COMPETITION AND SELECTION IN CREDIT MARKETS

Constantine Yannelis
Anthony Lee Zhang

Competition and Selection in Credit Markets
Constantine Yannelis and Anthony Lee Zhang
NBER Working Paper No. 29169
August 2021
JEL No. D14,D4,G20,G21,G5,L62

ABSTRACT

We present both theory and evidence that increased competition may decrease rather than increase consumer welfare in subprime credit markets. We present a model of lending markets with imperfect competition, adverse selection and costly lender screening. In more competitive markets, lenders have lower market shares, and thus lower incentives to monitor borrowers. Thus, when markets are competitive, all lenders face a riskier pool of borrowers, which can lead interest rates to be higher, and consumer welfare to be lower. We provide evidence for the model's predictions in the auto loan market using administrative credit panel data.

Constantine Yannelis
Booth School of Business
University of Chicago
5807 S. Woodlawn Avenue
Chicago, IL 60637
and NBER
constantine.yannelis@chicagobooth.edu

Anthony Lee Zhang
University of Chicago
Booth School of Business
5807 S Woodlawn Ave
Chicago, Illi 60637
anthony.zhang@chicagobooth.edu

1 Introduction

How do the effects of market concentration interact with lender screening in credit markets? The efficiency of lending markets can be hampered by information imperfections (Akerlof, 1970; Stiglitz and Weiss, 1981), but these harmful effects can be in part mitigated by imperfect competition (Petersen and Rajan, 1995; Mahoney and Weyl, 2017). We propose and test a new channel through which competition can have adverse effects on consumer credit markets.

There is a theoretical reason to believe that credit market competition can harm consumers in high-risk market segments. Lenders can invest in a fixed-cost screening technology, which screens out consumers who are likely to default, allowing lenders to charge lower interest rates to the remaining consumers.[1] Lenders in concentrated markets have higher incentives to invest in screening, since their fixed costs are divided among a larger customer base. As a result, when market competition increases, lenders have lower incentives to invest in screening. The population of borrowers becomes riskier, and interest rates can actually increase, leaving consumers worse off.

Our framework makes a simple and surprising empirical prediction. In low-risk market segments, loan rates should be positively associated with market concentration, as predicted by classical theory. In high-risk segments, where screening is more important, loan rates should actually be negatively associated with concentration measures. We test this prediction using administrative credit panel data from TransUnion. We focus on auto lending, a rapidly growing consumer lending sector with $1.4 trillion in outstanding loan volume. Auto lending is characterized by direct lending to consumers from banks and dealers, segmentation according to consumers' credit risk, and the absence of government subsidies and guarantees. All of these features make the auto lending market ideal to explore the predictions of the theoretical model. We build a nationally representative dataset, at the county by year level, split by VantageScore credit score bins. The model's predictions hold in the data: concentration is positively associated with interest rates for prime borrowers, and negatively associated with rates for subprime borrowers.

---

[1]Anecdotal evidence suggests that costly mechanisms which lenders use to screen borrowers are increasingly common. For example, lenders can invest in better predictive analytics, for example using machine learning or artificial intelligence. Lenders may also purchase data which can predict default. Some auto lenders also invest in physical technology, such as GPS technology to track cars in the event of a repossession.

We analyze a simple model in which lenders invest in a technology to screen out consumers likely to default, then compete by setting interest rates. In consumer lending settings, many screening technologies are characterized by large upfront costs and low variable costs: lenders often acquire alternative data on consumers, and hire employees to build default prediction models and loan decision software. To model this, we assume that the screening technology is fixed-cost: the lender's cost depends on the desired default rate, but not on the number of loans made. This implies that lenders who make more loans have larger incentives to invest in screening, since the cost of screening is distributed over a larger consumer base. We assume lenders are differentiated, so lenders are able to charge markups over their marginal costs. In equilibrium, consumers' interest rates are equal to the break-even interest rate, given consumers' default rates, plus a markup which depends on lenders' market power.

Competition has two opposing effects on interest rates. Competition tends to decrease interest rates by lowering lenders' market power: in more competitive markets, lenders set smaller markups over their break-even rates. However, more competition implies that each lender's market share is lower, so lenders have lower incentives to invest in screening out borrowers who are likely to default. Lenders thus face a riskier set of borrowers, and charge higher rates as a result. When the market contains a large fraction of high-risk consumers, the screening force can dominate, and equilibrium interest rates can actually increase, and consumer welfare can decrease, as markets become more competitive.

The main prediction of our model is that the relationship between concentration and interest rates depends on the level of default risk in the population. Interest rates should increase with market concentration in low-risk market segments, and interest rates should decrease with concentration in high-risk market segments. The model also makes two auxiliary predictions. The first is that concentrated markets should always have lower default rates, since lenders have higher incentives to screen. The second is that, in high-risk market segments, higher concentration can simultaneously lead to lower quantities and lower prices, as lenders with market power screen more intensively, but offer lower rates to borrowers that pass the screening. This cannot occur in an environment without some form of screening or rationing: demand curves slope downwards, so if lenders offered lower rates without screening, more customers will want to borrow.

Consistent with our predictions, we find an opposite relationship between interest rates and market concentration for low and high credit borrowers. For borrowers with high credit scores, above 600, we find the classical relationship that interest rates are higher in more concentrated markets. For borrowers with low credit scores, below 600, we find that interest rates are actually lower in areas with more concentrated markets. The relationship is true in the cross section, survives the inclusion of county and year fixed effects, and survives many alternative strategies for measuring market concentration and interest rates. We also find direct evidence that is suggestive that lenders engage in more screening in concentrated markets. In the cross-section, lenders in more concentrated markets are more likely to purchase CreditVision, an alternative data product sold by Transunion. Borrowers' average credit scores are also higher in more concentrated markets. We also find evidence for the other predictions of the model: in all markets, delinquency rates are decreasing in market concentration, and in high-risk markets, loan quantities are decreasing in market concentration, even though interest rates are also decreasing.

We further test our model's predictions using bank failures and mergers as quasi-random shocks to market competition. We link data on deposit shares from Federal Reserve call reports to data on bank mergers and acquisitions from the National Information Center. Following mergers, market concentration increases in counties with an acquired bank, and following large bank failures, market concentration tends to decrease. In both cases, consistent with our model's predictions, we find that increases in concentration lead to higher interest rates for high-credit-score consumers, but lower interest rates for low-credit-score consumers.

We discuss and rule out several potential alternative channels. The interactive effects of adverse selection and competition alone, as studied by Mahoney and Weyl (2017), cannot explain the results. With both competition and adverse selection, regardless of whether selection is adverse or advantageous, market power increases prices. Theories of screening through contract characteristics such as down payments also cannot explain our results (Veiga and Weyl (2016)). Moral hazard, or higher interest rates having a causal effect on delinquency, also cannot explain the results. In particular, moral hazard does not lead to higher competition correlating with high interest rates for low credit score borrowers. Markups charged by auto dealers cannot explain the asymmetry between low and high credit score individuals. Our re-

sults hold when we restrict the sample to pure auto lenders, or lenders (banks, credit unions, and other entities) who make multiple kinds of loans, suggesting that our results are not driven by heterogeneous funding costs for different kinds of lenders. Finally, since loan quantities are decreasing in market concentration, the primary effect of competition appears to be through screening, rather than improved collections technology, though this channel may also partially contribute to the effects that we find.

Our results imply that consumers may not always benefit from increased competition in credit markets. In business lending, competition is known to have potentially adverse effects, because it impairs the ability of banks to engage in relationship lending (Petersen and Rajan, 1995). This channel appears less relevant for consumer credit markets, since lender-borrower relationships are likely less important for households relative to firms. Our results show that competition can distort outcomes in consumer credit markets through a different channel.

The channel in our paper is also distinct from that of, for example, Mian and Sufi (2009) and Mahoney and Weyl (2017), who argue that credit market competition can lead to credit over-provision. In these papers, increased competition lowers interest rates, and attracts borrowers who are more likely to default. This may decrease social welfare, but always benefits borrowers. In contrast, in our model, increased competition causes lenders to invest less in screening, and instead to set higher interest rates for all consumers. This leads to inefficient credit allocation, because lenders are less able to tell which consumers are creditworthy. In our setting, increased competition can actually make consumers worse off.

This paper joins a literature on the effects of competition in credit markets. This paper presents a new model of competition in consumer credit markets, with the counterintuitive result that in selection markets greater competition can actually harm consumers. Petersen and Rajan (1995) study competition and relationship banking. Parlour and Rajan (2001) provide a theoretical model of competition in loan contracts with multiple borrowers. Drechsler, Savov and Schnabl (2017), Drechsler, Savov and Schnabl (2018) and Egan, Hortaçsu and Matvos (2017) study deposit market concentration. Calomiris (1999) studies the efficiency of bank mergers. Argyle, Nadauld and Palmer (2020b) study the real effects of search frictions in auto lending and Buchak and Jørring (2021) study the effects of competition on lending and discrimination in the mortgage market. There is also a large literature on the effects of bank

branching deregulation, for example Jayaratne and Strahan (1996), Economides, Hubbard and Palia (1996), Krozner and Strahan (1999). Einav, Jenkins and Levin (2012) present a model of subprime auto lending under imperfect competition with different risk types. Bank competition is also known to affect local industry structure (Cetorelli and Strahan, 2006). Livshits et al. (2016) study a model in which lenders have a fixed costs of contract design. The paper also joins work on the relationship between monitoring and competition in finance. Giroud and Mueller (2010) and Giroud and Mueller (2011) study the relationship between competition and corporate governance. Consistent with interactive effects of monitoring and competition, they point to monitoring playing a more important role in less competitive industries. [2]

This paper also joins a body of work on the interaction of adverse selection and competition. We show that, perhaps surprisingly, in some cases greater market concentration can benefit consumers. Previous papers have rather focused on the fact that selection can mitigate the harmful effects of market power. Mahoney and Weyl (2017) provide a model of imperfect competition, and show that in the presence of adverse selection many of the harmful effects of imperfect competition are mitigated. Lester, Shourideh, Venkateswaran and Zetlin-Jones (2019) analyze competition and adverse selection in a search-theoretic model, finding that increasing competition can decrease welfare. Crawford, Pavanini and Schivardi (2018) study the interaction of competition and adverse selection in corporate credit markets. Vayanos and Wang (2012) study the interaction of selection and competition in asset markets.

This paper is also related to a theory literature on competition between banks, when banks have imperfect information about borrowers' default risk. An early paper in this literature is Broecker (1990). He, Huang and Zhou (2020) studies competition in lending markets when borrowers can decide whether to share data with lenders. Our signal structure is similar to He, Huang and Zhou (2020): banks are able to screen out some bad types, and the remaining population has a mix of bad and good types. Hauswald and Marquez (2006) analyzes a related model, in which banks acquire information on borrowers distributed on a circle, and banks are better at acquiring information on borrowers they are closer to. The main difference between

---

[2]Most narrowly, this paper also joins a body of work on auto lending. For example, Adams, Einav and Levin (2009) study liquidity constraints in subprime auto lending, Argyle, Nadauld and Palmer (2020a) study the demand for maturity, Benmelech, Meisenzahl and Ramcharan (2017) study liquidity, Grunewald et al. (2020) study dealers' joint decisions of loan and car prices, and Einav, Jenkins and Levin (2013) study the introduction of credit scoring.

our paper and this literature is that we assume information acquisition is a fixed cost, rather than a variable cost. Variable-cost screening models are appropriate for modeling bank lending to firms, where banks make relatively few loans, and the evaluation and underwriting process is manual and relatively heterogeneous across firms. In contrast, consumer lending settings generally involve screening technologies with large fixed costs and lower variable costs: lending decisions are often made using software, default prediction models, and more automated processes.

The remainder of this paper is organized as follows. Section 2 presents our theoretical model, and shows that, with monitoring and adverse selection, greater market concentration can lead to a rise in prices. Section 3 presents data and institutional background. Section 4 presents empirical evidence consistent with our model. Section 5 discusses potential alternative channels. Section 6 concludes.

2   Model

We build a model in which differentiated lenders invest in a costly technology to screen potential borrowers, and then set loan rates for lending to the borrowers. There are N lenders, indexed by j.[3] Lenders compete in a two-stage game. In the first stage, lenders simultaneously choose how much to invest in screening out bad-type consumers. In the second stage, lenders compete a la Bertrand, simultaneously posting interest rates at which they are willing to lend to consumers.

Consumers. Each consumers wishes to borrow a unit of funds from lenders. There are two types of consumers: there is a unit mass of type G consumers, and a measure q of type B consumers. G consumers always pay back loans, and B consumers always default. We assume for simplicity that type B consumers default without paying any interest, and that recovery rates are always 0, so type B consumers cost lenders the principal of 1 and pay nothing. We relax the assumption of zero recovery rates in Appendix A.4. In Appendix A.6, we assume there is moral hazard for type G consumers: type G consumers default with some probability that is increasing in the interest rate they are charged. In both cases, our results are qualitatively

---

[3]In Appendix A.5, we show that N can be micro-founded as the equilibrium outcome from an entry game, in which lenders sequentially decide whether to pay a fixed entry cost to enter the market.

unchanged.

The willingness-to-pay of consumers for loans is independent of whether they are type B or G. Consumers' preferences over lenders are described by a Salop (1979) circle. The N lenders are uniformly spaced around a unit circle, and consumers are arranged uniformly on the circle. Consumers' preferences for lenders are a function of distance: the utility of a consumer who borrows from lender j, at loan rate r_j, is:

$$\mu - r_j - \theta x_j \tag{1}$$

where x_j is the distance between the consumer and lender j on the circle. The constant μ affects the total utility of the consumer for borrowing. We assume that μ is high enough that consumers do not choose the outside option in equilibrium. We also assume that $\mu < 1$, so that type B consumers' willingness-to-pay is never higher than the social cost, 1, of providing credit to them, so it is socially inefficient to provide credit to type-B consumers. We use the Salop circle because it is a simple model of imperfect competition, in which lenders set markups which depend on the number of competitors present. In Appendix A.8, we show that our main results also hold if borrowers' preferences over lenders are instead described by a logit model.

Costly screening. We assume that lenders can invest in a *fixed-cost screening technology* to imperfectly detect type-B consumers: by paying a fixed amount, lenders obtain a signal which is informative about consumers' types. We can think of auto lenders' screening technologies as, for example, acquiring alternative data, hiring employees to build default rate prediction models and automated decision-making software. These technologies have large fixed-cost components: building models or decision-making software has large upfront costs, but the variable cost of scaling a model to more consumers is low.[4]

The assumption that screening has fixed costs is the main difference between our model and much of the previous literature on bank lending, in which information acquisition is assumed to be a variable cost, scaling linearly with the number of loans made (Broecker (1990), Hauswald and Marquez (2003), Hauswald and Marquez (2006)). Information acquisition is likely to be

---

[4]In the main text, we assume screening is entirely fixed-cost for expositional simplicity. However, our main results hold screening has both fixed and variable cost components: as we show in Appendix A.4, any variable cost of lending would simply increase interest rates uniformly.

more variable-cost-intensive in firm lending settings, where banks make fewer loans, and tend to manually review and underwrite each loan, in contrast to the more automated process in consumer lending settings.

Formally, we assume that each lender j can pay a fixed cost to acquire an imperfect signal of borrowers' types. The signal has a "bad-news" structure: for type-G borrowers, the lender always observes a good signal. For type-B borrowers, the signal is good with probability α_j and bad with probability $1 - \alpha_j$, independently across borrowers. The cost of buying a signal with strength α_j is $\tilde{c}(\alpha_j)$. We assume $\tilde{c}(\alpha_j)$ is strictly decreasing in α_j: stronger signals (smaller α_j) are more expensive. Aggregating across borrowers, a measure $q(1 - \alpha_j)$ of type-B customers will receive bad signals: lender j knows with certainty that these customers are type-B, and will not lend to them.[5] A measure $q\alpha_j$ of type-B consumers receive good signals, and are indistinguishable from the unit measure of type-G consumers. Since the cost $\tilde{c}(\alpha_j)$ does not depend on the number of borrowers that the lender interacts with, signal purchasing is more cost-effective if the lender has a larger share of the market.

In the main text, we assume that type B consumers are ordered in terms of how easy they are to screen. Thus, if all firms attain a signal with strength α, they screen out exactly the same measure $q\alpha$ of consumers. This implies that, in any symmetric equilibrium where firms choose the same value of α, firms' signals about consumers are perfectly correlated: a consumer who is detected as a type-B by one firm is detected by all other firms, and any type-B consumers who are not detected are treated identically by all firms.[6]

If the population fraction of type-B consumers is q, and lender j chooses signal strength α_j, the default rate among borrowers with good signals is:

$$\delta_j = \frac{\alpha_j q}{1 + \alpha_j q} \tag{2}$$

---

[5]Since type-B consumers always default, and never pay interest or principal, there is no rate at which it is profitable to lend to them.

[6]If firms' signals are not perfectly independent across customers, firms would be able to infer additional information about customers' types from whether other firms are willing to lend to customers, so the inferred default rates among marginal consumers are different from default rates among average consumers. This complicates the model without adding significant insight, so we assume this away in the main text. However, we partially relax this assumption in appendix A.7, and show it does not affect our main results.

Inverting, in order to attain default rate δ_j, lenders must choose:

$$\alpha_j = \frac{\delta_j}{q(1-\delta_j)}$$

We can thus think of lenders as choosing a desired default rate δ_j. The cost of attaining default rate δ_j, when the population measure of bad types is q, is:

$$c_q(\delta_j) = \tilde{c}\left(\frac{\delta_j}{q(1-\delta_j)}\right) \tag{3}$$

Note that, since $\tilde{c}$ is decreasing, the function $c_q(\delta_j)$ is increasing in q, fixing δ_j: when there are more bad types in the population, it is more costly to attain any given default rate, because the lender must acquire a stronger signal (a lower value of α_j) to do so.

2.1 Equilibrium

We restrict to symmetric equilibria of the model, in which lenders make optimal screening and price-setting decisions. We solve the model backwards, solving for optimal price-setting decisions given default rates, and then solving for optimal screening decisions. First, note that in any symmetric equilibrium, lenders set identical interest rates, and each lender has market share:

$$s_j = \frac{1}{N}$$

Price setting. After screening is complete, lenders simultaneously set interest rates for customers who are not screened out as bad types. In the main text, we assume lenders can borrow at zero interest rates; we relax this in Appendix A.4. Lender j thus chooses r_j to maximize:

$$\Pi = \left(\frac{1}{1-\delta_j}\right) s_j \left(r_j(1-\delta_j) - \delta_j\right) \tag{4}$$

where s_j is j's market share, and the total measure of customers that pass j's screening is $\frac{1}{1-\delta_j}$. Intuitively, (4) says that, with probability $(1-\delta_j)$, the customer borrows a unit of funds, and pays back $1+r_j$, so the lender's profit is r_j. With probability δ_j, the customer defaults, and the

lender loses the principal of 1. Lenders' profits can be rearranged to:

$$\Pi = s_j \left(r_j - \frac{\delta_j}{1 - \delta_j} \right) \tag{5}$$

In Appendix A.1, we show that lender j's optimal interest rate r_j, in symmetric equilibrium, satisfies:

$$r_j - \frac{\delta_j}{1 - \delta_j} = \frac{\theta}{N} \tag{6}$$

The intuition for (6) is that $r_j - \frac{\delta_j}{1-\delta_j}$, the markup of r_j over the break-even interest rate, $\frac{\delta_j}{1-\delta_j}$, is higher when θ is higher, so consumers are more distance-sensitive and thus less price-sensitive, and when N is lower, so markets are more concentrated.

Optimal screening. Lender j chooses a desired default rate δ_j to maximize total lending profits, minus screening costs $c_q(\delta_j)$. That is, lender j solves:

$$\max_{\delta_j} \max_{r_j} s_j(r_j) \left(r_j - \frac{\delta_j}{1 - \delta_j} \right) - c_q(\delta_j) \tag{7}$$

In Appendix A.2, we characterize first- and second-order conditions for lenders' optimal information acquisition decisions. The first-order condition is:

$$\frac{s_j}{\left(1 - \delta_j\right)^2} = -c_q'(\delta_j) \tag{8}$$

Intuitively, the left-hand side of (8) is the marginal benefit of decreasing the default rate δ_j by a small amount, which is higher when j's market share, s_j, is higher.[7] The right-hand side is the marginal cost of decreasing δ_j, which depends on the fraction of type-B consumers in the population.

Combining (6) and (8), the following proposition states conditions on r, s, δ which characterize a symmetric equilibrium.

Proposition 1. *Necessary conditions for a symmetric equilibrium are that all lenders' market*

---

[7]While we focus on market shares, the average fixed cost decreases with market size. This is explored in the appendix, and we find empirical results consistent with stronger asymmetric effects in larger markets.

shares s, default rates δ, and interest rates r as are follows. Market shares of each lender are:

$$s = \frac{1}{N} \tag{9}$$

Lenders must set prices optimally:

$$r - \frac{\delta}{1-\delta} = \frac{\theta}{N} \tag{10}$$

All lenders must make optimal screening decisions:

$$\frac{1}{N(1-\delta)^2} = -c_q'(\delta) \tag{11}$$

$$c_q''(\delta)(1-\delta)^4 + \frac{2}{N}(1-\delta) > \theta \tag{12}$$

Total loan quantity is:

$$\frac{1}{1-\delta}$$

In any equilibrium, total consumer welfare of type G consumers is:

$$\mu - r - \frac{\theta}{4N} \tag{13}$$

Note that expression (13) for consumer surplus only accounts for type *G* consumers. This allows us to illustrate how non-defaulting consumers, who always have a willingness-to-pay which is higher than the cost of providing credit to them, are affected by the two forces of market power, and imperfect screening by lenders, which causes them to pool with type *B* consumers.[8]

We proceed to solve the model numerically. For our simulations, we assume costs take the form:

$$\tilde{c}(\alpha) = \frac{k}{\alpha}$$

---

[8]Since we have assumed that the social cost of providing credit to type *B* consumers is higher than their value for loans, it is never socially efficient to lend to type *B* consumers. Thus, screening also increases total social efficiency of credit allocation, by decreasing lending to type *B* consumers. We focus attention on the two-type case to capture the main qualitative insights of the setting. In a richer model with more than two types of consumers, screening could potentially have more complex distributional implications: an imperfect screening technology could reduce credit to some consumers who should receive credit in the first-best allocation.

Plugging into (3), this implies that:

$$c_q\left(\delta_j\right) = \frac{kq\left(1-\delta_j\right)}{\delta_j} = \frac{kq}{\delta_j} - kq \tag{14}$$

Expression (14) shows that screening costs are increasing in the parameter k, and the measure of type-B consumers, q. When q is large, screening costs are high, and it is more costly on the margin to decrease the default rate δ by any given amount. When q is 0, there are no bad types in the population, so it is costless for the firm to achieve a default rate of 0.

