relationship to CU choice, therefore, I analyze how the fraction of CU quotes can be related to CU choice. I consider three ways in which i chooses a lender from among those quoted.

- **Choice is uniformly random among quotes**

  Assume an individual chooses from among the quoted lenders randomly. That is, the probability of choosing a CU from among the quoted lenders is the fraction of CUs among quoted lenders: $\frac{c}{n}$. Then, prior to sampling quotes, the overall probability that i chooses a CU is equal to the instrument value:

  $$\mathbf{Pr}\left[i \longrightarrow C|n, B, C\right] = \rho.$$

- **Choice is the most preferred quote**

  Assume an individual chooses the quote that they most prefer. Further assume CUs are systematically, but probabilistically, more (or less) likely than banks to provide the best quote. More specifically, the borrower chooses a CU with probability $\frac{c}{\pi n}$, for some factor π. Prior to sampling, the probability that i chooses a CU is equal to the instrument value divided by π:

  $$\mathbf{Pr}\left[i \longrightarrow C|n, B, C\right] = \frac{\rho}{\pi}.$$

- **Choice is uniformly random, conditional on membership eligibility**

  Assume that, after receiving quotes from lenders, i learns they are unable to join a fraction μ of CUs because of membership restrictions and is otherwise indifferent between lenders. More specifically, the borrower chooses a CU from among those quoted with probability $\frac{(1-\mu)c}{n}$. Prior to sampling, the probability that i chooses a CU is equal to the instrument value multiplied by $1-\mu$:

$$\mathbf{Pr}\left[i \longrightarrow C|n, B, C\right] = (1-\mu)\rho.$$

Although these different scenarios above all rely on specific functional form assumptions, they are meant to illustrate that, even in the presence of frictions, an expected relationship between the instrument and CU choice exists. As in Section 5.1.1, let the probability that a borrower samples a CU be given by:

$$\rho = \frac{\sum_{l \in C} \frac{1}{d(i,l)}}{\sum_{l \in C} \frac{1}{d(i,l)} + \sum_{l \in B} \frac{1}{d(i,l)}}.$$

Further, the probability that a borrower chooses a CU conditional on it being in among the quoted lenders is given by $\frac{c}{n}\delta$, where δ is a multiplicative adjustment factor.

I derive the unconditional probability that a borrower chooses a CU as a function of n, B, C, let $c = y+1$ and $n = m+1$, then

$$\begin{aligned}
\mathbf{Pr}\left[i \longrightarrow c|n, B, C\right] &= \sum_{c=1}^{n} \frac{n!}{(c-1)!(n-c)!}\rho^{c}(1-\rho)^{n-c}\frac{c}{n}\delta. \\
&= \frac{\delta}{n}\sum_{y=0}^{m} \frac{(m+1)!}{y!(m-y)!}\rho^{y+1}(1-\rho)^{m-y}. \\
&= \frac{\delta}{n}(m+1)\rho\sum_{y=0}^{m} \frac{m!}{y!(m-y)!}\rho^{y}(1-\rho)^{m-y}. \\
&= \delta\rho.
\end{aligned}$$

The last equality follows from the binomial theorem. By setting $\delta = 1$, $\delta = \frac{1}{\pi}$, and $\delta = (1-\mu)$, we obtain the results in the three choice regimes outlined at the end of Section 5.1.1.