Figure 1 shows equilibrium outcomes, as we vary the number of lenders, for different levels of q. Throughout, we fix $k = 0.001$, as this generates realistic numbers for interest rates. The x-axis of each plot is the Herfindahl-Hirschman index:

$$HHI \equiv \sum_j s_j^2$$

In our model, this is simply $\frac{1}{N}$. The top-left panel shows the interest rate, r. When q is low, so the consumer population is low-risk, the relationship between r and competition is consistent with classical theory: interest rates are higher when HHI is higher and markets are more concentrated. However, when q is high, so the population is risky, we obtain the opposite result: interest rates are actually lower when when HHI is higher and markets are more concentrated.

To illustrate the forces driving this results, the top-middle plot shows the default rate δ, and the top-right plot shows the equilibrium markup charged by lenders over the break-even price, $r - \frac{\delta}{1-\delta}$. The top-middle plot shows that the default rate δ is always decreasing in HHI, regardless of q. Intuitively, when HHI is higher, firms have larger market shares, and thus higher incentives to invest in screening, lowering equilibrium default rates. The slope of the HHI-δ curve depends on q, the level of risk in the population.

The top right plot shows the markup that firms charge over the break-even price in equilibrium, $r - \frac{\delta}{1-\delta}$. Markups are always higher in more concentrated markets, since firms have more market power. However, unlike default rates, the effect of concentration on markups is insensitive to the average riskiness of the borrower population.

The net effect of concentration on interest rates combines the effects of concentration on

default rates and on markups. When q is low and the population is low-risk, the markup effect tends to dominate, so interest rates are increasing in concentration. When q is high and the population is high-risk, the screening effect tends to dominate, and interest rates are decreasing in concentration. The effect of concentration on default rates can be strong enough that the welfare of good-type consumers is actually higher in more concentrated markets: the bottom left panel shows that, when q is large, consumer surplus increases as *HHI* increases.

The bottom right panel of Figure 1 shows total loan quantities as a function of concentration. In our model, loan quantity tends to decrease when concentration is higher. This is true even when q is high, and concentration is negatively correlated with interest rates. This is because lenders screen more in more concentrated markets, removing bad types from the population, and thus decreasing equilibrium loan quantities, even though interest rates are lower.

Based on Figure 1, we can derive three testable predictions to bring to the data.

Prediction 1. *When the level of default risk in the consumer population is low, higher concentration tends to increase interest rates. When the population default risk is high, concentration tends to decrease interest rates.*

This follows from the top left panel of Figure 1. The next prediction concerns default rates, from the top-middle panel.

Prediction 2. *Higher concentration always leads to lower default rates.*

Finally, the bottom-right panel of Figure 1 makes a prediction about loan quantities.

Prediction 3. *When the population default risk is high, higher concentration can simultaneously lead to lower interest rates and lower loan quantities.*

The intuition behind Prediction 3 is that, in high-risk markets, lenders screen more, limiting the set of consumers that receive loans, and offer lower interest rates to these consumers. This cannot occur in an environment without screening: market demand curves slope downwards, so lower prices will always lead to higher quantities, if all customers are allowed to borrow at the market price.[9]

---

[9]We note that, in the baseline model, quantities are always decreasing in concentration, because there is no

3 Data and Institutional Background

3.1 Institutional Background

While the model presented in section 2 can broadly apply in consumer credit markets, we focus our empirical analysis on auto loans for three reasons. First, the institutional details pertaining to auto lending are relatively simple and direct relative to other large consumer loan markets, like the mortgage and student loan markets. Second, the auto loan market is largely segmented along borrower riskiness. Borrowers with different credit scores and risk tend to purchase different vehicles and utilize different lenders. Finally, unlike mortgage and student loans, auto lending is typically not guaranteed, and so losses are directly incurred to the lender– an important feature of our model. Relatedly, securitization can also change lenders' incentives to screen borrowers (Keys, Mukherjee, Seru and Vig, 2008), and one attractive feature of the market is that auto loans are also less likely to be securitized relative to mortgage loans.

Auto loans are the third largest source of household debt in the United States, following mortgage and student loans. The Federal Reserve Bank of New York reports approximately $1.4 trillion in outstanding auto loan debt in 2020.[10] The vast majority of auto purchases in the US are financed. Over 95% of American households own cars and the National Association of Auto Dealers estimates that in 2019, 85% of new vehicles and 55% of used vehicles were purchased using auto loans. According to Experian, in 2020 31.2% of auto loans were made by captive subsidiaries, 30.2% were made by banks, 18.7% were made by credit unions, 12.4% by finance companies and the remaining 7.6% by dealers themselves.[11]

There are two types of auto lending, direct and indirect. Direct lending implies that consumers take a loan directly from a financial institution, and use that to purchase a vehicle. The consumer will submit information to a lender, and the lender will decide whether to approve the loan. Under indirect lending, the consumer applies for a loan through the dealer

---

extensive margin: customers never choose not to borrow. Thus, total loan quantity depends only on lenders' screening decisions: when markets are more concentrated, lenders screen more intensively, so total loan quantity decreases. In a richer model, such as that of Appendix A.8, prices would also affect customers' choices on the extensive margin, so higher concentration can conceivably lead to higher quantities, if lenders can lower prices sufficiently to attract many more type-G customers into the market.

[10]Levitin (2020) provides a detailed description of the auto loan market.

[11]In the appendix, we show that our results hold if we restrict to banks and other lenders that do not exclusively offer auto loans.

and dealers obtain financing through third party lenders. Dealers typically have relationships with several lenders, and after providing lenders with borrower information the dealers solicit offers for the minimum interest rate that a lender will provide.

Importantly for linking to our model, auto loans typically remain on lenders' books. Hence lenders incurs costs if borrowers default. In 2020, only 14% of auto loans were securitized according to SPG Global.[12] A slightly higher fraction of subprime loans are securitized, but the vast majority – three quarters – of subprime auto loans are not securitized.

3.1.1 Monitoring

Beyond the most basic form of screening, accepting or rejecting clients based on credit scores, there are a number of ways in which auto lenders monitor borrowers. These methods typically incur costs to lenders. For example, some auto lenders install GPS tracking to be better able to repossess vehicles in the case of default. Lenders can also invest in predictive analytics, sometimes using machine learning and artificial intelligence, to identify which borrowers are less likely to default, even within subprime categories. Modeling prepayment risk can also allow lenders to go beyond traditional credit scoring. Some lenders also verify a car's condition before lending to risky borrowers, since lemons are useless as collateral. All of these actions would lead to better screening and monitoring of borrowers, but would incur costs for the lender either through building better predictive analytics, paying to inspect vehicles or installing additional features to track vehicles.

In our analysis, we will use county-year HHI as a measure of market competition. This implicitly assumes it is a lender's local market share which matters for screening decisions, which would be reasonable if lenders' screening decisions are local in nature. There are a number of reasons why this might be the case. First, most auto loans are made through lenders' relationships with dealers. If lenders need to invest in dealer-specific information acquisition, for example to determine dealerships with higher default rates, these investments would be location-specific in nature. Second, some information used to estimate default risks is very local in nature. For example, borrowers living in certain neighborhoods may have different

---

[12]Similarly, Klee and Shin (2020), using data from SIFMA, state that the quantity of outstanding auto ABS was around $225 billion in 2018, which is around 18% of the total outstanding volume of auto loans.

default risks than others; detecting these relationships and using them to price loans may require location-specific data acquisition and analysis.[13]

3.2 Data

3.2.1 Booth TransUnion Consumer Credit Panel

Our main data source is the Booth TransUnion Consumer Credit Panel. The data is an anonymized 10% sample of all TransUnion credit records from 2009 to 2020. Individuals who were in the initial sample in 2000 have their data continually updated, and each year 10% of new first time individuals in the credit panel are added. A small fraction of individuals also leave the panel each year, for example due to death or emigration.[14] We define markets at the county level, and our main analysis dataset consists of new loans at the county by credit score by year level. We drop observations with fewer than ten loan contracts annually.

We can observe basic information about a loan, including the original balance, the current balance, scheduled payments, and maturity of the loan. We can also observe other borrower-level information, including VantageScore and geographical variables. Interest rates are not directly observed, and thus we back these out using scheduled payments. We take the first observation for each loan and use the amortization formula to calculate interest rates $A = \frac{P \times i}{1-(1+i)^{-n}}$, where A is the monthly payment due, P is the principal amount on the loan, n is the maturity in months, and i is the interest rate, we solve for i using a root-solving algorithm, after removing missing observations for each requisite variable.[15] The identities of the lender that originated the loan, as well as the loan customer, are anonymized by the data provider.

Total loan volumes in our data are also comparable to measures from other datasets: we plot the time series of the total number and dollar volume of car loans from different datasets in Appendix Figure A.4. Appendix Figure A.5 shows the distribution of consumers by credit

---

[13]Address history and property values appear to be important components of some alternative credit data products. See, for example, LexisNexis RiskView.

[14]Keys, Mahoney and Yang (2020) provide more details about the Booth TransUnion Consumer Credit Panel and Herkenhoff, Phillips and Cohen-Cole (2019) and Braxton, Herkenhoff and Phillips (2020) more generally discuss TransUnion data. All tables and figures that list TransUnion as a source have statistics 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.

[15]We drop a small number of observations where predicted interest rates are either negative or implausibly large. We further winsorize rates at the .2% level.

score in our data. The average interest rate in our sample is 9.3%, which compares to averages rates of 5.9% for new vehicles, and 9.5% for used vehicles, from the National Association of Auto Dealers. Interest rates are much higher for groups with lower credit scores. In the lowest credit score groups in our sample (below 600), the average interest rate is 15.07%. This is approximately four times the average interest rate in the highest credit score group (above 800), which is 3.67%.[16] These patterns likely reflect greater charge-off probabilities for low credit-score borrowers. In the lowest credit score group, the average 90-day delinquency rate is 34.26%, while it is 1.06% for the highest credit score group.

Our primary measure of market competitiveness is the Herfindahl–Hirschman Index, or HHI. We construct HHI using the volume of auto loans, that is

$$HHI_{ct} = \sum_{l=1}^{N} s_{clt}^2 \tag{15}$$

where s_{clt} is a lender l's dollar share of auto lending in a county c in year t within a credit score range. An HHI of zero means the market is perfectly competitive, while an HHI of one indicates monopoly. The average HHI in our sample is .05, suggesting that the auto lending market is on average quite competitive. Appendix Figure A.6 shows the geographic distribution of HHI.[17]

TU also provides novel data on Creditvision, a proprietary product which lenders can purchase. Creditvision contains additional information on consumer behavior and histories. Lenders who access Creditvision have additional tools that can be used to screen borrowers, including predictive modeling, purpose-built scores, propensity models, attributes, algorithms, and estimators. The Creditvision data is available at the state level, from 2016 to 2019. We observe the total number of lenders active in a state, and the total number of lenders that purchased the product from TU.

---

[16]These are quite similar to rates published by Experian in 2020. The average rate for Deep Subprime borrowers with credit scores below 580 is 14.39% for new cars, and 20.45% for used cars. For Subprime borrowers with credit scores between 580 and 620 the corresponding rates are 11.92% and 17.74% respectively. For Super Prime borrowers with scores above 720, the average rate for a new car loan is 3.65% and the average rate for a used car loan is 4.29%.

[17]When we restrict to bank lenders, the HHI is .12. This is comparable to estimates in the literature. For example, Kahn, Pennacchi and Sopranzetti (2005) estimate an average MSA-level HHI of of .14 for commercial banks' personal and auto loan market shares from 1989-1997, and Drechsler, Savov and Schnabl (2017) estimate an average county-level HHI of .22 for banks' deposit market shares from 1994-2014.

3.2.2 Bank Merger Data

We complement our main analysis with data on bank mergers. We obtain deposit shares from Federal Reserve Call reports. The bank mergers data is collected from the transformations file from the National Information Center (NIC). We assume that a county is affected by a merger if an acquired bank has positive deposit market share in the county. In cases with more than one merger, we use the first merger event. Between 2009 and 2019 there are a total of 1,442 mergers, covering 1,812 distinct counties.

Appendix Figure A.12 shows the geographical distribution of mergers. More than half of the counties in our sample are affected by a merger at some point in our sample period, and the affected counties seem to be fairly uniformly distributed across the US. Appendix Figure A.13 shows a binscatter of bank deposit market HHI, measured using the summary of deposits data, on the x-axis, with auto lending HHI, measured in the TransUnion data, on the y-axis. Auto lending HHI and bank deposit HHI are very strongly correlated, suggesting a tight link between the two.

4 Empirical Evidence

4.1 Interest Rates, Competition and Credit Scores

Prediction 1 of the model states that, when the population is low-risk, and screening costs are low, we see the classical relationship that competition tends to be associated with lower interest rates. When the population is high-risk, and screening costs are high, we see the opposite relationship that competition should be associated with higher interest rates. This prediction is borne out by the data.

Figure 2 presents our main result. The figure panels show median interest rates in a county for given credit score ranges, broken down in twenty equal-sized bins of HHI, our measure of the competitiveness of a market. The left panel shows the relationship for borrowers with VantageScore scores below 600, while the right panel shows the relationship for borrowers with VantageScore scores above 600. The two panels display strikingly different patterns, consistent with our model presented in Section 2. The left panel, which covers high-risk borrowers,

shows a strong, linear and negative relationship: in contrast with standard economic theory, interest rates are actually decreasing in market concentration. The right panel, which shows the relationship for low-risk borrowers, shows precisely the opposite relationship. Consistent with a standard framework, we see that interest rates are increasing in market concentration. The magnitudes of both relationships are fairly large: an increase in HHI from 0.05 to 0.15 is associated with approximately a 1% decrease in interest rates for high-risk borrowers, and a 1% increase in interest rates for low-risk borrowers.

Figure 3 shows the same relationship, broken down into finer credit score categories. We split the sample into six credit score bins. We see the strongest negative relationship between interest rates and concentration for deep subprime borrowers, with credit scores lower than 550. We see a flatter relationship for credit scores between 550 and 600, and for credit scores above 600 we generally see the classical relationship that interest rates are rising in concentration.

Table 2 presents similar information to the figures by presenting regression coefficients. More specifically, the table shows point estimates β and standard errors from specifications similar to

$$ln(r_{ct}) = \alpha_c + \alpha_t + \beta \, ln(HHI_{ct}) + \varepsilon_{ct} \tag{16}$$

where r_{ct} is the average interest rate for auto loans in a county, and HHI_{ct} is the Herfindahl–Hirschman Index measuring market concentration. We cluster standard errors at the county level. The main coefficient of interest β captures the effect of market concentration on interest rates. We run estimates of specification (16) separately by different credit score buckets.

We additionally include county fixed effects α_c, which absorb time invariant county specific factors, such as geographic areas having riskier drivers, and α_t time fixed effects absorbing economy wide temporal shocks. The inclusion of time trends is particularly important, as they allow us to rule out that the observed patterns in Figure 2 are driven by temporal trends in both interest rates, credit scores and HHI. For example, in the absence of time fixed effects, it is possible that the differing relationships between the slopes of interest rates and market concentration are simply driven by a decline in the fraction of low-credit score borrowers coinciding with movements in interest rates and market concentration.

The top panel of Table 2 splits the sample between credit scores above and below 600, and

gradually adds in fixed effects. The first column of each triplet has no fixed effects, the second column adds in year fixed effects, while the final columns adds in both year and county fixed effects. In all three cases, we see a similar pattern and magnitudes. For borrowers with lower credit scores, we see that a 1 percent increase in market concentration, as measured by HHI, is associated with a .07 to .08 percent *decrease* in interest rates. The relationship is statistically significant at the 1% level in all specifications. For high-credit score borrowers, we observe the opposite relationship. A 1 percent increase in market concentration, as measured by HHI, is associated with a .19 to .30 percent *increase* in interest rates. The relationship is slightly weaker for high-credit score borrowers, significant at the 5% or 10% level in each specification.

The bottom panel of Table 2 splits the sample into finer credit score bins, including county and year fixed effects. We find a strong and highly significant negative relationship between interest rates and concentration for borrowers with credit-scores below 600, and a positive and significant or marginally significant relationship for borrowers with credit-scores above 600. The relationship is generally increasing as credit scores improve, consistent with the predictions of the model.

Table 3 presents an alternative specification, interacting HHI with credit score groups (above or below 600.) Specifically, the table shows variants of the coefficient γ and ϕ from the equation

$$ln(r_{cst}) = \alpha_{cs} + \alpha_t + \gamma ln(HHI_{ct}) \times \mathbb{1}[CreditScore_{Low}] + \phi ln(HHI_{ct}) + \varsigma_{cst} \qquad (17)$$

$\mathbb{1}[CreditScore_{Low}]$ is an indicator of borrowers being in the low credit score group and α_{cs} are county by score group fixed effects. Standard errors are clustered at the county level. The coefficient ϕ captures the effect of market concentration on interest rates for high credit score borrowers, while $\gamma + \phi$ captures the effect for low credit score borrowers.

Column (1) of Table 3 presents the baseline result. Consistent with the estimates in Table 2, we see a positive effect of market concentration on interest rates for high credit score borrowers, and a negative effect for low credit score borrowers. Effects are highly statistically significant, at the 1 percent level. This result holds without weighting counties (column 2), using only large counties (column 3), using all counties (column 4), and winsorizing interest rates at the 1% level instead of the .2% level (column 5). Column 6 shows the results from computing HHIs using loan number as a measure of market share, instead of loan values. The

effect of HHI in the high-credit-score group becomes insignificant, but the effect for low-credit-score consumers remains negative and significant. Column 7 uses the number of lenders active in a county as a measure of competition, instead of HHIs: the effect is insignificant for high-credit-score consumers, but for low-credit-score consumers, counties with more lenders tend to have higher interest rates. Appendix Figures A.7 and A.8 repeat the binscatters of Figure 3 with these different specifications; the stylized facts from the baseline specification hold in all cases.

For large lenders, screening decisions may be made at more aggregated levels than counties: for example, lenders may use the same analytics and decision software across branches in multiple regions. Thus, larger lenders' screening decisions should be less sensitive to local market HHIs. On the other hand, if lenders set loan markups at the level of local markets, the competition channel should affect large and small lenders similarly. This implies that interest rates should be more *positively* correlated with HHIs for larger lenders, since the competition channel plays a larger role than the screening channel.

To test this hypothesis, Appendix Table A.3 interacts HHI with the total volume of outstanding loans and the number of counties in which a lender is active. As predicted, we find that the interaction effect between HHI and both lender size measures is positive, for both prime and subprime groups. In words, local concentration is more positively associated with interest rates for larger lenders. Extrapolated out of sample, the estimates in Appendix Table A.3 suggest that under a monopoly, interest rates will be 9.1 percentage points lower for subprime borrowers for a lender operating in a single county, but only 5.9 percentage points low for a lender operating in all counties in the United States. Notice, however, that the relationship between local HHIs and interest rates is negative even for large lenders; this suggests that, even for large lenders, local factors may play some role in screening.[18]

4.2 Direct Evidence of Screening

We next provide suggestive evidence that lenders engage in more screening in more concentrated markets. Figure 4 presents more direct evidence that lenders engage in more screening

---

[18]One example of this could be investing in relationships with dealers, or otherwise acquiring local information. For example, Bank of America operates a large dealer network dealer network.

in more concentrated markets. The figure shows that in the cross section, lenders in more concentrated markets purchase more information on borrowers, and borrowers have higher credit scores.

The top panels show the fraction of lenders purchasing Creditvision, by ventiles of HHI. Creditvision is a proprietary product from TU, which contains additional information on consumer behavior and histories. Creditvision includes predictive modeling, purpose-built scores, propensity models, attributes, algorithms, and estimators. The data is available at the state level, from 2016 to 2019. We construct auto loan HHI at the state level. The left panel shows the relationship between Creditvision purchasing and HHI based on the volume of loans, while the right panel shows the same relationship using a measure of HHI based on the number of loans.

Both panels show a similar relationship consistent with our model: in more concentrated markets, lenders are more likely to purchase additional information from TU. Regressing the fraction of lenders purchasing Creditvision on the volume-based measure of HHI yields an OLS coefficient of 0.965 and a standard error of 0.437, clustered at the state level. A similar regression using the number of loans based measure of HHI yields a coefficient of 0.881 with a standard error of 0.335.

The bottom panels show average credit scores in a county, by ventiles of HHI. Again the left panel shows the relationship between credit scores and volume based HHI, while the right panel shows the relationship using a measure of HHI based on the number of loans. Consistent with more screening in concentrated markets, we see higher credit scores in more concentrated markets. Regressing the fraction of lenders purchasing Creditvision on the volume-based measure of HHI yields an OLS coefficient of 22.97 with a standard error of 4.354, clustered at the county level. A similar regression using the number of loans based measure of HHI yields a coefficient of 22.80 and a standard error of 1.443.

4.3   Loan Delinquencies

We next turn to the relationship between market concentration and loan delinquency rates. Prediction 2 of the model states that lower competition always leads to lower default rates. This prediction is again consistent with the data.

We estimate variations of equation (16), in which the outcome is the fraction of loans that ever become delinquent for more than ninety days. Table 4 presents estimates of the relationship between delinquency and market concentration. Consistent with the predictions of the model and concentration leading to more monitoring, we see a negative relationship between delinquency and concentration. The effect sizes change significantly when county fixed effects are included, suggesting that it is important to control for regional heterogeneity. We find a significant negative relationship between delinquency rates and HHI for borrowers with low credit scores, and a statistically insignificant relationship for high credit score borrowers.

The bottom panel of Table 4 splits the sample in finer credit score bins, again including county and year fixed effects. Consistent with the theory, in all estimates the observed relationship is negative. The results suggest that a 1 percent increase in market concentration is associated with a .3 to .12 percent reduction in delinquency rates. The estimates are statistically significant at the 5% or higher level, except for the most creditworthy borrowers. This lack of statistical significance in the top bin may simply be due to a lack of variation in the dependent variable, as very few borrowers with high credit scores become delinquent on their loans.

4.4   Loan quantities

Next, we test Prediction 3, regarding the relationship between concentration, interest rates, and loan quantities in low credit-score buckets. In Table 5, we estimate panel regressions, in which we regress the log of loan quantity on the log of HHI. We use only panel regressions, since cross-sectional variation in loan quantities likely is driven largely by the size and demographic composition of counties.

For all credit score buckets, the panel coefficients are negative and statistically significant, implying that increases in county HHIs are associated with decreases in the number of loans made. Quantitatively, a 1% increase in HHI is associated with a 2.7% to 5.2% decrease in loans made. The coefficients tend to be larger for lower credit score groups.

For consumers with credit scores above 600, we found that higher concentration is associated with higher interest rates. These patterns are consistent with classical intuitions about market power: in concentrated markets, firms set higher prices, reducing equilibrium quanti-

23

ties.

For consumers with credit scores below 600, however, we showed in Table 2 that higher concentration is associated with lower interest rates. Market power alone cannot explain why higher concentration would be associated with lower interest rates, and also lower loan quantities: in the absence of some kind of screening or credit rationing, lower interest rates should induce more customers to borrow. Prediction 3 of our model does explain this: with higher concentration, firms invest more in screening, limiting loans to a subset of customers who are less likely to default, thus decreasing loan quantities.

4.5   Bank Failures, Mergers and Acquisitions

One significant concern is that the regressions presented earlier in this section could be biased, as HHI and interest rates are simultaneously determined. While it is unclear why simultaneity would generate an asymmetric pattern between interest rates and market concentration along credit scores, we use bank mergers and failures as quasi-exogeneous shifters of market concentration, as a natural experiment to provide additional evidence of an asymmetric relationship between rates and concentration. Bank mergers are commonly used to induce variation in market concentration, for example Sapienza (2002); Scharfstein and Sunderam (2016); Garmaise and Moskowitz (2006); Buchak and Jørring (2021); Liebersohn (2017) and Favara and Giannetti (2017).

Bank Failures. We first use bank failures during the 2008 financial crisis to generate variation in market concentration. Wachovia, Washington Mutual and Countrywide were the three largest household lenders that failed during the crisis. The failure of these lenders led to significantly lower market concentration in areas where these lenders had a greater market share prior to the crisis, due to other firms purchasing assets and capturing market share.[19] Buchak and Jørring (2021) note that frictions relating to market entry may have led to persistent effects of these banks' failure on concentration.

We use a standard 2SLS approach, and first use the market share of failed banks to predict market concentration, using the following equation.

---

[19]The effects were substantial, particularly in certain regions. Wachovia was the ninth-largest auto lender in the United States in 2006, prior to failure. Mayer et al. (2014) provide a discussion of the failure of Countrywide.

$$ln(HHI_{ct}) = \psi Share_c + X_{ct} + e_{ct} \tag{18}$$

where $Share_c$ is the market share of the three failed lenders in 2008, and X_{ct} are controls including the number of lenders and the sum of deposits in the county in 2008. We then run the second stage, estimating the effect of predicted market concentration $ln(\hat{HHI}_{ct})$ on interest rates.

$$ln(r_{ct}) = \zeta ln(\hat{HHI}_{ct}) + X_{ct} + v_{ct} \tag{19}$$

Standard errors are clustered at the county level. The variation we use for the instrument is cross sectional, so we do not include county or time fixed effects in this analysis. The identifying assumption is that the market share of the three failed banks is uncorrelated with auto loan interest rates in subsequent periods, other than through effects on market concentration. In particular, any correlation between the three banks' market shares and interest rates that differs for high and low-credit score borrowers would be of particular concern, since we are interested in the asymmetry of the price-concentration relationship for high and low credit score borrowers.

We test the inclusion restriction, and show that bank failure predicts concentration in the appendix. The first stage is shown in Appendix Tables A.6 and A.7. The tables show that, consistent with prior work and institutional details, market share of failed banks is associated with lower market concentration. F-statistics for the first stage are between 40 and 65, and thus well above standard levels for rule of thumb tests regarding weak instruments. Additionally, the first stage is quite similar for borrowers in high and low credit score groups.

Figure 6 shows estimates of ζ from equation (19), along with a 95% confidence interval, split by the credit score group. The results are consistent with the theoretical predictions in Section 2, as well as the earlier empirical results. We see a negative effect of concentration on interest rates in low credit score groups, and a standard positive effect for borrowers with high credit scores. The relationship is monotonic, with estimates of ζ rising with credit quality for each group.

Bank Acquisitions. We next use bank mergers as a source of variation which increases

market concentration. This strategy allows us to use panel variation. Figure 7 shows that there is little trend in concentration prior to the merge.[20] Specifically, the figure plots coefficients ζ_i from the following regression:

$$ln(HHI_{cst}) = \alpha_{cs} + \alpha_t + \alpha_y + \sum_{i=-5}^{t} \zeta_i \mathbb{1}[i = t] + \xi_{cst} \tag{20}$$

where α_{cs} are county by credit score group (above or below 600) fixed effects, α_t are year fixed effects and α_y merger year fixed effects. $\mathbb{1}[i = t]$ are indicators for time periods before and after acquisitions in a county. We define an acquisition as occurring if an acquired lender has any deposits in a county. Standard errors are clustered at the county level. The results suggest that concentration rises following bank mergers, which is consistent with prior work.[21]

Figure 8 shows a similar analysis, replacing the outcome in equation (20) with logged interest rates, $ln(r_{ct})$ and splitting the sample by high and low credit score groups, which we define as above and below a 600 VantageScore. Again consistent with the predictions in Section 2, we see that while interest rates rise for high-credit score borrowers following a merger event, they actually fall for low-credit score borrowers.

Figure 9 shows the relationship between interest rates and concentration for finer credit score groups. The figure shows estimates of the coefficient from a regression of interest rates on the number of years following an acquisition. While the estimates are less precise than those from Figure 9, we again see a weakly monotonic relationship with estimates of the interest rate concentration relationship rising with credit quality for each group.

---

[20]We present further results in the appendix. Figure A.14 shows that there is a drift upwards in HHI post-merger, as measured both by auto lending and bank deposit HHI.

[21]While we do find a relationship between acquisitions and HHI, the relationship is not strong during our sample period. F-statistics are between 6 and 10, and thus below or near thresholds for standard weak instrument tests. For this reason we do not use acquisitions as an instrument, but rather explore reduced form relationships. In Appendix Tables A.8 and A.9 we show that the first stage is similar for both high and low credit score borrowers.

5 Alternative Channels and Robustness

5.1 Additional Robustness Checks

We conduct a number of robustness checks of our results. Appendix Table A.2 regresses interest rates on concentration at the lender level. That is, we run specifications similar to:

$$ln(r_{clt}) = \alpha_c + \alpha_t + \alpha_l + \beta ln(HHI_{ct}) + \varepsilon_{clt} \tag{21}$$

where r_{clt} is the average interest rate charged by lender l in county c at time t, and HHI_{ct} is the HHI for county c, time t. In addition to county and time fixed effects, specification (21) allows us to add lender fixed effects, so that β is identified using within-lender variation in interest rates: how much a given lender tends to charge higher interest rates in counties with higher HHIs.

Columns 1 and 3 of Appendix Table A.2 do not include lender fixed effects: coefficient estimates are qualitatively and quantitatively similar to our baseline regressions in Table 2. Columns 2 and 4 add lender fixed effects: the coefficients essentially do not change. Our findings imply that, if a lender makes auto loans in multiple counties, she tends to charge lower rates in more concentrated markets for low-credit-score consumers, and higher rates in more concentrated markets for high-credit-score consumers. The fact that lenders' rates differ across regions, in the ways predicted by our theory, lends support for our hypothesis that lenders' monitoring and price-setting decisions are made at the local level, rather than at the lender level.

In columns 5-8, we divide lenders into two subsamples: lenders who only make auto loans (columns 5-6) and lenders who make multiple kinds of loans (columns 7-8). In both cases, we include county, year, and lender fixed effects. For both subsamples, the results are quantitatively similar, though column 8 shows that the effect of concentration on interest rates for high-credit-score customers in the non-auto-lender group is not significant.

Appendix Figure A.9 further shows that the results hold if restricted to lenders that were in the sample from 2009 onwards and remained until the end. In other words, the main result is not driven by the entry or exit of lenders.

Finally, we do a back-of-the-envelope calculation of the costs of the screening technology which are implied by our data. The average subprime loan in our data is $17,394. Recovery rates for subprime loans are roughly 40%,[22] implying that a lender would be willing to pay approximately $104 to decrease the default rate on a single subprime loan by 1%. The size-weighted median lender in our data makes 55,830 subprime loans per year, whereas the 25th percentile lender makes 3,700 subprime loans annually. This implies that the median lender would be willing to pay roughly $5.8 million per year for a technology that would decrease default rates by 1%, whereas the 25th percentile lender would only be willing to pay $384,000. Thus, a technology costing between $384,000 and $5.8 million per year would be worth paying for larger lenders, but not smaller ones. This seems like a reasonable range for the cost of an auto lending group building a data analytics team, for example.[23]

As a placebo test, in the online appendix we replicate our analysis in the mortgage market. We would not expect the screening channel to hold for the same groups in the mortgage market due to securitization regulations and the restrictions on pricing. Appendix figure A.15 and table A.10 reproduces the main analysis, using mortgage rather than auto loan interest rates. In both cases, we do not find any asymmetry and observe the interest rates are higher in areas with more market concentration. We caution though that the sample is small, and only 2.9% of mortgage loans are originated to borrowers with credit scores below 600. It is possible that the screening costs channel is present to some degree in the mortgage market for mid-prime borrowers, and we leave this to future work.

5.2 Alternative explanations

In this subsection, we discuss other possible explanations of our results, and show they cannot simultaneously rationalize all the stylized facts that we observe.

Adverse selection. In our model, the primary channel driving our results is costly infor-

---

[22]See S&P Global.

[23]If screening involves fixed costs, aggregate efficiency would be higher if lenders pooled their screening efforts: for example, if all lenders contributed to a centralized database, and then used these data for default prediction and pricing. However, this does not appear to have happened in car loans markets, or consumer credit markets more generally. This kind of pooling may be difficult because of coordination costs, as well as free-rider problems – individual lenders would want to use data from other lenders, but keep their own data so as to maintain a competitive advantage.

mation acquisition, which improves the quality of the borrower pool. A classic force, which we assume away for tractability, is adverse selection: the idea that, as prices vary, the riskiness of the borrower pool varies endogenously. It is difficult to rationalize our results using existing theories of selection. Mahoney and Weyl (2017) develop a general model of competition and selection. Proposition 1 of Mahoney and Weyl (2017) states that, regardless of whether selection is adverse or advantageous, market power always tends to increase prices. Thus, adverse selection alone, in the model of Mahoney and Weyl (2017), cannot match the result we find, that higher concentration is actually associated with lower interest rates in low credit score submarkets.[24]

Crawford et al. (2018) also study competition and adverse selection. The online appendix of Crawford et al. shows that, when adverse selection is severe, it is possible for prices and competition to have a U-shaped relationship: prices are decreasing and then increasing as market competition increases. This is due to a second-derivative effect: optimal markups depend on the marginal sensitivity of default rates to prices, which can vary with the level of market competition. It is possible that related forces are at work in our setting, but these second derivatives have ambiguous signs, so the direction of the effect is in general unclear. Moreover, we do not empirically observe a U-shaped relationship between competition and interest rates in the subprime market, and theories of adverse selection and competition would also not explain our results on the association between competition and lenders' direct investments in screening technologies.

Screening with down payments. In auto lending markets, lenders can also screen borrowers using down payments (Einav, Jenkins and Levin (2012), Einav, Jenkins and Levin (2013)). If borrowers' default rates are correlated with their preferences over down payments, lenders could separate high- and low-risk borrowers by varying how much down payments are required to take out a loan. We cannot directly test the effects of competition on down payments, since we do not observe the value of cars purchased and thus cannot estimate down payments. However, the down payment screening channel seems to be unable to match all of the stylized facts that we observe. Veiga and Weyl (2016) analyze an adverse selection model in which firms choose a single product's price and "quality", which can be thought of as down payments. Veiga

---

[24]See DeFusco, Tang and Yannelis (2021) for a further discussion of adverse selection in consumer credit markets.

and Weyl finds that firms' down payment choices incorporate a "sorting" effect: firms will tend to increase down payments if, among marginal consumers, consumers who dislike high down-payments are also more likely to default. When markets are more competitive, the elasticity of residual demand facing any given lender increases, so the sorting incentive increases; the Veiga and Weyl model thus predicts that down payments should on average increase, as lenders try to cream-skim profitable consumers. High down payments should induce more borrowers with high default rates to thus leave the market, so the population of borrowers should actually have lower default rates in more competitive markets; this is the opposite of what we find in figure 5 and table 4. Moreover, it is unclear how screening through down payments could explain why quantities are lower, and default rates higher, in more competitive markets. Down payments thus seem to be unable to explain all of our stylized facts.

Lester et al. (2019) analyze a model of competition, adverse selection, and screening through contract characteristics, allowing lenders to offer a menu of contracts rather than a single product. Lester et al. find that, when markets are more competitive, cream-skimming incentives are greater, so equilibria tend to be more separating: equilibrium menus tend to assign different contracts to different types of consumers, screening using consumers' differential valuations of contract characteristics. When adverse selection is severe, the effect of competition on welfare and contract characteristics is non-monotone: the distortion to high types' contracts is first decreasing and then increasing in market competitiveness. It is possible that competition also affects properties of lenders' equilibrium menus of contracts in our setting; however, the Lester et al. model would not explain our results on lenders' direct investments in screening technologies.

Moral hazard/direct effects. Another reason why interest rates may be correlated with default rates is that higher interest rates may have a direct effect on default rates, by increasing borrowers' payments. We formally show in appendix A.6 that, in a model with fixed screening costs and moral hazard, the main conclusions of our model continue to hold. However, moral hazard alone, without fixed costs, has difficulty explaining all of our stylized facts. Moral hazard could explain our finding that default rates are high in competitive and low credit score submarkets, based on the fact that interest rates are high. However, this does not explain why high competition correlates with high interest rates only for low credit score borrowers,

and why high competition correlates with low interest rates and high default rates for high credit score borrowers.

Competition and loan standards. A number of papers argue that increased competition in credit markets can lead to more credit and lower lending standards in equilibrium (Mian and Sufi (2009), Favara and Imbs (2015), Mahoney and Weyl (2017)). Our channel is related to these, but distinct. In particular, the models of Mian and Sufi (2009) and Mahoney and Weyl (2017) predict that increased competition should always be associated with *lower* interest rates. In our model, fixed screening costs imply that higher competition can in fact lead to *higher* interest rates, since all firms have lower incentives to screen consumers in equilibrium.

On the one hand, the high-level policy conclusions from our work and earlier work are similar: in financial markets with information frictions, competition may not improve market outcomes. However, in our model, increased competition in credit markets can actually also make creditworthy consumers worse off, an outcome which is not possible in the model of Mahoney and Weyl (2017).

Dealer markups. Another hypothesis is that vertical competition is driving these effects. It is possible that, in more competitive areas, lenders pay dealers higher markups. Thus, competition could decrease the rates lenders receive, but customer-facing rates, which is what we observe in our data, may be higher, because dealers are able to extract greater markups for intermediation in more competitive markets.

There are a few reasons why it is difficult to explain our results using this channel. First, it is unclear why dealer markups should be more important for lower credit score individuals, whereas our information acquisition theory predicts in particular that competition tends to lead to higher rates in low credit score areas. Second, this theory does not explain why default rates should increase with competition, except through a causal effect of higher rates on default. Third, Appendix Table A.2 shows that our main result holds for lenders who only make auto loans, as well as lenders who make multiple kinds of loans, such as banks. The auto-loan-only group includes dealers and integrated lenders, who would have no incentives to charge markups. The fact that our stylized facts hold within each of these groups suggests that dealer markups do not explain the entirety of our effect.

Heterogeneous funding costs. A related hypothesis is that larger lenders may have lower

costs of funding loans. Similar to the dealer markups hypothesis, this could explain why more competitive areas have higher prices, but cannot explain why the sign of the effect differs for high and low credit score groups, and why default rates are higher in more competitive areas. Moreover, Appendix Table A.2 shows that our main result holds for lenders who only make auto loans, as well as lenders who make multiple kinds of loans, such as banks. These two groups are likely to have different funding costs on average: the fact that our stylized facts hold within each of these groups suggests that funding costs do not explain our result.

Improved collections technology. Another possible, and closely related channel is that lenders invest in collections technology, rather than screening technology, which improves lenders' recovery rates conditional on default. We do not observe recovery rates in our data, so we cannot test this hypothesis directly. This channel would produce very similar implications to our theory about interest rates: for low credit-score groups, rates could be lower in more concentrated markets, since firms with market power can achieve better recovery rates. However, this channel would not explain why default rates are lower in more concentrated areas, or why loan quantities are lower.

6 Concluding Remarks

This paper presents a model of competition in consumer credit markets, with selection and lender monitoring. The model shows that, in the presence of lender monitoring, the effect of market concentration on prices depends on the riskiness of borrowers. In markets with lower risk borrowers, we see the standard classical relationship: more competition leads to lower prices. However, in markets with a greater portion of high-risk borrowers, increased competition can actually increase prices. We provide empirical support for the model's counterintuitive predictions in the auto loan market: in markets with high-risk borrowers, increased competition is associated with higher prices.

Our results have implications for competition policy in lending markets. Competition appears not to improve market outcomes in subprime credit markets, so antitrust regulators may want to allow some amount of concentration in these markets. Our results also suggest, however, that there is some degree of inefficiency in the industrial organization of these markets:

firms appear to be making screening decisions independently, even though there are returns to scale in screening. The industry as a whole could potentially achieve better outcomes at lower costs if firms could pool efforts in developing screening technologies. It is conceivable that developments in fintech, such as the rise of alternative data companies, could eventually improve the efficiency of screening in these markets.

There remain significant avenues for future work. In terms of empirical work, while we show that greater competition in low credit-score markets can lead to higher prices in the auto loan market, other consumer loan markets such as the credit cards and mortgages may see different patterns. There also exists space to extend the theoretical model. Many consumer loan markets, like the mortgage market and student loan market see a high proportion of loan securitization or guaranteed loans and interest rate subsidies through government programs. This may lead to very different competitive effects, as lenders do not incur direct losses which can impact incentives to monitor (Keys, Mukherjee, Seru and Vig, 2008). Future work should explore how securitization guaranteed lending and subsidies interact with monitoring and competition.

References

Adams, William, Liran Einav, and Jon Levin, "Liquidity Constraints and Imperfect Information in Subprime Lending," *American Economic Review*, 2009, *99* (1), 49–84.

Akerlof, George A., "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 1970, *84* (3), 488–500.

Argyle, Bronson S, Taylor D Nadauld, and Christopher J Palmer, "Monthly Payment Targeting and the Demand for Maturity," *The Review of Financial Studies*, 2020, *33* (11), 5416–5462.

Argyle, Bronson, Taylor D Nadauld, and Christopher Palmer, "Real Effects of Search Frictions in Consumer Credit Markets," Technical Report, National Bureau of Economic Research 2020.

Benmelech, Efraim, Ralf R Meisenzahl, and Rodney Ramcharan, "The Real Effects of Liquidity During the Financial Crisis: Evidence from Automobiles," *The Quarterly Journal of Economics*, 2017, *132* (1), 317–365.

Braxton, J Carter, Kyle F Herkenhoff, and Gordon M Phillips, "Can the Unemployed Borrow? Implications for Public Insurance," *National Bureau of Economic Research*, 2020, (27026).

Broecker, Thorsten, "Credit-Worthiness Tests and Interbank Competition," *Econometrica: Journal of the Econometric Society*, 1990, pp. 429–452.

Buchak, Greg and Adam Jørring, "Do Mortgage Lenders Compete Locally? Evidence Beyond Interest Rates," *Working Paper*, 2021.

Calomiris, Charles W, "Gauging the Efficiency of Bank Consolidation During a Merger Wave," *Journal of Banking & Finance*, 1999, *23* (2-4), 615–621.

Cetorelli, Nicola and Philip E Strahan, "Finance as a Barrier to Entry: Bank Competition and Industry Structure in Local US Markets," *The Journal of Finance*, 2006, *61* (1), 437–461.

Crawford, Gregory S, Nicola Pavanini, and Fabiano Schivardi, "Asymmetric Information and Imperfect Competition in Lending Markets," *American Economic Review*, 2018, *108* (7), 1659–1701.

DeFusco, Anthony, Huan Tang, and Constantine Yannelis, "Measuring the Welfare Effects of Adverse Selection in Consumer Credit Markets," *Working Paper*, 2021.

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl, "The Deposits Channel of Monetary Policy," *The Quarterly Journal of Economics*, 2017, *132* (4), 1819–1876.

_ , _ , **and** _ , "Banking on Deposits: Maturity Transformation Without Interest Rate Risk," Technical Report, National Bureau of Economic Research 2018.

Economides, Nicholas, R. Glenn Hubbard, and Darius Palia, "The Political Economy of

Branching Restrictions and Deposit Insurance: A Model of Monopolistic Competition among Small and Large Banks," *The Journal of Law and Economics*, 1996, *39* (2), 667–704.

Egan, Mark, Ali Hortaçsu, and Gregor Matvos, "Deposit Competition and Financial Fragility: Evidence from the US Banking Sector," *American Economic Review*, 2017, *107* (1), 169–216.

Einav, Liran, Mark Jenkins, and Jonathan Levin, "Contract Pricing in Consumer Credit Markets," *Econometrica*, 2012, *80* (4), 1387–1432.

_ , _ , **and** _ , "The Impact of Credit Scoring on Consumer Lending," *The RAND Journal of Economics*, 2013, *44* (2), 249–274.

Favara, Giovanni and Jean Imbs, "Credit Supply and the Price of Housing," *American Economic Review*, 2015, *105* (3), 958–92.

_ **and Mariassunta Giannetti**, "Forced Asset Sales and the Concentration of Outstanding Debt: Evidence from the Mortgage Market," *The Journal of Finance*, 2017, *72* (3), 1081–1118.

Garmaise, Mark J and Tobias J Moskowitz, "Bank Mergers and Crime: The Real and Social Effects of Credit Market Competition," *The Journal of Finance*, 2006, *61* (2), 495–538.

Giroud, Xavier and Holger M Mueller, "Does Corporate Governance Matter in Competitive Industries?," *Journal of Financial Economics*, 2010, *95* (3), 312–331.

_ **and** _ , "Corporate Governance, Product Market Competition, and Equity Prices," *The Journal of Finance*, 2011, *66* (2), 563–600.

Granja, Joao and Nuno Paixao, "Market Concentration and Uniform Pricing: Evidence from Bank Mergers," *Available at SSRN 3488035*, 2019.

Grunewald, Andreas, Jonathan Lanning, David Low, and Tobias Salz, "Auto Dealer Loan Intermediation: Consumer Behavior and Competitive Effects," *Consumer Financial Protection Bureau Office of Research Working Paper*, 2020, (2020-01).

Hauswald, Robert and Robert Marquez, "Information Technology and Financial Services Competition," *The Review of Financial Studies*, 2003, *16* (3), 921–948.

_ **and** _ , "Competition and Strategic Information Acquisition in Credit Markets," *The Review of Financial Studies*, 2006, *19* (3), 967–1000.

He, Zhiguo, Jing Huang, and Jidong Zhou, "Open Banking: Credit Market Competition When Borrowers Own the Data," Technical Report, National Bureau of Economic Research 2020.

Herkenhoff, Kyle, Gordon Phillips, and Ethan Cohen-Cole, "How Credit Constraints Impact Job Finding Rates, Sorting & Aggregate Output," *National Bureau of Economic Research*, 2019, (22846).

Jayaratne, Jith and Philip Strahan, "The Finance Growth Nexus," *The Quarterly Journal of Economics*, 1996, *111* (3).

Kahn, Charles, George Pennacchi, and Ben Sopranzetti, "Bank Consolidation and the Dynamics of Consumer Loan Interest Rates," *The Journal of Business*, 2005, *78* (1), 99–134.

Keys, Benjamin J, Neale Mahoney, and Hanbin Yang, "What Determines Consumer Financial Distress? Place-and Person-Based Factors," Technical Report, National Bureau of Economic Research 2020.

Keys, Benjamin, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans," *Quarterly Journal of Economics*, 2008, *125* (1), 307–362.

Klee, Elizabeth and Chaehee Shin, "Post-crisis Signals in Securitization: Evidence from Auto ABS," 2020.

Krozner, Randall S. and Philip E. Strahan, "What Drives Deregulation? Economics and Politics of the Relaxation of Bank Branching Restrictions," *The Quarterly Journal of Economics*, 1999, *114* (4), 1437–1467.

Lester, Benjamin, Ali Shourideh, Venky Venkateswaran, and Ariel Zetlin-Jones, "Screening and Adverse Selection in Frictional Markets," *Journal of Political Economy*, 2019, *127* (1), 338–377.

Levitin, Adam J, "The Fast and the Usurious: Putting the Brakes on Auto Lending Abuses," *The Georgetown Law Journal*, 2020, *108*, 1257.

Liebersohn, Jack, "How Does Competition Affect Bank Lending? Quasi-experimental Evidence from Bank Mergers," Technical Report, Working paper, MIT Sloan School of Management 2017.

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

Mahoney, Neale and E Glen Weyl, "Imperfect competition in selection markets," *Review of Economics and Statistics*, 2017, *99* (4), 637–651.

Mayer, Christopher, Edward Morrison, Tomasz Piskorski, and Arpit Gupta, "Mortgage Modification and Strategic Behavior: Evidence from a Legal Settlement with Countrywide," *American Economic Review*, 2014, *104* (9), 2830–57.

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

Parlour, Christine A and Uday Rajan, "Competition in Loan Contracts," *American Economic Review*, 2001, *91* (5), 1311–1328.

Petersen, Mitchell A and Raghuram G Rajan, "The Effect of Credit Market Competition on Lending Relationships," *The Quarterly Journal of Economics*, 1995, *110* (2), 407–443.

Salop, Steven C, "Monopolistic Competition with Outside Goods," *The Bell Journal of Economics*, 1979, pp. 141–156.

Sapienza, Paola, "The Effects of Banking Mergers on Loan Contracts," *The Journal of Finance*, 2002, *57* (1), 329–367.

Scharfstein, David and Adi Sunderam, "Market Power in Mortgage Lending and the Transmission of Monetary Policy," *Unpublished working paper. Harvard University*, 2016, *2.*

Stiglitz, Joseph and Andrew Weiss, "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 1981, *71* (3), 393–410.

Train, Kenneth E, *Discrete Choice Methods With Simulation*, Cambridge university press, 2009.

Vayanos, Dimitri and Jiang Wang, "Liquidity and Asset Returns Under Asymmetric Information and Imperfect Competition," *The Review of Financial Studies*, 2012, *25* (5), 1339–1365.

Veiga, André and E Glen Weyl, "Product Design in Selection Markets," *The Quarterly Journal of Economics*, 2016, *131* (2), 1007–1056.

Figure 1: Model Simulations
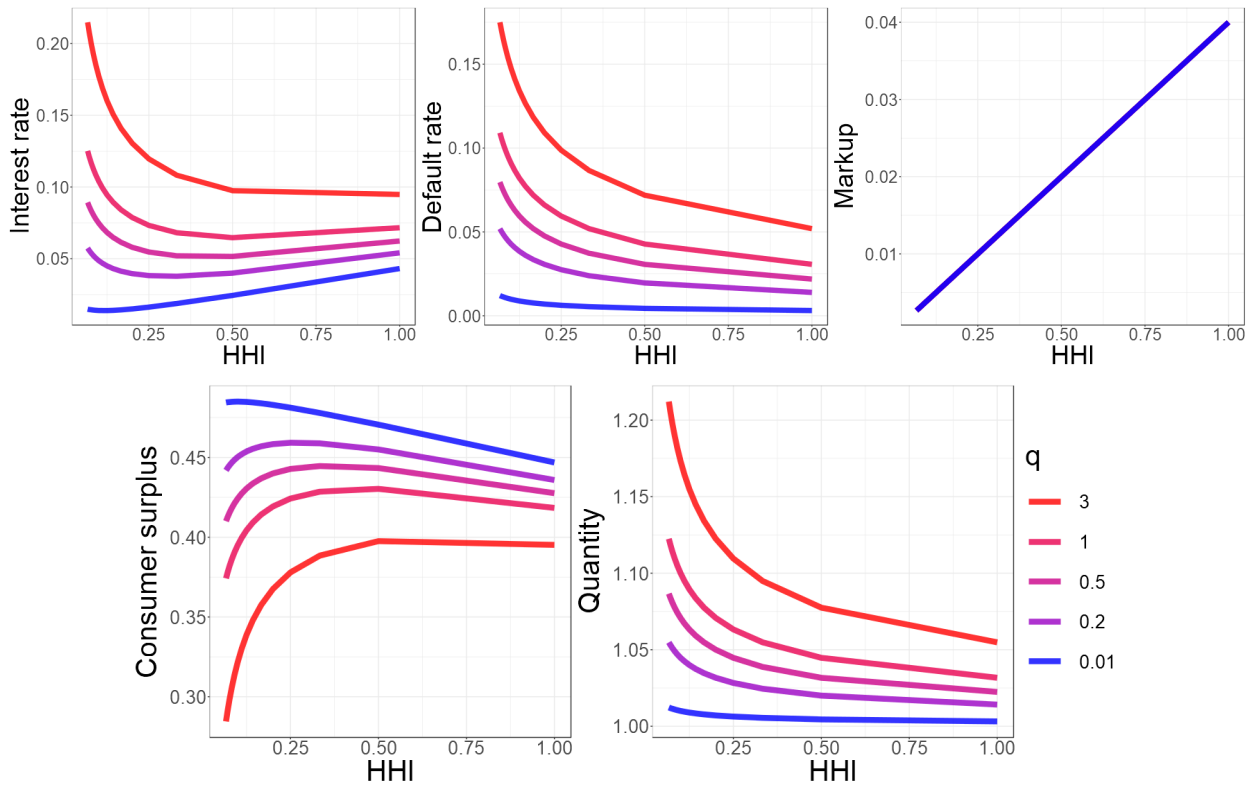
The above figure shows model outcomes, for different values of HHI and q. The top left panel shows equilibrium interest rates, r. The top middle panel shows default rates, δ. The top right panel shows markups over the break-even interest rate, $r - \frac{\delta}{1-\delta}$. The bottom left panel shows consumer surplus, (13). The bottom right panel shows total loan quantity. All simulations use $k = 0.001, \theta = 0.04$.

Figure 2: Interest Rates and Competition
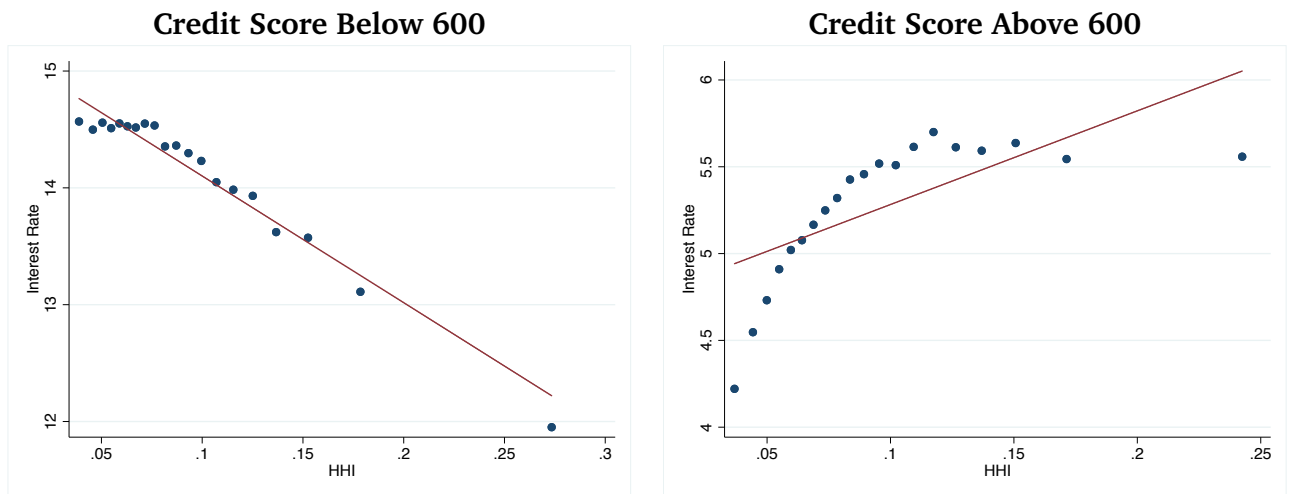
Credit Score Below 600	Credit Score Above 600

The above figure shows median interest rates in a county for given credit score ranges, broken down by ventile of HHI. HHI is defined using the volume of auto loans, that is $HHI = \sum_i^N s_i^2$, where s_i is a lender's share of auto lending in a county within a credit score range. Credit score ranges are denoted above each panel. Credit scores are given by VantageScore ratings. Source: TransUnion.

Figure 3: Interest Rates and Competition
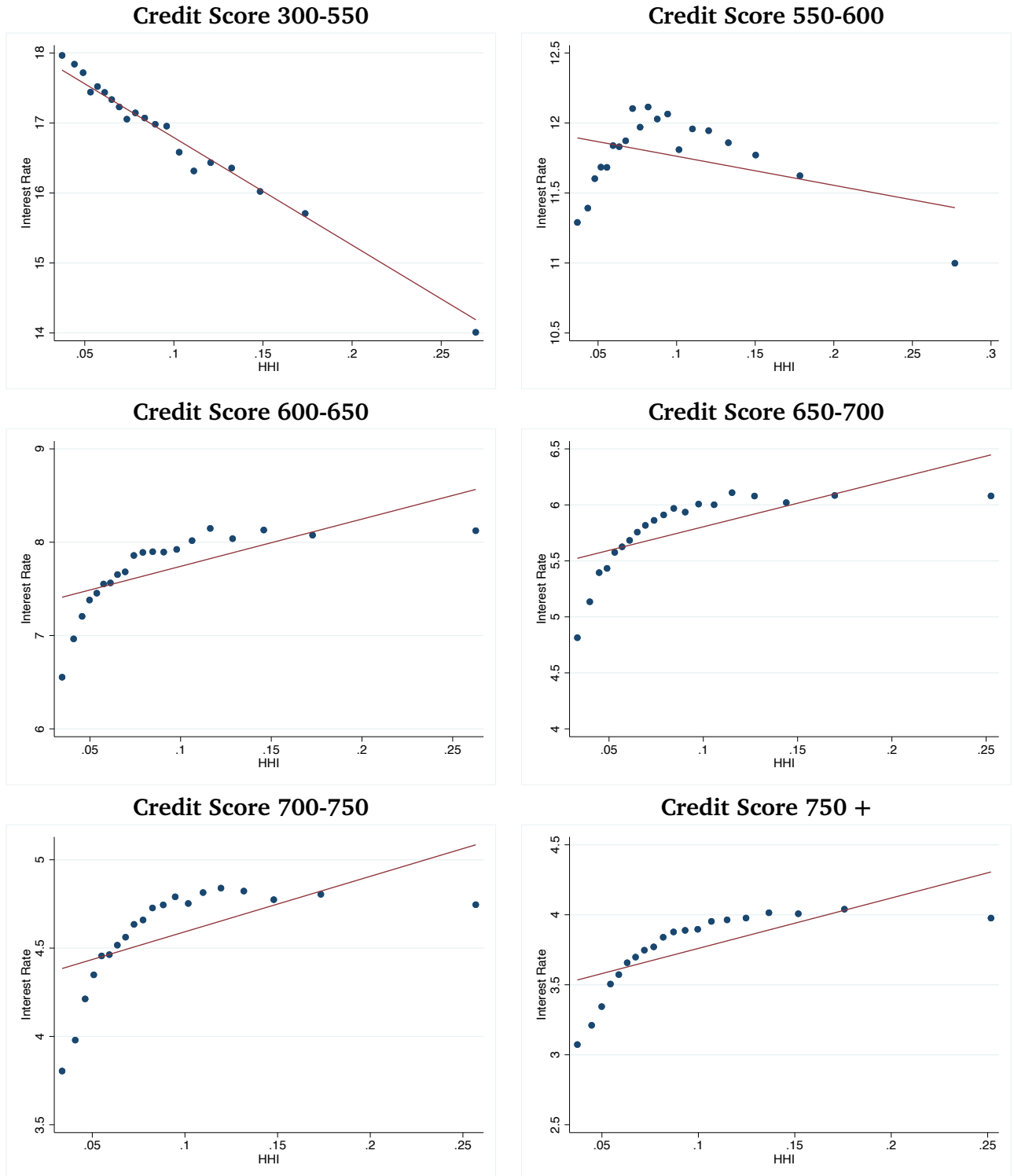
The above figure shows median interest rates in a county for given credit score ranges, broken down by ventile of HHI. HHI is defined using the volume of auto loans, that is $HHI = \sum_i^N s_i^2$, where s_i is a lender's share of auto lending in a county within a credit score range. Credit score ranges are denoted above each panel. Credit scores are given by VantageScore ratings. Source: TransUnion.

Figure 4: Screening and Competition

The above figure shows mean utilization of Creditvision (top row) or mean credit scores (bottom row), broken down by ventile of HHI. In the left panels, HHI is defined using the volume of auto loans, and in the right panels HHI is defined using the number of auto loans; that is, $HHI = \sum_i^N s_i^2$, where s_i is a lender's share of auto lending in a county within a credit score range. Take-up of Creditvision is at the state level. Credit scores are given by average VantageScore ratings in a county. Source: TransUnion.

Figure 5: Delinquency Rates and Competition

Credit Score Below 600

Credit Score Above 600

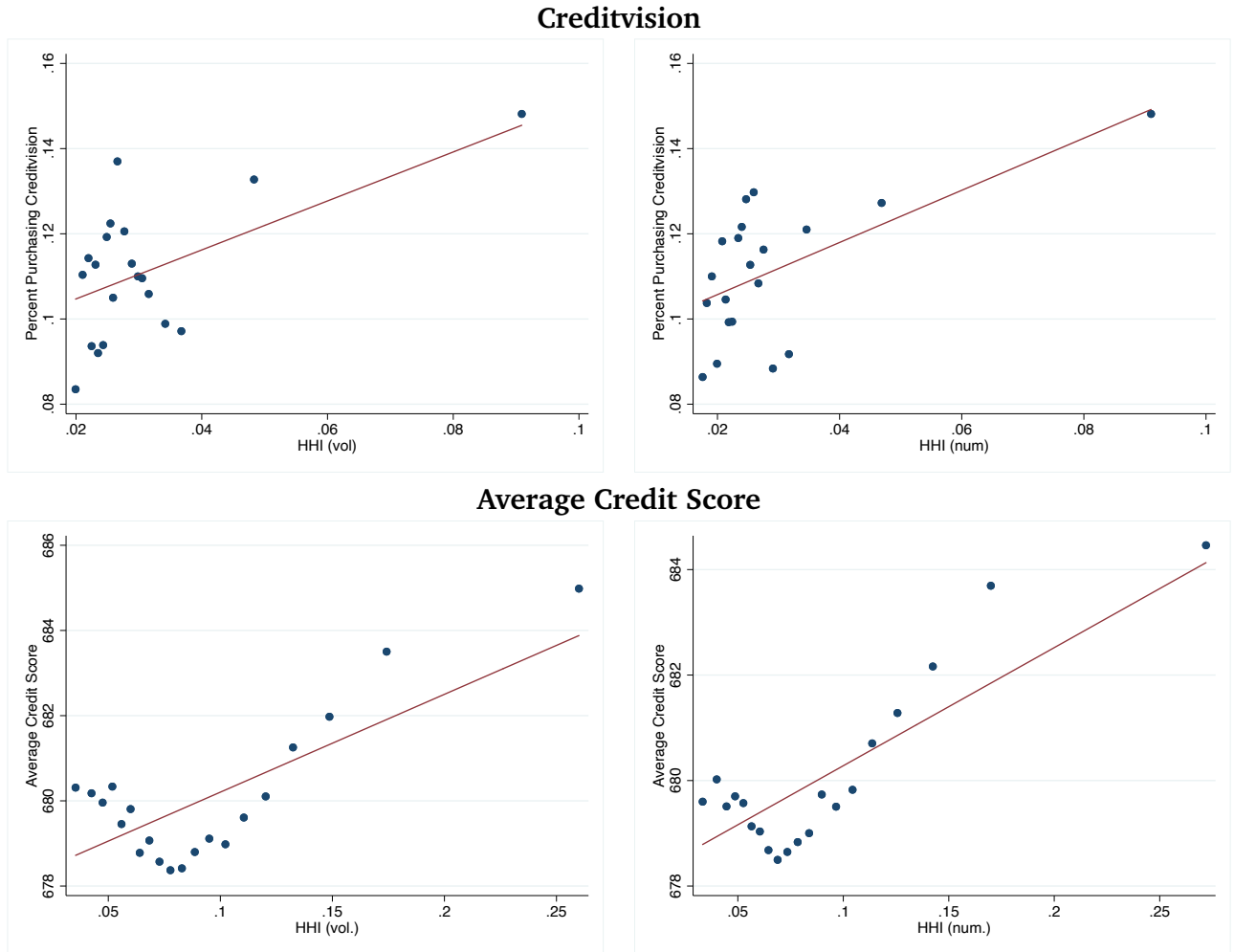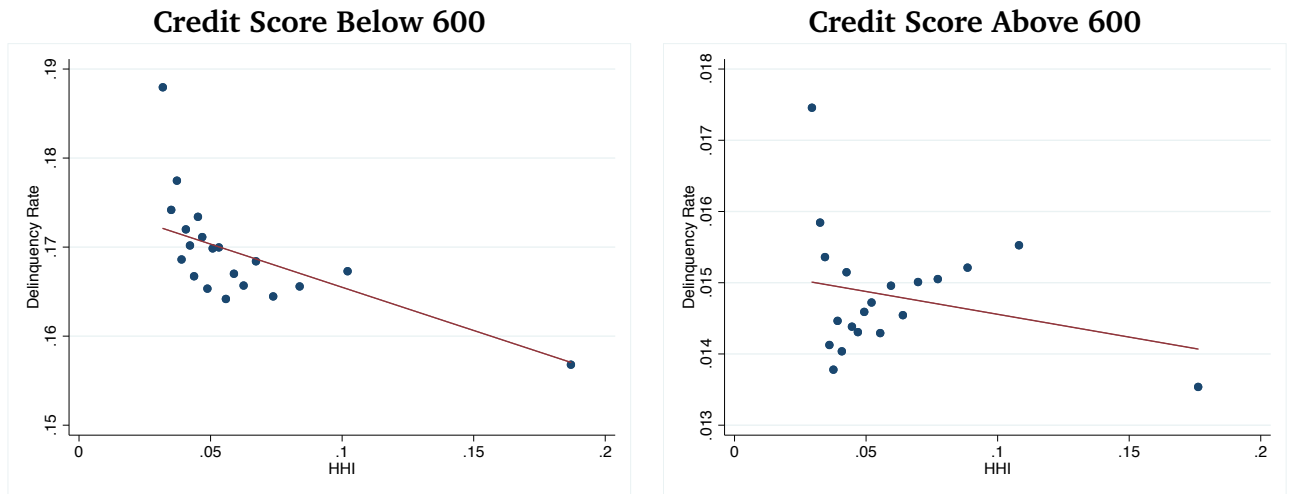The above figure shows mean ninety day delinquency rates in a county for given credit score ranges, broken down by ventile of HHI. HHI is defined using the volume of auto loans, that is $HHI = \sum_i^N s_i^2$, where s_i is a lender's share of auto lending in a county within a credit score range. Credit score rangers are denoted above each panel. Credit scores are given by VantageScore ratings. Source: TransUnion.

Figure 6: Interest Rates and HHI by Credit Score Group: Bank Failures
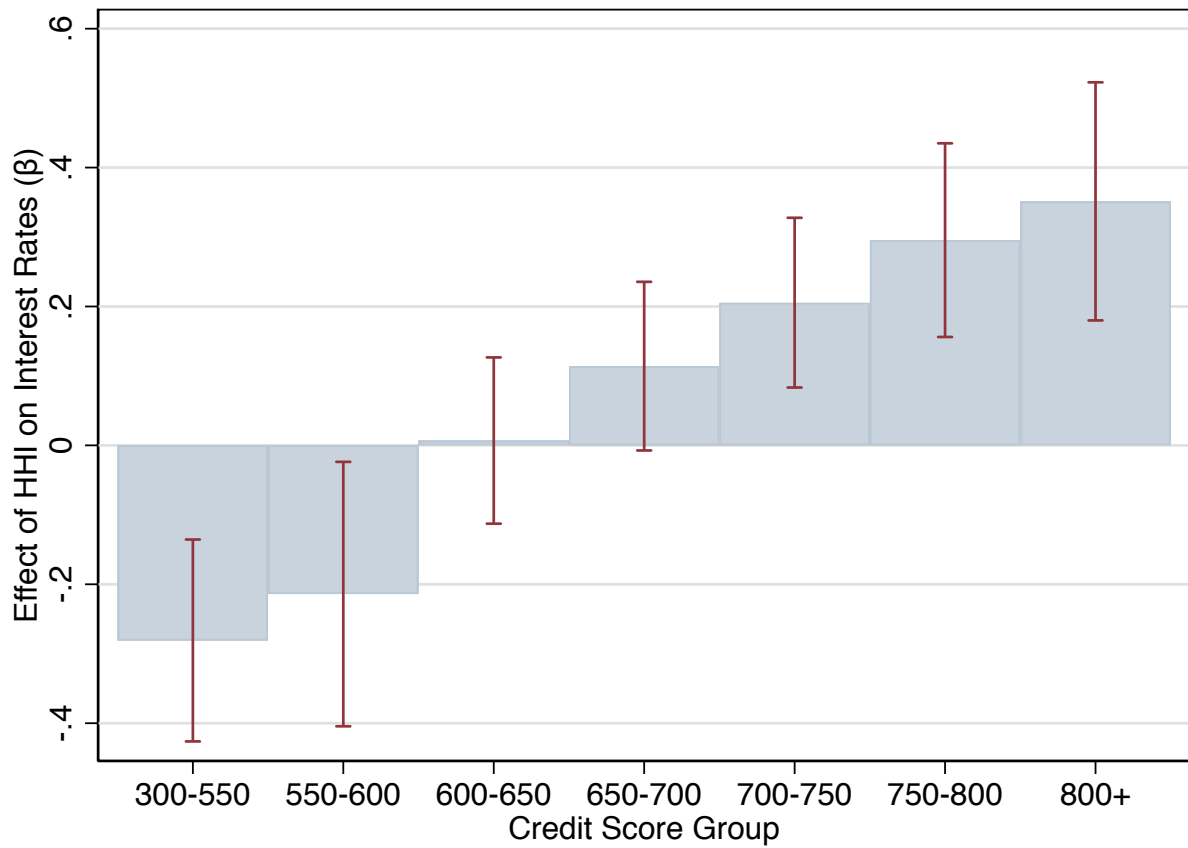
The above figure shows point estimates and a 95% confidence interval from a regression of log interest rates on log HHI. Log HHI is instrumented using the share of deposits from the largest three banks that failed in the financial crisis, Wachovia, Washington Mutual and Countrywide. HHI is defined using the volume of auto loans, that is $HHI = \sum_i^N s_i^2$, where s_i is a lender's share of auto lending in a county within a credit score range. Standard errors are clustered at the county level. Source: TransUnion and Federal Reserve Call Reports.

Figure 7: HHI Following Acquisition Event Study
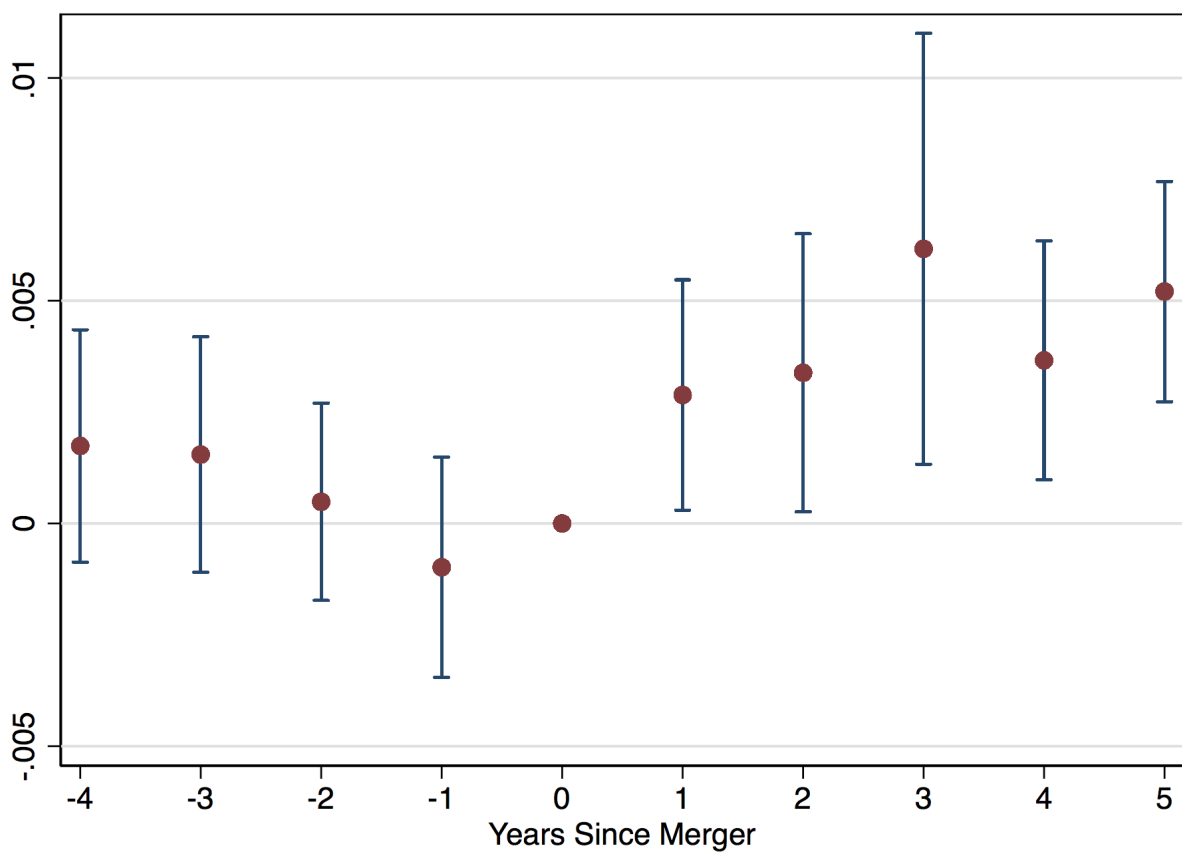
The above figure shows point estimates and a 95% confidence interval from a regression of log HHI on indicators of periods before and after a bank merger, including county, year and merger year fixed effects. HHI is defined using the volume of deposit shares, that is $HHI = \sum_{i}^{N} s_i^2$, where s_i is a lender's share of deposits. Standard errors are clustered at the county level. Source: TransUnion and Federal Reserve Call Reports.

Figure 8: Interest Rates Following Acquisition Events
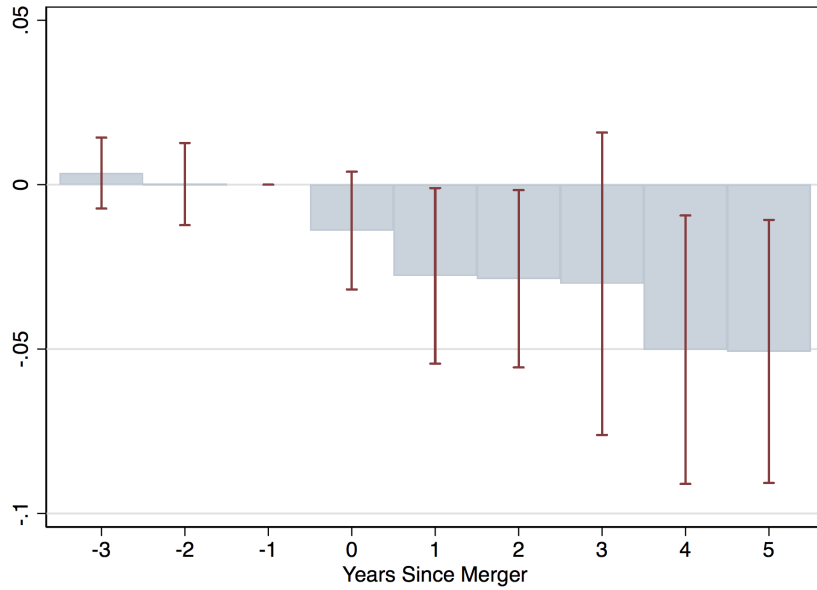
Credit Score 300-600	Credit Score 600-850
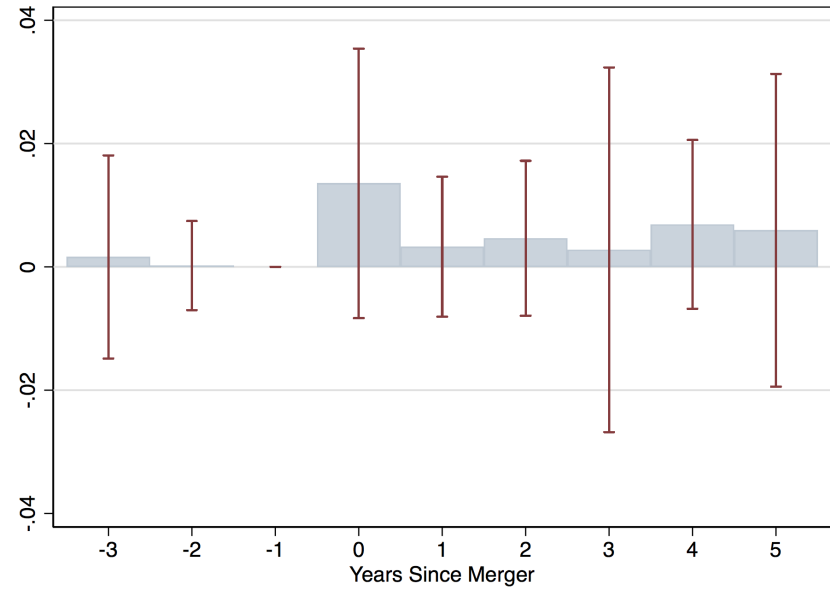
The above figure shows point estimates and a 95% confidence interval from a regression of log interest rates on indicators of periods following a bank merger, including county, year and merger year fixed effects. The left panel shows estimates for borrowers with credit scores below 600, while the right panel shows estimates for borrowers with credit scores above 600. Standard errors are clustered at the county level. Source: TransUnion and Federal Reserve Call Reports.

Figure 9: Interest Rates and HHI by Credit Score Group: Bank Acquisitions
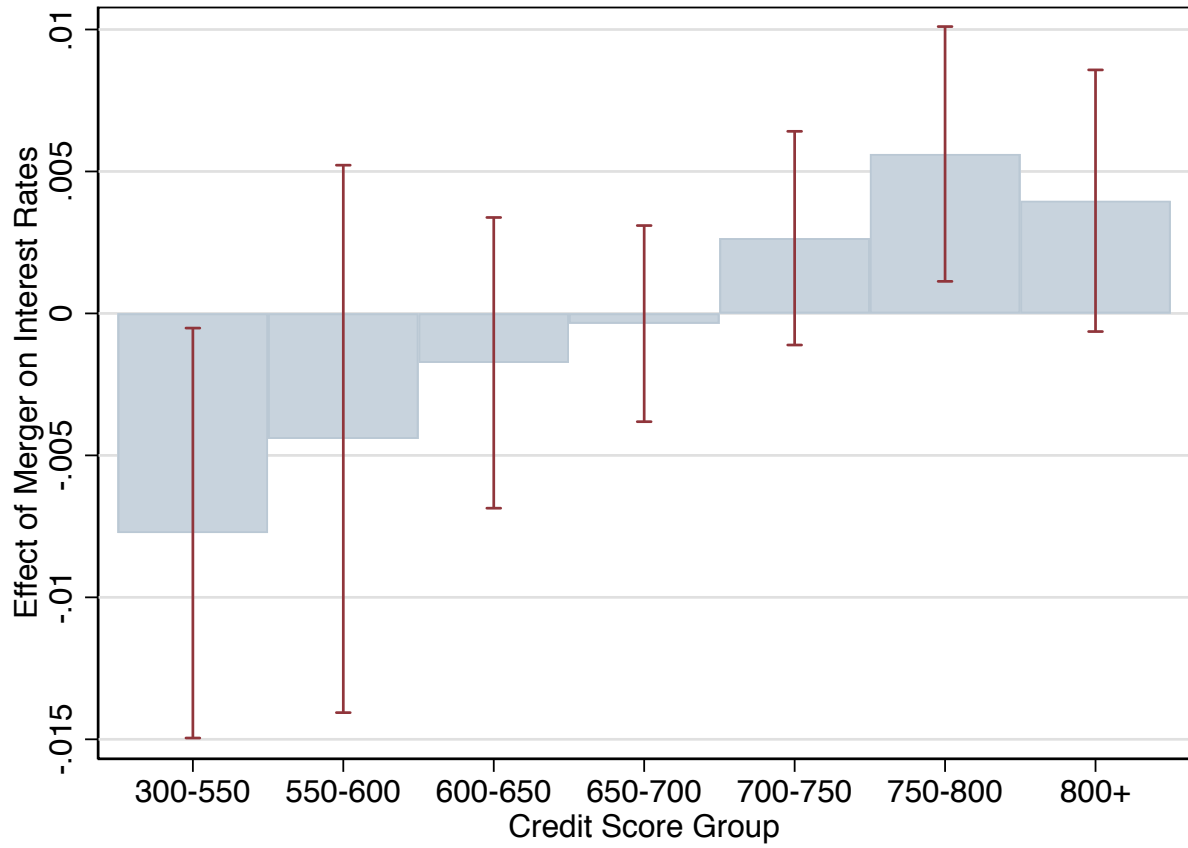
The above figure shows point estimates and a 95% confidence interval from a regression of log interest rates on the number of years since a merger occurred, including county, year and merger year fixed effects. Standard errors are clustered at the county level. Source: TransUnion and Federal Reserve Call Reports.

Table 1: Summary Statistics

This table shows summary statistics for the main analysis variables. Observations are at the county-year level and weighted by the number of observations. Source: TransUnion.

	Minimum	Mean	SD	Max
Loan information				
Principal	$1,477	$21,456	$2,883	$74,779
Monthly Payment	$163	$409	$51	$9,097
Maturity	10	62	3.3	98
Interest Rate	2.2%	7.8%	1.3%	22.0%
Credit Score	436	677	19.4	822
30-day Delinquency	0%	13.5%	6.2%	100.0%
60-day Delinquency	0%	6.9%	3.9%	75.0%
90-day Delinquency	0%	3.4%	2.2%	66.7%
Market information				
HHI (Deals)	0.026	0.058	0.036	1
HHI (Volume)	0.028	0.059	0.034	1
Number of lenders	1	252	195	903
Share of lenders only in auto market	0	44%	6.7%	100%

Table 2: Interest Rates and Market Competition

The table shows the relationship between interest rates and HHI, split by credit score. The top panel splits by above and below a 600 score, while the bottom panel shows a more detailed sample split. The inclusion of fixed effects is denoted beneath each column. Regressions are weighted by the number of loans. Standard errors are clustered at the county level. Credit scores are given by VantageScore ratings. Source: TransUnion. $^{*}p < .1,$ $^{**}p < .05,$ $^{***}p < .01.$

Panel A: Interest Rate and HHI

	(1)	(2)	(3)	(4)	(5)	(6)
			Ln(Interest Rate)			
	Credit Score 300-600			Credit Score 600-850		
Ln(HHI)	-0.0632***	-0.0723***	-0.0641***	0.269**	0.207*	0.134**
	(0.0189)	(0.0200)	(0.0127)	(0.109)	(0.111)	(0.0628)
Year Fixed Effects	No	Yes	Yes	No	Yes	Yes
County Fixed Effects	No	No	Yes	No	No	Yes
Observations	27,887	27,887	27,826	31,773	31,773	31,733
R^2	0.013	0.045	0.571	0.013	0.043	0.845

Panel B: Interest Rate and HHI by Credit Score

	(1)	(2)	(3)	(4)	(5)	(6)
			Ln(Interest Rate)			
	Credit Score 300-550	Credit Score 550-600	Credit Score 600-650	Credit Score 650-700	Credit Score 700-750	Credit Score 750+
Ln(HHI)	-0.0715***	-0.0267***	0.0467*	0.106*	0.0471	0.103
	(0.0129)	(0.00956)	(0.0242)	(0.0541)	(0.0315)	(0.0731)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25,985	27,019	29,162	30,093	29,794	29,989
R^2	0.392	0.565	0.619	0.665	0.760	0.844

Table 3: Interest Rates and Market Competition

The table shows the relationship between interest rates and HHI, interacting HHI with credit scores being below 600. Each column presents a slightly different specification. The first column presents our baseline. The second column does not weight the sample using the number of loans. The third column restricts to counties with more than 25 loan contracts, instead of 10. The fourth column includes all counties. The fifth column winsorizes at the 1% level. The sixth column computes HHI using the number of loans. The seventh column interacts with the number of lenders rather than HHI. The inclusion of fixed effects is denoted beneath each column. Regressions are weighted by the number of loans, except for column (2). Standard errors are clustered at the county level. Credit scores are given by VantageScore ratings. Source: TransUnion. $^*p < .1$, $^{**}p < .05$, $^{***}p < .01$.

	(1) Baseline	(2) Unweighted	(3) Large Counties	(4) All Counties	(5) Winsorized	(6) Loan HHI	(7) Num. Lenders
Ln(HHI)	0.164***	0.118***	0.168***	0.213***	0.0421***		
	(0.0437)	(0.00938)	(0.0459)	(0.0591)	(0.00925)		
Ln(HHI) X CS < 600	-0.428***	-0.406***	-0.429***	-0.442***	-0.396***		
	(0.0106)	(0.00237)	(0.0107)	(0.0146)	(0.00271)		
Ln(HHI$_{Loan}$)						-0.0435	
						(0.0916)	
Ln(HHI$_{Loan}$) X CS < 600						-0.420***	
						(0.0105)	
Ln(N$_{Lender}$)							0.00954
							(0.137)
Ln(N$_{Lender}$) X CS < 600							0.286***
							(0.00898)
County X Credit Score Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	59,620	59,620	49,317	68,482	59,620	59,620	59,620
R^2	0.800	0.768	0.800	0.799	0.968	0.797	0.801

Table 4: Delinquency Rates and Market Competition

The table shows the relationship between ninety day delinquency rates and HHI, split by credit score. The top panel splits by above and below a 600 score, while the bottom panel shows a more detailed sample split. The inclusion of fixed effects is denoted beneath each column. Regressions are weighted by the number of loans. Standard errors are clustered at the county level. Credit scores are given by VantageScore ratings. Source: TransUnion. $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$.

Panel A: Delinquency and HHI

	(1)	(2)	(3)	(4)	(5)	(6)
			Ln(Delinquency)			
	Credit Score 300-600			Credit Score 600-850		
Ln(HHI)	-0.0974**	-0.0534	-0.0290**	-0.00550	-0.0225	-0.0272**
	(0.0381)	(0.0421)	(0.0143)	(0.0202)	(0.0183)	(0.0125)
Year Fixed Effects	No	Yes	Yes	No	Yes	Yes
County Fixed Effects	No	No	Yes	No	No	Yes
Observations	27,887	27,887	27,826	31,773	31,773	31,733
R^2	0.027	0.710	0.827	0.089	0.617	0.825

Panel B: Delinquency and HHI by Credit Score

	(1)	(2)	(3)	(4)	(5)	(6)
			Ln(Delinquency)			
	Credit Score 300-550	Credit Score 550-600	Credit Score 600-650	Credit Score 650-700	Credit Score 700-750	Credit Score 750+
Ln(HHI)	0.00641	-0.0651***	-0.0528***	-0.0543***	-0.0301***	-0.0147**
	(0.0131)	(0.0145)	(0.0151)	(0.0139)	(0.0114)	(0.00746)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25,985	27,019	29,162	30,093	29,794	29,989
R^2	0.762	0.724	0.694	0.621	0.514	0.503

Table 5: Loans and Market Competition

The table shows the relationship between the number of loans rates and HHI, split by credit score. The first two rows split the sample by above and below a 600 score, while columns (3) through. (8) show a more detailed sample split. The inclusion of fixed effects is denoted beneath each column. Standard errors are clustered at the county level. Credit scores are given by VantageScore ratings. Source: TransUnion. $^*p < .1$, $^{**} p < .05$, $^{***} p < .01$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Ln(Loans)				
	Credit Score 300-600	Credit Score 600-850	Credit Score 300-550	Credit Score 550-600	Credit Score 600-650	Credit Score 650-700	Credit Score 700-750	Credit Score 750+
Ln(HHI)	-0.105***	-0.0480***	-0.0663***	-0.0669***	-0.0412***	-0.0350***	-0.0336***	-0.0286***
	(0.0112)	(0.0129)	(0.00873)	(0.00794)	(0.00731)	(0.00740)	(0.00689)	(0.00942)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,826	31,733	25,985	27,019	29,162	30,093	29,794	29,989
R^2	0.974	0.983	0.977	0.979	0.982	0.985	0.986	0.983

Appendix

A Proofs and supplementary material for Section 2

A.1 Optimal price setting

From (5) in the main text, lenders' profits are:

$$\Pi = \left(\frac{1}{1-\delta_j}\right) s_j \left(r_j \left(1-\delta_j\right) - \delta_j\right) = s_j \left(r_j - \frac{\delta_j}{1-\delta_j}\right)$$

As in the main text, lenders' first-order condition for optimal price-setting is:

$$s_j - \frac{\partial s_j}{\partial r_j} \left(r_j - \frac{\delta_j}{1-\delta_j}\right) = 0 \tag{22}$$

Now, in any symmetric equilibrium of the Salop circle model, we have $s_j = \frac{1}{N}$. To calculate the demand slope, $\frac{\partial s_j}{\partial r_j}$, consider a pair of banks j, j' next to each other. Let x represent the distance of a consumer to bank j, so the distance to the neighboring bank is $\frac{1}{N} - x$. Given r_j, r'_j, the set of consumers who choose j satisfies:

$$-r_j - \theta x \leq -r_{j'} - \theta \left(\frac{1}{N} - x\right)$$

The marginal consumer has:

$$x = \frac{1}{2N} + \frac{r_{j'} - r_j}{2\theta}$$

The market share of j takes into account two neighbors j', j'', hence:

$$s_j = 2x = \frac{1}{N} + \frac{r_{j'} + r_{j''} - 2r_j}{2\theta}$$

Hence,

$$\frac{\partial s_j}{\partial r_j} = -\frac{1}{\theta} \tag{23}$$

Thus, expression (22) becomes:

$$r_j - \frac{\delta_j}{1-\delta_j} = \frac{\theta}{N} \tag{24}$$

A.2 Optimal information acquisition

From (7), lenders choose δ_j to solve:

$$\max_{\delta_j} \max_{r_j} s_j(r_j)\left(r_j - \frac{\delta_j}{1-\delta_j}\right) - c_q(\delta_j) \tag{25}$$

We can differentiate (25) using the envelope theorem. By the chain rule,

$$\frac{d}{d\delta_j} \max_{r_j} s_j(r_j)\left(r_j - \frac{\delta_j}{1-\delta_j}\right) - c_q(\delta_j) =$$
$$\frac{\partial}{\partial\delta_j}\left[\max_{r_j} s_j(r_j)\left(r_j - \frac{\delta_j}{1-\delta_j}\right) - c_q(\delta_j)\right] + \frac{dr_j}{d\delta_j}\frac{\partial}{\partial r_j}\left[\max_{r_j} s_j(r_j)\left(r_j - \frac{\delta_j}{1-\delta_j}\right) - c_q(\delta_j)\right]$$

Now, the first-order condition from appendix A.1 implies that

$$\frac{\partial}{\partial r_j}\left[\max_{r_j} s_j(r_j)\left(r_j - \frac{\delta_j}{1-\delta_j}\right) - c_q(\delta_j)\right] = 0$$

Hence,

$$\frac{d}{d\delta_j} \max_{r_j} s_j(r_j)\left(r_j - \frac{\delta_j}{1-\delta_j}\right) - c_q(\delta_j) =$$
$$\frac{\partial}{\partial\delta_j}\left[\max_{r_j} s_j(r_j)\left(r_j - \frac{\delta_j}{1-\delta_j}\right) - c_q(\delta_j)\right] = -\frac{s_j(r_j^*)}{(1-\delta_j)^2} - c_q'(\delta_j)$$

Hence, the first-order condition is:

$$-\frac{s_j}{(1-\delta_j)^2} - c_q'(\delta_j) = 0 \tag{26}$$

Rearranging, we get (8).

A.2.1 Second-order condition

In order for (26) to be a maximum, a second-order condition must also hold. Differentiating again, we have:

$$\frac{\partial^2}{\partial\delta_j^2} \max_{r_j} s_j(r_j)\left(r_j - \frac{\delta_j}{1-\delta_j}\right) - c_q(\delta_j) = \frac{\partial}{\partial\delta_j}\left[-\frac{s_j}{(1-\delta_j)^2} - c_q'(\delta_j)\right]$$

This is:

$$\frac{\partial}{\partial \delta_j}\left[-\frac{s_j}{\left(1-\delta_j\right)^2}-c_q'\left(\delta_j\right)\right]=-\frac{1}{\left(1-\delta_j\right)^2}\frac{\partial s_j}{\partial \delta_j}-\frac{2s_j}{\left(1-\delta_j\right)^3}-c_q''\left(\delta_j\right) \tag{27}$$

This must be negative, at the optimal choice of δ_j for j. Expression (30) of appendix A.2.2 below shows that the derivative of j's optimal market share s_j as δ_j varies, holding fixed all other agents' interest rates and market shares, is:

$$\frac{ds_j}{d\delta_j}=\frac{\theta}{\left(1-\delta_j\right)^2}$$

Substituting this into (27), we get:

$$\frac{1}{\left(1-\delta_j\right)^2}\frac{\theta}{\left(1-\delta_j\right)^2}-\frac{2s_j}{\left(1-\delta_j\right)^3}-c_q''\left(\delta_j\right)<0$$

This rearranges to:

$$c_q''\left(\delta_j\right)\left(1-\delta_j\right)^4+2s_j\left(1-\delta_j\right)>\theta$$

This is (12).

A.2.2 Characterizing $\frac{ds_j}{d\delta_j}$

For analytical convenience, define $\zeta_j \equiv \frac{\delta_j}{1-\delta_j}$, so that (6) becomes:

$$r_j-\zeta_j=\frac{\theta}{N} \tag{28}$$

This gives:

$$\frac{dr_j}{d\zeta_j}=1$$

Now,

$$\frac{dr_j}{ds_j}=\frac{1}{\theta} \tag{29}$$

Hence,

$$\frac{ds_j}{d\zeta_j}=\frac{\frac{dr_j}{d\zeta_j}}{\frac{dr_j}{ds_j}}=\theta$$

Now, from the definition of ζ_j, we have:

$$\frac{d\zeta_j}{d\delta_j} = \frac{1}{\left(1-\delta_j\right)^2}$$

Hence,

$$\frac{ds_j}{d\delta_j} = \frac{ds_j}{d\zeta_j}\frac{d\zeta_j}{d\delta_j} = \frac{\theta}{\left(1-\delta_j\right)^2} \tag{30}$$

A.3 Consumer surplus

In equilibrium, the average consumer's distance from a lender is:

$$\frac{1}{4N}$$

Plugging into expression (1) for consumer welfare, expected welfare of type G consumers is:

$$\mu - r - \frac{\theta}{4N}$$

A.4 Nonzero Funding Costs, Nonzero Recovery Rates

In the main text, for expositional simplicity we assume that the cost of funds and the recovery rate are both zero. We also assume there are no variable costs of screening. In this section, we relax each of those assumptions and show that the main model predictions hold. Suppose lenders have some positive variable cost ρ for each loan they make. ρ could reflect lenders' funding costs, or a component of screening costs which is variable and scales with the number of consumers. Suppose also that recovery rates are nonzero: lenders can recover $1 - \phi$ on average when borrowers default. Lenders' profits are thus:

$$\Pi = \left(\frac{1}{1-\delta_j}\right)s_j\left(r_j\left(1-\delta_j\right) - \delta_j\phi - \rho\right) \tag{31}$$

In words, (31) says that lenders have to pay ρ to borrow a unit of funds to lend to customers. With probability δ_j, the borrower defaults and the lender loses a fraction ϕ of the principal, and with probability $\left(1-\delta_j\right)$ the borrower pays r_j to the lender. Profits rearrange to:

$$\Pi = s_j\left(r_j - \frac{\delta_j\phi + \rho}{1-\delta_j}\right)$$

Lender j's optimal markup thus satisfies:

$$r_j - \frac{\delta_j \phi + \rho}{1 - \delta_j} = \frac{\theta}{N} \tag{32}$$

Comparing (32) to the markup equation (6) in the main text, lenders simply set markups above a different marginal cost, $\frac{\delta_j \phi + \rho}{1 - \delta_j}$, which reflects lenders' cost of funds and expected recovery rates. All other equilibrium conditions are unchanged.

A.5 Entry costs

In the baseline model, we have taken the number of firms N as exogenous. N can be micro-founded using a simple extension to the model with fixed entry costs. Fixed entry costs may differ across markets due to heterogeneous regulatory intensity, labor and rent costs, and other such factors.

Suppose there are a countably infinite number of potential entrants, who are exogenously ordered. In the first stage, firms sequentially decide whether to pay entry cost C_e to enter the market. If N firms enter, they are uniformly spaced around a Salop circle, as in the main text. Firms then play the screening and price-setting game in the baseline model: firms decide how much to invest in costly screening, and then set prices.

If there are N entrants, firms' profits, net of screening costs, are:

$$\frac{1}{1 - \delta(N)} \left(\frac{\theta}{N} \right) - c_q(\delta(N))$$

where $\delta(N)$ is the solution to (11), the equilibrium amount of screening done if there are N firms. Firms will enter until the marginal entrant's expected profit is negative. Hence, for any entry cost C_e, the equilibrium number of firms $N(C_e)$ satisfies:

$$\frac{1}{1 - \delta(N(C_e))} \left(\frac{\theta}{N(C_e)} \right) - c_q(\delta(N(C_e))) > C_e \tag{33}$$

$$\frac{1}{1 - \delta(N(C_e) + 1)} \left(\frac{\theta}{N(C_e) + 1} \right) - c_q(\delta(N(C_e) + 1)) < C_e \tag{34}$$

That is, firms make profits greater than C_e with $N(C_e)$ entrants, but not with $N(C_e)+1$ entrants. Using expressions (33) and (34), we can simulate the equilibrium number of entrants as a function of C_e. The results are shown in Appendix Figure A.1. When entry costs are higher, the equilibrium number of entrants decreases.

A.6 Moral hazard

Here, we consider how our model's conclusions change if there is moral hazard. As in the baseline model, suppose that type-B borrowers always default. However, suppose that type-G borrowers also default with some probability $\phi(r)$, which is an increasing function of the interest rate r that they face. Banks can invest to screen out type-B borrowers, as in the main text. If there are a fraction δ of type-B borrowers and the interest rate is r, the population default rate is thus:

$$\psi(r,\delta) \equiv \delta + (1-\delta)\phi(r) \tag{35}$$

If a lender charges interest rate r and has market share s_j, her expected profits are:

$$\left(\frac{1}{1-\delta_j}\right)s_j\left(r_j\left(1-\psi\left(r,\delta_j\right)\right)-\psi\left(r,\delta_j\right)\right)$$

In words, the lender faces a quantity $\frac{s_j}{1-\delta_j}$ of consumers. The default rate among consumers is $\psi(r,\delta_j)$. Thus, with probability $1-\psi(r,\delta_j)$, the lender is paid r_j, and with probability $\psi(r,\delta_j)$, the lender loses the principal and the interest payment.

Price setting. Conditional on δ_j, the lender chooses r_j to maximize:

$$s_j\left(r_j\left(1-\psi\left(r_j,\delta_j\right)\right)-\psi\left(r_j,\delta_j\right)\right)$$

We can write this as:

$$\underbrace{\left(1-\psi\left(r_j,\delta_j\right)\right)s_j}_{\text{Market size}}\underbrace{\left(r_j-\frac{\psi\left(r_j,\delta_j\right)}{1-\psi\left(r_j,\delta_j\right)}\right)}_{\text{Markup}}$$

Differentiating with respect to r, we have:

$$-\frac{d\psi}{dr_j}s_j\left(r_j-\frac{\psi\left(r_j,\delta_j\right)}{1-\psi\left(r_j,\delta_j\right)}\right)+\left(1-\psi\left(r_j,\delta_j\right)\right)s_j'\left(r_j-\frac{\psi\left(r_j,\delta_j\right)}{1-\psi\left(r_j,\delta_j\right)}\right)+$$

$$\left(1-\psi\left(r_j,\delta_j\right)\right)s_j-\left(1-\psi\left(r_j,\delta_j\right)\right)s_j\left(\frac{\partial}{\partial r_j}\left(\frac{\psi\left(r_j,\delta_j\right)}{1-\psi\left(r_j,\delta_j\right)}\right)\right)=0$$

Using (23) from Appendix A.1, we have $s'_j = -\frac{1}{\theta}$. Rearranging,

$$\left(r_j - \frac{\psi(r_j, \delta_j)}{1 - \psi(r_j, \delta_j)}\right)\left(-(1 - \psi(r_j, \delta_j))\frac{1}{\theta} - \frac{d\psi}{dr_j}s_j\right) =$$
$$-\left[(1 - \psi(r_j, \delta_j))s_j - (1 - \psi(r_j, \delta_j))s_j\left(\frac{\partial}{\partial r_j}\left(\frac{\psi(r_j, \delta_j)}{1 - \psi(r_j, \delta_j)}\right)\right)\right] \quad (36)$$

Now, we have:

$$\frac{d\psi}{dr_j} = (1 - \delta_j)\frac{d\phi}{dr_j}$$

$$\frac{\partial}{\partial r_j}\left(\frac{\psi(r_j, \delta_j)}{1 - \psi(r_j, \delta_j)}\right) = \frac{1 - \delta_j}{(1 - \psi)^2}\frac{d\phi}{dr_j}$$

Hence, (36) becomes:

$$\left(r_j - \frac{\psi(r_j, \delta_j)}{1 - \psi(r_j, \delta_j)}\right)\left(-(1 - \psi(r_j, \delta_j))\frac{1}{\theta} - (1 - \delta_j)\frac{d\phi}{dr_j}s_j\right) =$$
$$-\left[(1 - \psi(r_j, \delta_j))s_j - (1 - \psi(r_j, \delta_j))s_j\left(\frac{1 - \delta_j}{(1 - \psi(r_j, \delta_j))^2}\frac{d\phi}{dr_j}\right)\right]$$

Rearranging, we have:

$$r_j - \frac{\psi(r_j, \delta_j)}{1 - \psi(r_j, \delta_j)} =$$
$$\left((1 - \psi(r_j, \delta_j))\frac{1}{\theta} + (1 - \delta_j)\frac{d\phi}{dr_j}s_j\right)^{-1}\left[1 - \frac{1 - \delta_j}{(1 - \psi)^2}\frac{d\phi}{dr_j}\right](1 - \psi(r_j, \delta_j))s_j \quad (37)$$

When there is no moral hazard, $\phi(r) = 0$, so $\frac{d\phi}{dr_j} = 0$, so (37) reduces to (6) in the main text. When there is moral hazard, the RHS of (37) tends to be lower, so lenders' markups above the break-even interest rate are lower. Intuitively, increasing interest rates increases costs, so lenders set lower markups in response.

Screening incentives. Recall we defined:

$$\psi(r, \delta) = \delta + (1 - \delta)\phi(r)$$

This gives:

$$\frac{\partial \psi}{\partial \delta} = 1 - \phi(r)$$

Lenders' optimal profits are the solution to:

$$\max_{\delta_j} \max_{r_j} \left(\frac{1}{1-\delta_j}\right) s_j \left(r_j \left(1 - \psi(r, \delta_j)\right) - \psi(r, \delta_j)\right) - c_q(\delta_j) \tag{38}$$

Using (35), the objective in (38) becomes:

$$\left(\frac{1}{1-\delta_j}\right) s_j \left(r_j \left(1 - \left(\delta_j + \left(1-\delta_j\right)\phi(r)\right)\right) - \left(\delta_j + \left(1-\delta_j\right)\phi(r)\right)\right) - c_q(\delta_j)$$

This rearranges to:

$$= s_j \left(r_j - \frac{\delta_j}{1-\delta_j}\right) + s_j \left(-\phi\left(1+r_j\right)\right) - c_q(\delta_j) \tag{39}$$

As in Appendix A.2, the envelope theorem applies, so lenders' investment incentives are determined by differentiating (39) with respect to δ_j, holding s_j and r_j fixed. This means that optimal investments are determined by the first-order-condition:

$$\frac{-s_j}{\left(1-\delta_j\right)^2} - c_q'(\delta_j) = 0$$

which is identical to (8) in the main text. Hence, if there is moral hazard for the type-G consumers, this actually does not change incentives for screening out type-B consumers, in this model.

Simulations. To demonstrate the effects of moral hazard on outcomes, we simulate lenders' price-setting decisions. Since screening incentives are unchanged from the main text, we will simply hold default rates δ_j contant. We will assume that moral hazard has a linear effect on type-G customers' default rates:

$$\phi(r) = \chi r$$

This gives:

$$\psi(r_j, \delta_j) = \delta_j + \left(1-\delta_j\right)\chi r$$

Plugging in to (37), we have:

$$r_j - \frac{\psi\left(r_j, \delta_j\right)}{1 - \psi\left(r_j, \delta_j\right)} =$$

$$\left(\left(1 - \psi\left(r_j, \delta_j\right)\right)\frac{1}{\theta} + \left(1 - \delta_j\right)\chi s_j\right)^{-1}\left[1 - \frac{1 - \delta_j}{\left(1 - \psi\right)^2}k\right]\left(1 - \psi\left(r_j, \delta_j\right)\right)s_j \quad (40)$$

In any symmetric equilibrium, we have $s_j = \frac{1}{N}$. Using (40), we can simulate markups, for different values of χ. Figure A.2 shows how markups vary with χ. When χ is larger, so moral hazard is a greater concern, markups are lower for any given value of N, because banks internalize the fact that raising interest rates tends to increase default rates.

A.7   Winner's curse

In this appendix, we relax the assumption that the results of lenders' screening decisions are perfectly correlated. For simplicity, we assume that each customer can only take loans from the two banks on the Salop circle she is nearest to: the transportation costs to other lenders are high enough that there is no interest rate she is willing to borrow at. This simplifies the derivations, since each customer only has two choice of lenders.

Rather than the default rate δ_j, it is convenient to work with measures of type-B consumers. There is a total measure q of type-B consumers. As in the main text, suppose that each lender can invest $\tilde{c}\left(\alpha_j\right)$ to create a test that perfectly identifies a fraction $1 - \alpha_j$ of type-B consumers, out of a total measure q, and screens them out. In contrast to the main text, we assume that test results are independent across banks. From the perspective of a given bank, consider a customer located between lender j and her neighbor j'. There are four types of type-B consumers, who appear with the following probabilities.

1. $\alpha_j \alpha_{j'}$: Pass both tests

2. $\alpha_j\left(1 - \alpha_{j'}\right)$: Pass j's test, but not my neighbor's

3. $\left(1 - \alpha_j\right)\alpha_{j'}$: Pass my neighbor's test, but not mine

4. $\left(1 - \alpha_j\right)\left(1 - \alpha_{j'}\right)$: Fail both tests

Categories 3 and 4 of consumers have failed j's test, so j knows that they are type-B consumers and never lends to them. Category 2 of consumers has failed j''s test, so j' does not lend to them: hence, these customers will borrow from j at any price. Category 1 of consumers passes both tests, so they are price sensitive: they will choose the bank that offers higher utility, net of transportation costs.

Let $s_i(r_i)$ represent lender j's market share, among price-elastic consumers. Suppose all lenders are choosing some screening probability α_{-j}. The profit of lender j an be written as:

$$\Pi_i = \underbrace{s_i(r_i)r_i}_{Good\ type\ profits} - \underbrace{s_i(r_i)q\alpha_j\alpha_{-j}}_{Group\ 1\ losses} - \underbrace{\frac{2}{N}q\alpha_j(1-\alpha_{-j})}_{Group\ 2\ losses} - \tilde{c}(\alpha_j) \tag{41}$$

In words, for the measure $s_i(r_i)$ of price-elastic type-G consumers, the lender makes r_i. For the measure $q\alpha_j\alpha_{-j}$ of type-1 consumers, who have passed j's test as well as the neighbors, j loses the principal. There is a measure $\frac{2}{N}\alpha_j(1-\alpha_{-j})$ of type-B consumers who have passed j's test, but not the neighbor's (note the 2 is because j has two neighbors), and j loses the principal on these. Finally, j pays the screening cost $\tilde{c}(\alpha_j)$.

Price-setting. Differentiating (41) with respect to r_i, we have:

$$s_i'(r_i - q\alpha_j\alpha_{-j}) + s_i = 0 \tag{42}$$

Using (23) from Appendix A.1, we have $s_j' = -\frac{1}{\theta}$. Hence, (42) rearranges to:

$$r_i - q\alpha_i\alpha_j = \theta s_i \tag{43}$$

(43) is effectively a markup formula. The marginal cost of increasing lending is equal to the fraction of type-B consumers, among consumers who are marginal with respect to price. Among price-elastic consumers, for every type-G, there are $q\alpha_i\alpha_j$ type-B's. Hence, this is the relevant marginal cost for lenders' markups. In symmetric equilibrium, $s_i = \frac{1}{N}$, so markups over marginal costs still decrease as N increases.

Screening. To solve for optimal screening decisions, we differentiate (41) with respect to α_j. Using the envelope theorem, we can ignore effects of changes in r_j. This gives:

$$\frac{d\Pi_i}{d\alpha_i} = -s_i(r_i)q\alpha_{-j} - \frac{2q(1-\alpha_j)}{N} - \tilde{c}'(\alpha_j)$$

$$-\tilde{c}'(\alpha_j) = s_i(r_i)q\alpha_{-j} + \frac{2q(1-\alpha_j)}{N}$$

In symmetric equilibrium, $s_i(r_i) = \frac{1}{N}$, hence we have:

$$-\tilde{c}'(\alpha_j) = \frac{(2-\alpha_{-j})q}{N} \tag{44}$$

(44) says that, when α_{-j} is lower – when j's neighbors screen more intensely – j also increases

screening intensity. This is because customers who fail j's neighbors' tests will tend to borrow from j, which increases j's incentives to invest in screening. However, as in the baseline model, (44) shows that screening incentives are lower when N is larger. Hence, both forces in the baseline model are still present when we assume lenders face a winner's curse.

A.8 Logit demand model

In this appendix, we consider an alternative model of preferences, the logit model, and show that our results still hold in this setting. As in the main text, there is a unit mass of type-G consumers, and some measure of type-B consumers. The screening technology is identical to the main text: lenders pay a fixed cost $c_q(\delta)$ to lower the population default rate to δ. As in the main text, the willingness-to-pay of consumers for loans is independent of whether they are type B or G. Unlike the main text, we assume consumers' preferences over banks are described by a logit model. The utility of consumers' outside option, of not borrowing, is normalized to 0. The utility that consumer i attains if she borrows from lender j, at loan rate r_j, is:

$$\mu - \alpha r_j + \epsilon_{ij} \tag{45}$$

where ϵ_{ij} is i.i.d. type-1 extreme value. Hence, consumers have logit demand, with mean utility μ for borrowing. μ can be thought of as consumers' mean utility for car loans, relative to the outside option. α determines how sensitive consumers are to interest rates. ϵ_{ij} is an idiosyncratic preference that consumer i has for lender j.[25] Given lenders' interest rates r_j, the market share of lender j is:

$$s_j = \frac{\exp(\mu - \alpha r_j)}{1 + \sum_j \exp(\mu - \alpha r_j)} \tag{46}$$

Lenders' profits are still, as in the main text:

$$\Pi = \left(\frac{1}{1-\delta_j}\right) s_j \left(r_j(1-\delta_j) - \delta_j\right) = s_j \left(r_j - \frac{\delta_j}{1-\delta_j}\right)$$

Appendix A.9.1 shows that lenders' optimal prices satisfy:

$$r_j - \frac{\delta_j}{1-\delta_j} = \frac{1}{\alpha(1-s_j)} \tag{47}$$

---

[25]Technically, idiosyncratic terms are needed in Bertrand models so that demand is not perfectly elastic, so firms set prices above marginal cost in equilibrium. In our setting, these preference shocks could represent either consumers' actual preferences over lenders, or could represent in reduced-form dealers' relationships with lenders, or consumer search costs.

The intuition for (47) is that $r_j - \frac{\delta_j}{1 - \delta_j}$, the markup of r_j over the break-even interest rate, $\frac{\delta_j}{1 - \delta_j}$, is higher when consumers' price sensitivity, α, is lower, and when firm j's market share s_j is higher.

Appendix A.9.2 shows that lenders' FOC for optimal information acquisition is unchanged from the main text:

$$\frac{s_j}{\left(1 - \delta_j\right)^2} = -c_q'\left(\delta_j\right) \tag{48}$$

However, the second-order condition changes to:

$$c_q''(\delta)(1 - \delta)^4 + 2s(1 - \delta) > \frac{\alpha}{\frac{1}{s(1-s)} + \frac{1}{(1-s)^2}} \tag{49}$$

A symmetric equilibrium is described by a triple s, r, δ which solves (46), (47), and (48).[26] Equilibrium consumer surplus, for type G consumers, follows from the standard logit surplus formula, from, for example, Train (2009) :

$$\log\left(1 + \sum_{j=1}^{J} \exp\left(\mu - \alpha r\right)\right) + C \tag{50}$$

where C is a constant.

We proceed to solve the model numerically. As in the main text, we parametrize costs as:

$$\tilde{c}(\alpha) = \frac{k}{\alpha}$$

implying that:

$$c_q\left(\delta_j\right) = \frac{kq\left(1 - \delta_j\right)}{\delta_j}$$

Figure A.3, analogous to Figure 1, shows equilibrium outcomes, as we vary the number of lenders. The findings are identical to the main text: Concentration increases markups, but decreases default rates. The net effect of concentration on interest rates is positive when k is lower and the population is low-risk, and negative when k is high and the population is high-risk. Quantities always tend to decrease as concentration increases. The main difference of the logit model to the Salop circle model is that consumer surplus tends to increase as markets

---

[26]We note that these conditions are necessary, but not sufficient, for equilibrium, and that the model does not always have unique equilibria, even when $c_q(\cdot)$ is convex. Intuitively, this is because, in the logit model, costly information acquisition creates increasing returns: if firms acquire more information, default rates δ are lower, allowing firms to charge lower interest rates, which then increases market shares and increases firms' incentives for information acquisition. Moreover, nontrivial equilibria are not guaranteed to exist: for some choices of $c_q(\cdot)$, there is no $\delta < 1$ which satisfies all three conditions simultaneously.

become more competitive. Intuitively, this is because gains-from-variety are much larger in the logit model than in the Salop circle model, so there is a stronger mechanical effect of increasing N on consumer surplus. However, when k is very large – the blue line – increasing N can still decrease consumer surplus in the logit model.

A.9   Proofs for Appendix A.8

A.9.1   Optimal price setting

From (46), market shares are:

$$s_j = \frac{\exp(\mu - \alpha r_j)}{1 + \sum_j \exp(\mu - \alpha r_j)}$$

The derivative with respect to the interest rate r_j is:

$$\frac{\partial s_j}{\partial r_j} = -\alpha s_j (1 - s_j)$$

From (5), lenders choose interest rates to solve:

$$\max_{r_j} (s_j(r_j)) \left(r_j - \frac{\delta_j}{1 - \delta_j} \right)$$

Differentiating with respect to r_j, we have:

$$s_j - \frac{\partial s_j}{\partial r_j} \left(r_j - \frac{\delta_j}{1 - \delta_j} \right) = 0$$

$$s_j - \alpha s_j (1 - s_j) \left(r_j - \frac{\delta_j}{1 - \delta_j} \right) = 0$$

This simplifies to:

$$r_j - \frac{\delta_j}{1 - \delta_j} = \frac{1}{\alpha (1 - s_j)}$$

This is (47).

A.9.2   Optimal information acquisition

From (7), lenders choose δ_j to solve:

$$\max_{\delta_j} \max_{r_j} s_j(r_j) \left(r_j - \frac{\delta_j}{1 - \delta_j} \right) - c_q(\delta_j) \tag{51}$$

64

As in the baseline model, the envelope theorem gives:

$$\frac{d}{d\delta_j}\left[\max_{r_j} s_j(r_j)\left(r_j - \frac{\delta_j}{1-\delta_j}\right)\right] = \frac{\partial}{\partial\delta_j}\left[s_j(r_j^*)\left(r_j^* - \frac{\delta_j}{1-\delta_j}\right)\right]$$

That is, we can hold the interest rate r_j^* constant at the optimal choice when taking the derivative. Thus, we have:

$$\frac{\partial}{\partial\delta_j}\left[s_j(r_j^*)\left(r_j^* - \frac{\delta_j}{1-\delta_j}\right)\right] = -\frac{s_j(r_j^*)}{(1-\delta_j)^2}$$

Note that this is always negative, since $s_j > 0$. That is, higher default rates always decrease the objective. Combining this with the derivative of $c_q(\delta_j)$, we get the first-order condition:

$$-\frac{s_j}{(1-\delta_j)^2} - c_q'(\delta_j) = 0 \tag{52}$$

Rearranging, we get (8).

Second-order condition. In order for (52) to be a maximum, a second-order condition must also hold. Differentiating again, we have:

$$\frac{\partial^2}{\partial\delta_j^2}\max_{r_j} s_j(r_j)\left(r_j - \frac{\delta_j}{1-\delta_j}\right) - c_q(\delta_j) = \frac{\partial}{\partial\delta_j}\left[-\frac{s_j}{(1-\delta_j)^2} - c_q'(\delta_j)\right]$$

This is:

$$\frac{\partial}{\partial\delta_j}\left[-\frac{s_j}{(1-\delta_j)^2} - c_q'(\delta_j)\right] = -\frac{1}{(1-\delta_j)^2}\frac{\partial s_j}{\partial\delta_j} - \frac{2s_j}{(1-\delta_j)^3} - c_q''(\delta_j) \tag{53}$$

This must be negative, at the optimal choice of δ_j for j. Expression (30) of appendix A.2.2 below shows that the derivative of j's optimal market share s_j as δ_j varies, holding fixed all other agents' interest rates and market shares, is:

$$\frac{ds_j}{d\delta_j} = \left(\frac{-\alpha}{(1+\delta_j)^2}\right)\left(\frac{1}{\frac{1}{s_j(1-s_j)} + \frac{1}{(1-s_j)^2}}\right)$$

Substituting this into (53), we get:

$$\frac{1}{(1-\delta_j)^2}\left(\frac{\alpha}{(1+\delta_j)^2}\right)\left(\frac{1}{\frac{1}{s_j(1-s_j)} + \frac{1}{(1-s_j)^2}}\right) - \frac{2s_j}{(1-\delta_j)^3} - c_q''(\delta_j) < 0$$

This rearranges to:

$$c_q''\left(\delta_j\right)\left(1-\delta_j\right)^4 + 2s_j\left(1-\delta_j\right) > \frac{\alpha}{\frac{1}{s_j\left(1-s_j\right)} + \frac{1}{\left(1-s_j\right)^2}}$$

This is (49).

Characterizing $\frac{ds_j}{d\delta_j}$. First, consider the market share equilibrium conditions, (46) and (47). For analytical convenience, define $\zeta_j \equiv \frac{\delta_j}{1-\delta_j}$, so that (47) becomes:

$$r_j - \zeta_j = \frac{1}{\alpha\left(1-s_j\right)} \tag{54}$$

As δ_j varies, c_j also varies, tracing out (46). From (46), we have:

$$\frac{dr_j}{ds_j} = \frac{-1}{\alpha s_j\left(1-s_j\right)} \tag{55}$$

Now, differentiating (28) totally, we have:

$$dr_j - d\zeta_j = d\left(\frac{1}{\alpha\left(1-s_j\right)}\right) = \frac{1}{\alpha\left(1-s_j\right)^2}ds_j \tag{56}$$

We can solve (56) and (29) for $\frac{d\zeta_j}{ds_j}$, to get:

$$\frac{d\zeta_j}{ds_j} = \frac{-1}{\alpha s_j\left(1-s_j\right)} - \frac{1}{\alpha\left(1-s_j\right)^2}$$

Inverting,

$$\frac{ds_j}{d\zeta_j} = \frac{-\alpha}{\frac{1}{s_j\left(1-s_j\right)} + \frac{1}{\left(1-s_j\right)^2}}$$

Now, from the definition of c_j, we have:

$$\frac{d\zeta_j}{d\delta_j} = \frac{1}{\left(1-\delta_j\right)^2}$$

Hence,

$$\frac{ds_j}{d\delta_j} = \frac{ds_j}{d\zeta_j}\frac{d\zeta_j}{d\delta_j} = \left(\frac{-\alpha}{\left(1-\delta_j\right)^2}\right)\left(\frac{1}{\frac{1}{s_j\left(1-s_j\right)} + \frac{1}{\left(1-s_j\right)^2}}\right) \tag{57}$$

66

B Bank Merger Data

We combine the Bank Mergers data with the quarterly Summary of Deposits data. The bank mergers data is collected from the Transformations file from the National Information Center (NIC) (previously provided by the Federal Reserve Bank of Chicago until 2015), which lists all mergers and acquisitions events of banks and bank holding companies that have occurred since 1976, along with the names of the surviving and non-surviving entities. We exclude government- assisted mergers which resulted in failed banks ceasing to exist.

From the Summary of Deposits data, we initially match commercial banks to their ultimate parent bank holding company using the relationships files from the National Information Center. The latter reports the entire history of a bank's control relationships with others over time, and the effective dates when such relationships hold. We use these links where ownership/control is in any BHC or bank and for which the exact ownership percent is reported. We iterate through the entire chain of relationships for a given entity to infer the ultimate BHC at each point in time, and we adjust the deposit amount recorded in the Summary of Deposits for the ultimate ownership share in the original entity.

When matching the Summary of Deposits data with the bank mergers data, we follow Granja and Paixao (2019): we measure the level of deposits of each acquired bank during the merger year as the level of deposits as of June 30th of the merger year if the merger occurs after June 30th; otherwise, we measure deposits as of June 30th of the previous year.

Between 2009 and 2019 there are a total of 1442 mergers, covering 1812 distinct counties. Table A.4 shows the number of mergers in each year, as well as the distribution of the number of mergers across counties for each year. Table A.5 shows the largest 50 mergers in the sample, by number of counties affected, listed by chronological order.

Figure A.1: Equilibrium Entry
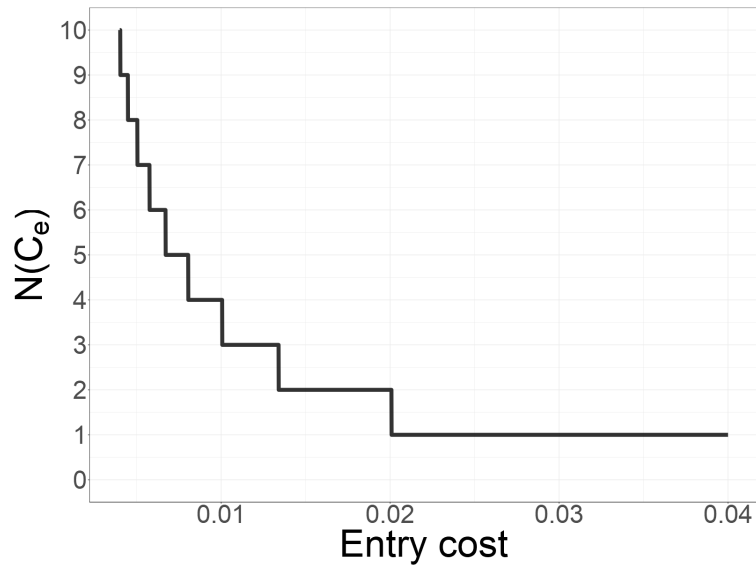
The above figure shows the equilibrium number of entrants as a function of the fixed entry cost C_e, in the entry model of Appendix A.5. The x-axis shows the fixed entry cost C_e. The y-axis shows the equilibrium number of entrants, $N(C_e)$, which we calculate by solving (33) and (34). We use the cost function (14), and we set $\theta = 0.04, \mu = 0.5, k = 10^{-5}$.

Figure A.2: Markups under moral hazard
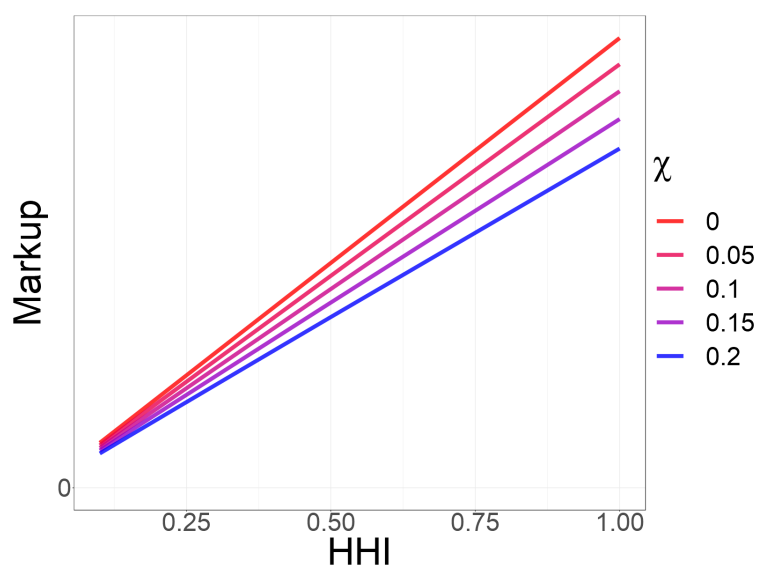
The above figure shows equilibrium markups when there is moral hazard, for different values of χ. The x-axis shows the HHI, as we vary the number of lenders N. The y-axis shows the equilibrium markup, $r_j - \frac{\psi(r_j, \delta_j)}{1 - \psi(r_j, \delta_j)}$, calculated as the RHS of (40). We set $\delta = 0.1, \theta = 0.04$.

Figure A.3: Logit Model Simulations
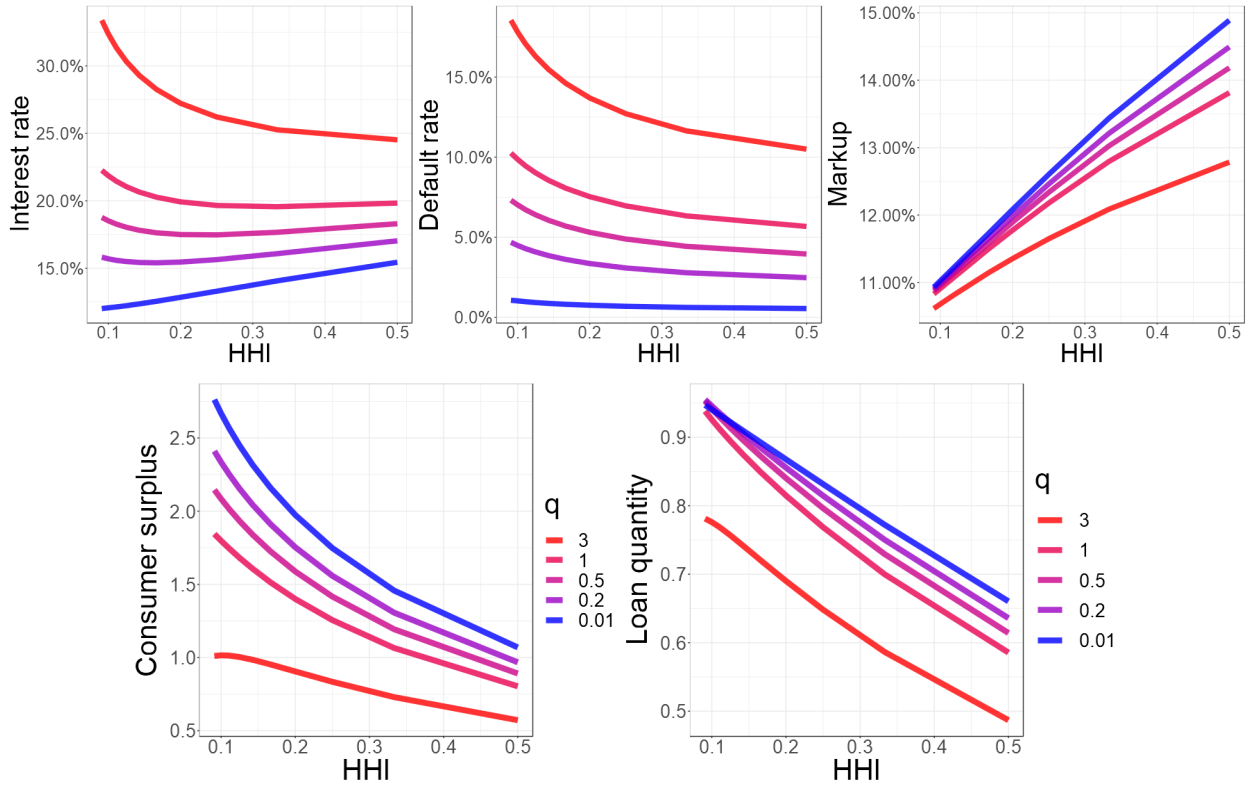
The above figure shows model outcomes, for different values of HHI and q, in the logit model described in Appendix A.8. The top left panel shows equilibrium interest rates, r. The top middle panel shows default rates, δ. The top right panel shows markups over the break-even interest rate, $r - \frac{\delta}{1-\delta}$. The bottom left panel shows consumer surplus, (50). The bottom right panel shows total loan quantity. All simulations use $\alpha = 10, \mu = 1.5, k = 0.001$.

Figure A.4: Auto Loans in TU and Comparable Datasets
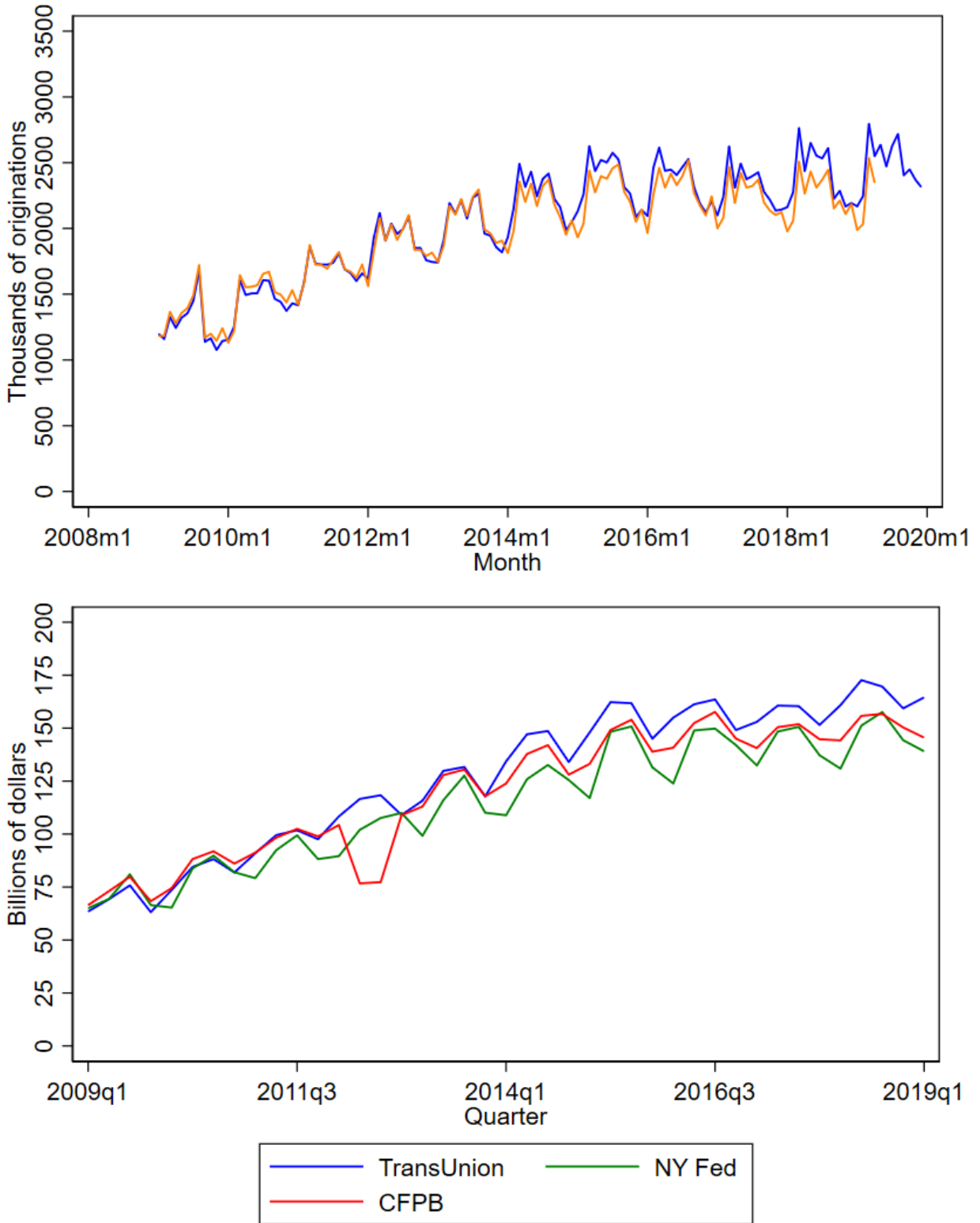
The above figure shows monthly auto loan originations. The top panel shows the number of loans, while the bottom panel shows the volume. The blue line shows the TU data series, while the red line shows a similar CFPB data series. Source: TransUnion, CFPB, and NY Fed.

Figure A.5: Distribution of Credit Scores
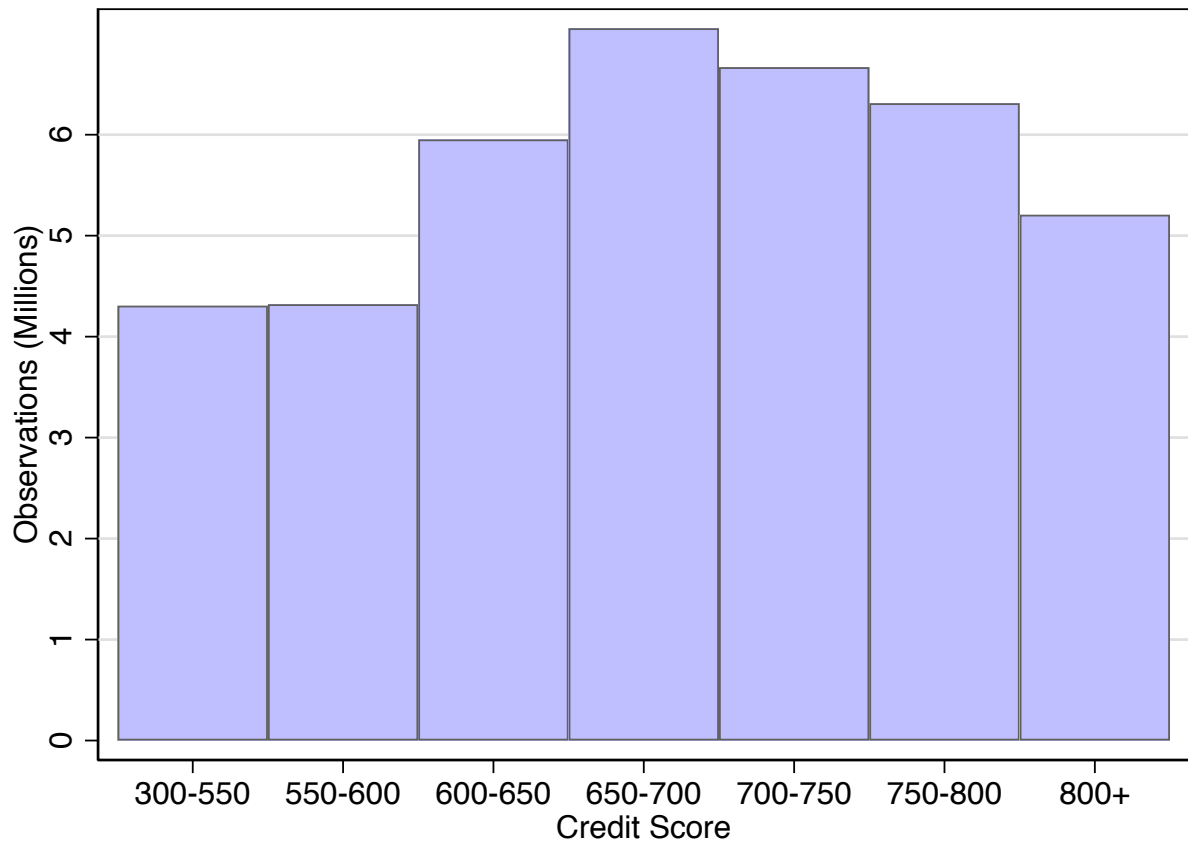
The above figure shows the total number of observations within each credit rating bin. Credit scores are given by VantageScore ratings. Source: TransUnion

Figure A.6: Geographic Distribution of Market Competition
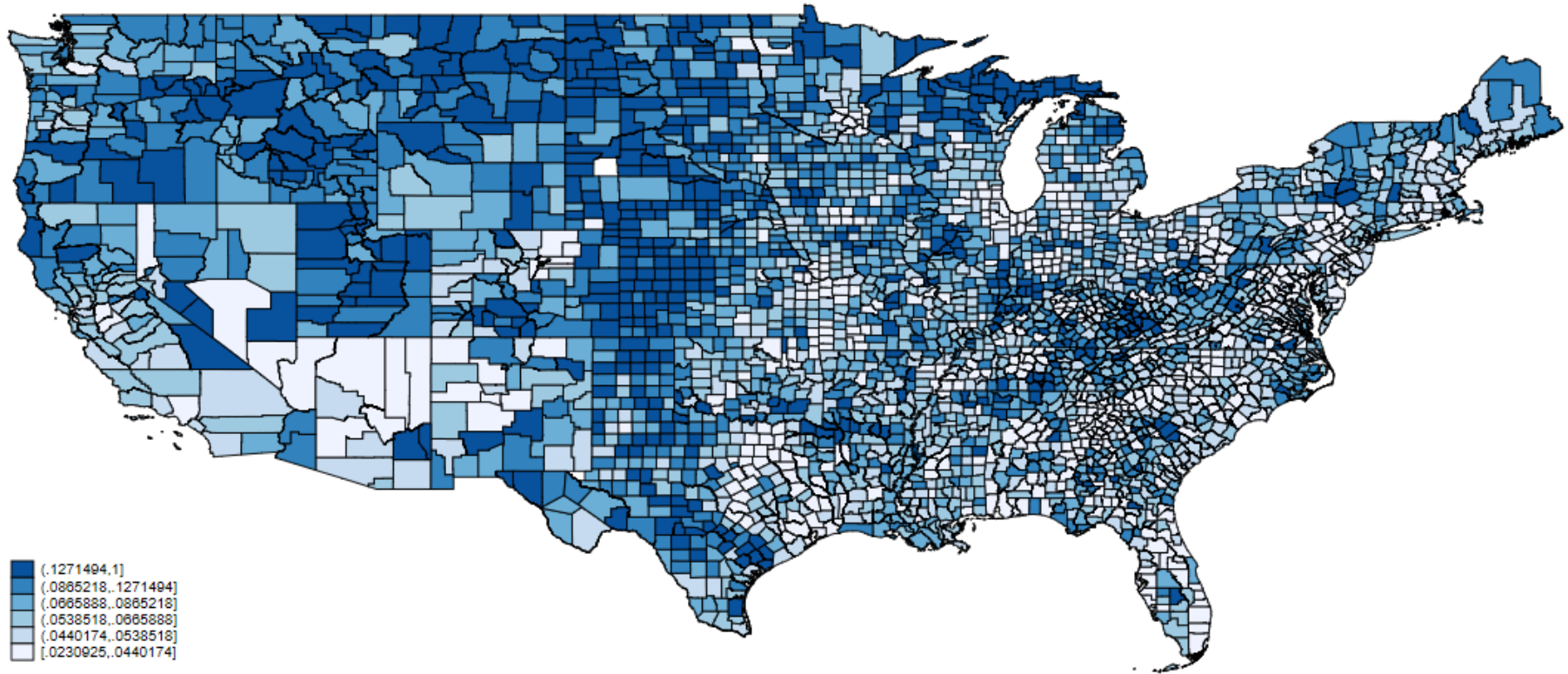
The above figure shows the average auto loan market HHI in each mainland US county in 2009. Darker shades show more concentrated auto lending markets. Source: TransUnion

Figure A.7: Interest Rates and Competition: Alternative Measurement
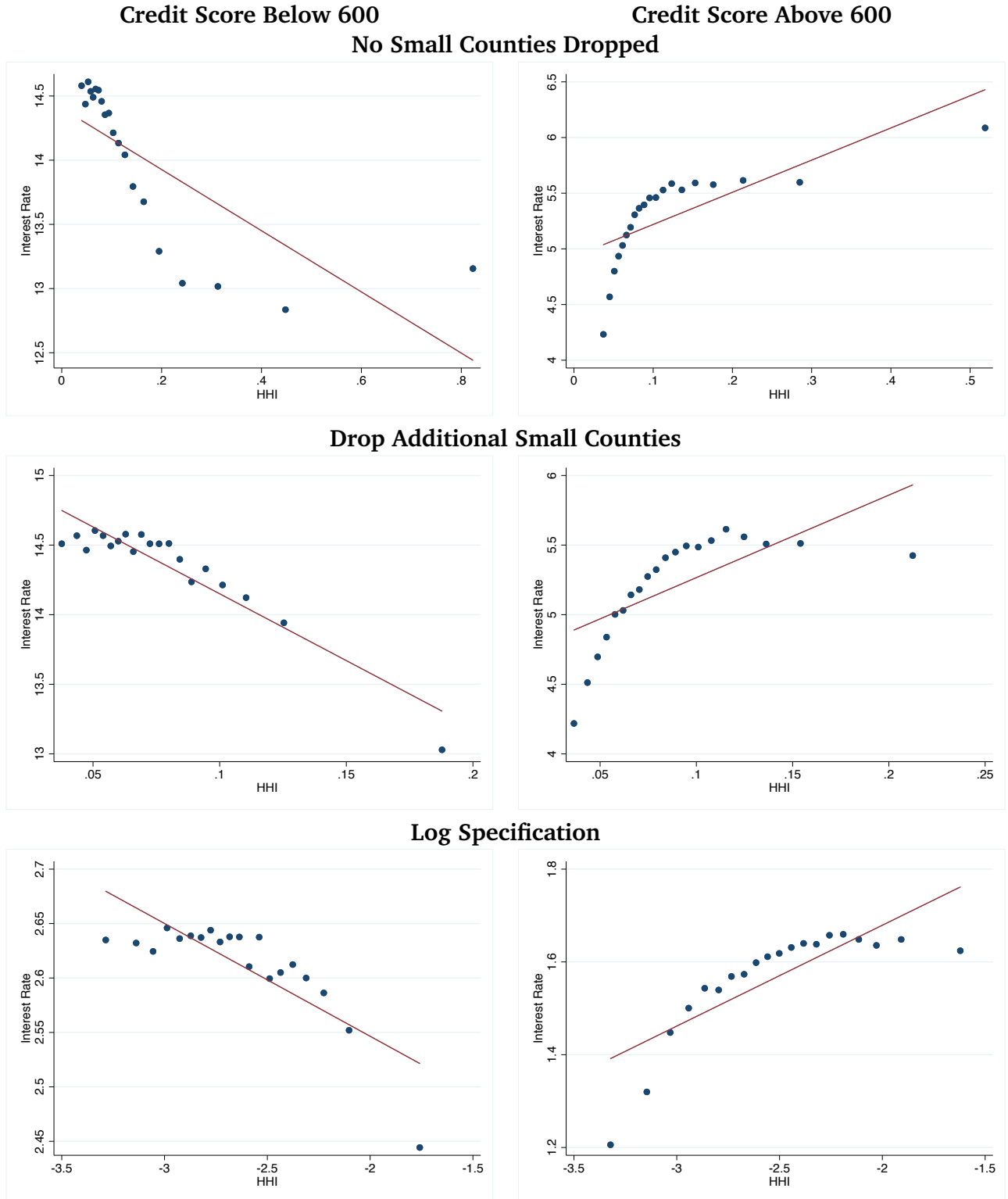
Credit Score Below 600 **Credit Score Above 600**

No Small Counties Dropped

Drop Additional Small Counties

Log Specification

The above figure shows median interest rates in a county for given credit score ranges, broken down by ventile of HHI using alternative samples or measurement. The measurement or sample is noted above each panel. HHI is defined using the volume of auto loans, that is $HHI = \sum_i^N s_i^2$, where s_i is a lender's share of auto lending in a county within a credit score range. Credit score ranges are denoted above each panel. Credit scores are given by VantageScore ratings. Source: TransUnion.

74

Figure A.8: Interest Rates and Competition: Alternative Measurement
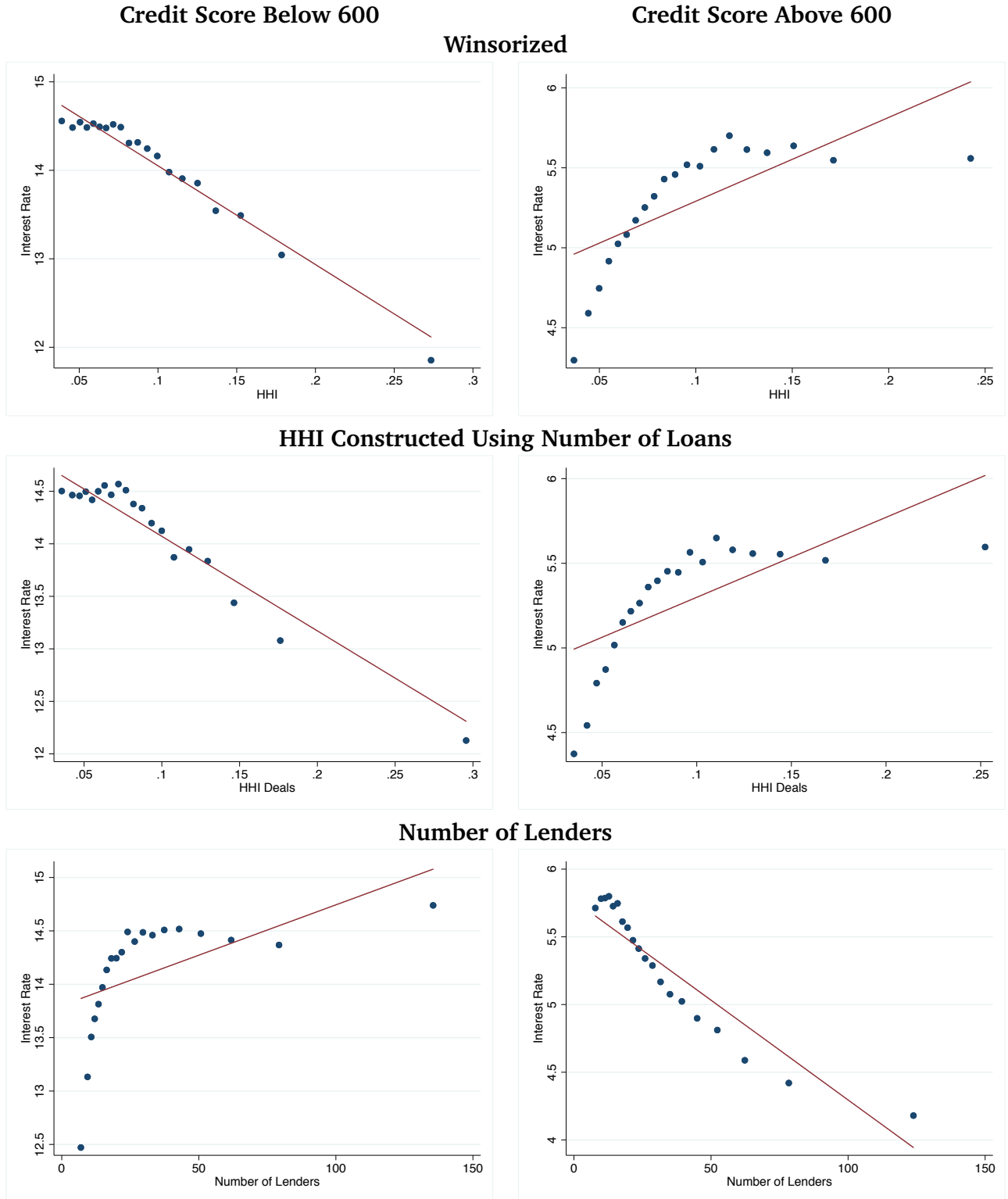
Credit Score Below 600   Credit Score Above 600

Winsorized

HHI Constructed Using Number of Loans

Number of Lenders

The above figure shows median interest rates in a county for given credit score ranges, broken down by ventile of HHI using alternative measurement. The measurement change is noted above each panel. For the top panel, HHI is defined using the volume of auto loans, that is $HHI = \sum_i^N s_i^2$, where s_i is a lender's share of auto lending in a county within a credit score range. For the middle panel, HHI is constructed using lenders' market shares by number of loans, rather than total loan amounts. Credit score ranges are denoted above each panel. Credit scores are given by VantageScore ratings. Source: TransUnion.

Figure A.9: Interest Rates and Competition Excluding New Entrants
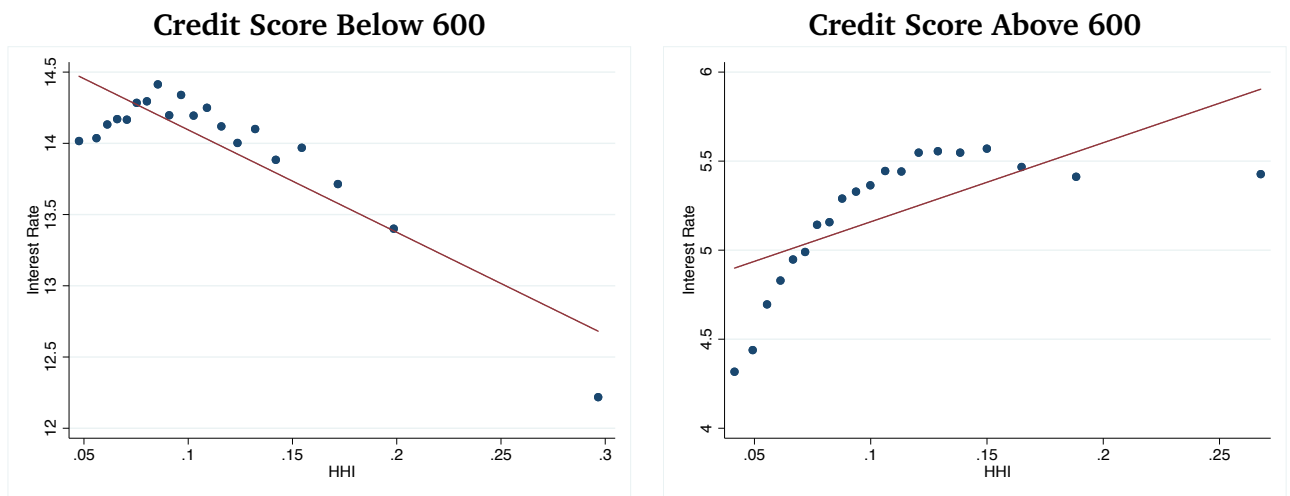
Credit Score Below 600

Credit Score Above 600

The above figure shows median interest rates in a county for given credit score ranges, broken down by ventile of HHI. The sample is restricted to lenders that operated since 2009, and thus excludes new entrants. HHI is defined using the volume of auto loans, that is $HHI = \sum_{i}^{N} s_{i}^{2}$, where s_{i} is a lender's share of auto lending in a county within a credit score range. Credit score ranges are denoted above each panel. Credit scores are given by VantageScore ratings. Source: TransUnion.

Figure A.10: Maturity and Competition
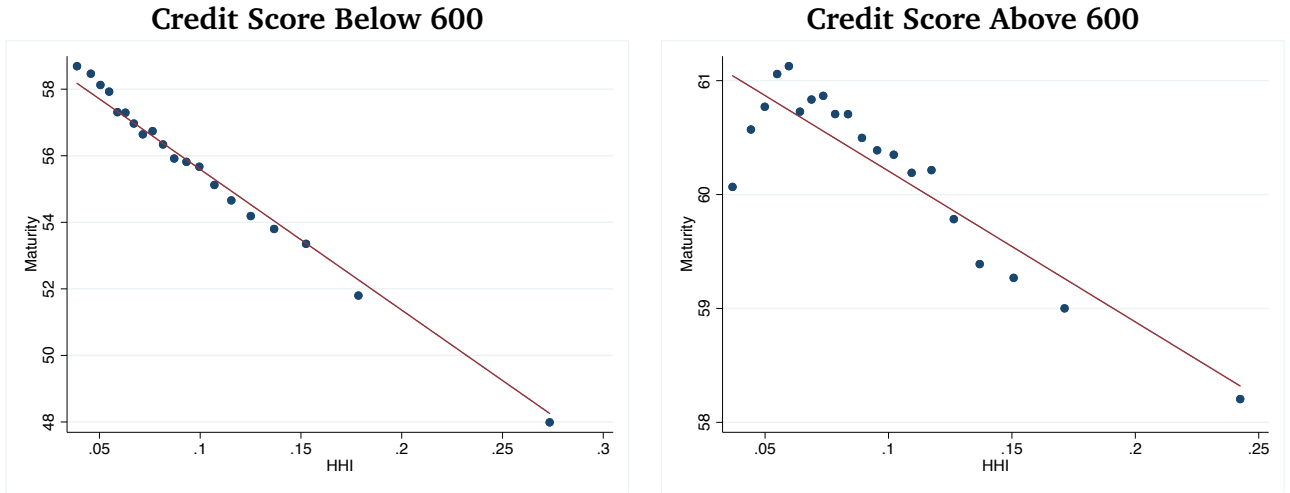
The above figure shows mean loan maturity in a county for given credit score ranges, broken down by ventile of HHI. HHI is defined using the volume of auto loans, that is $HHI = \sum_i^N s_i^2$, where s_i is a lender's share of auto lending in a county within a credit score range. Credit score ranges are denoted above each panel. Credit scores are given by VantageScore ratings. Source: TransUnion.

Figure A.11: Geographic Distribution of Failed Bank Deposit Shares
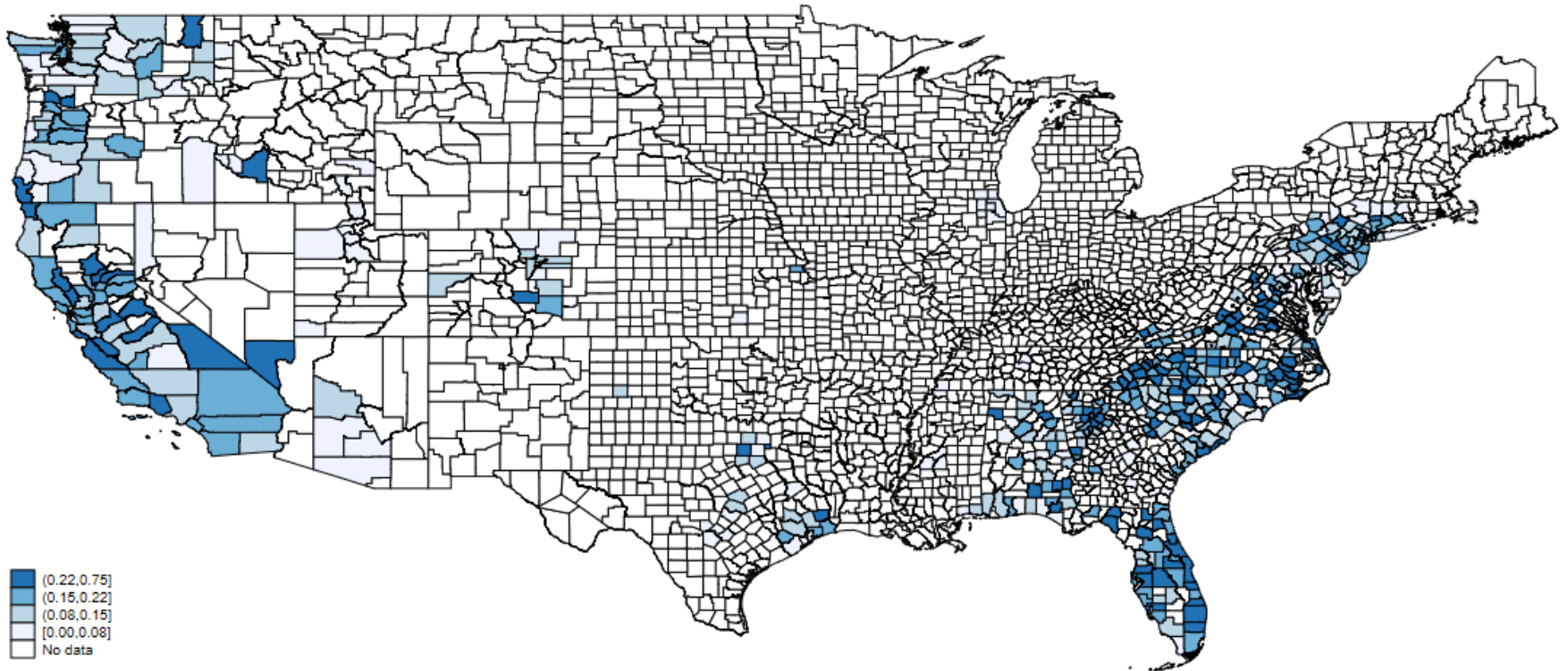
The above figure shows the deposit share in 2008 of the three largest banks that failed during the 2008 crisis, Wachovia, Washington Mutual and Countrywide. Source: Federal Reserve Call Reports

Figure A.12: Geographic Distribution of Mergers

Non-merger County (1301)
Merger County (1803)
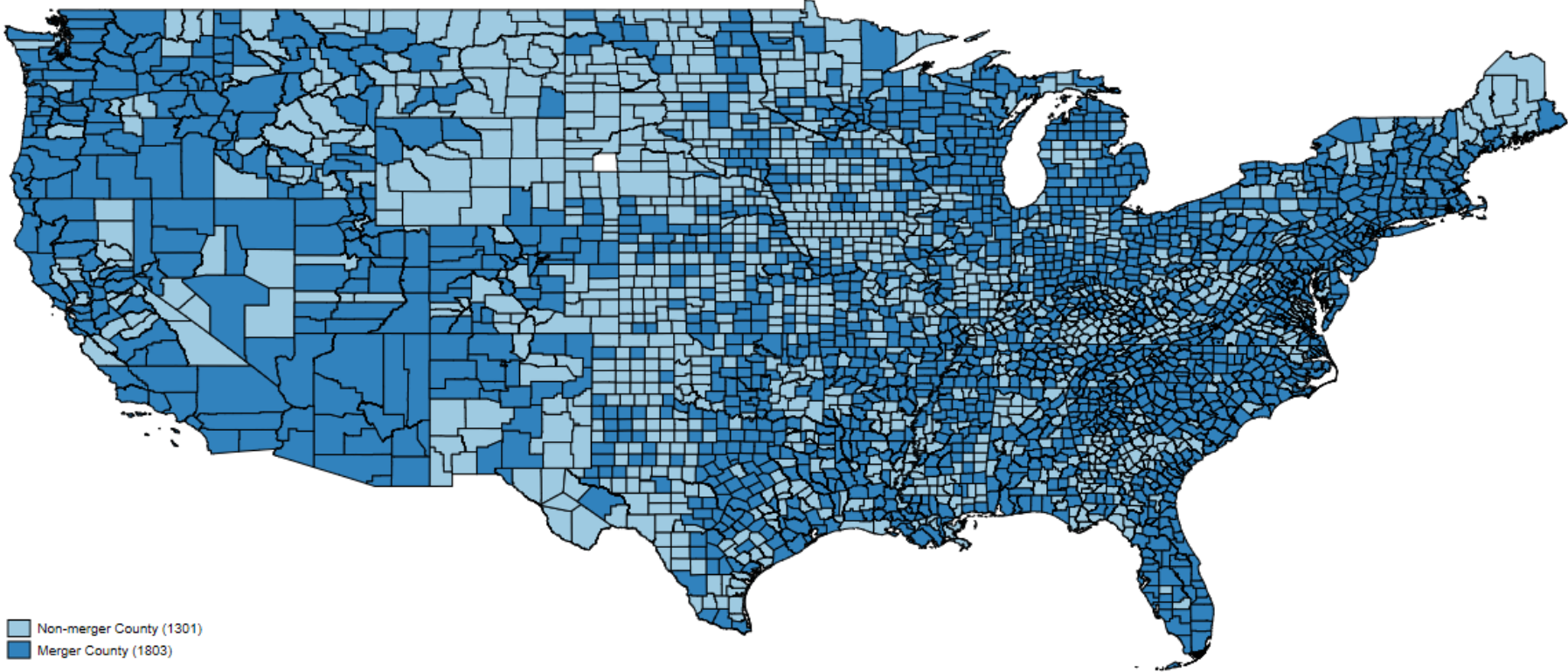
The above figure shows counties affected by bank mergers. Source: NIC and Federal Reserve Call Reports

Figure A.13: Auto Lending and Bank HHI
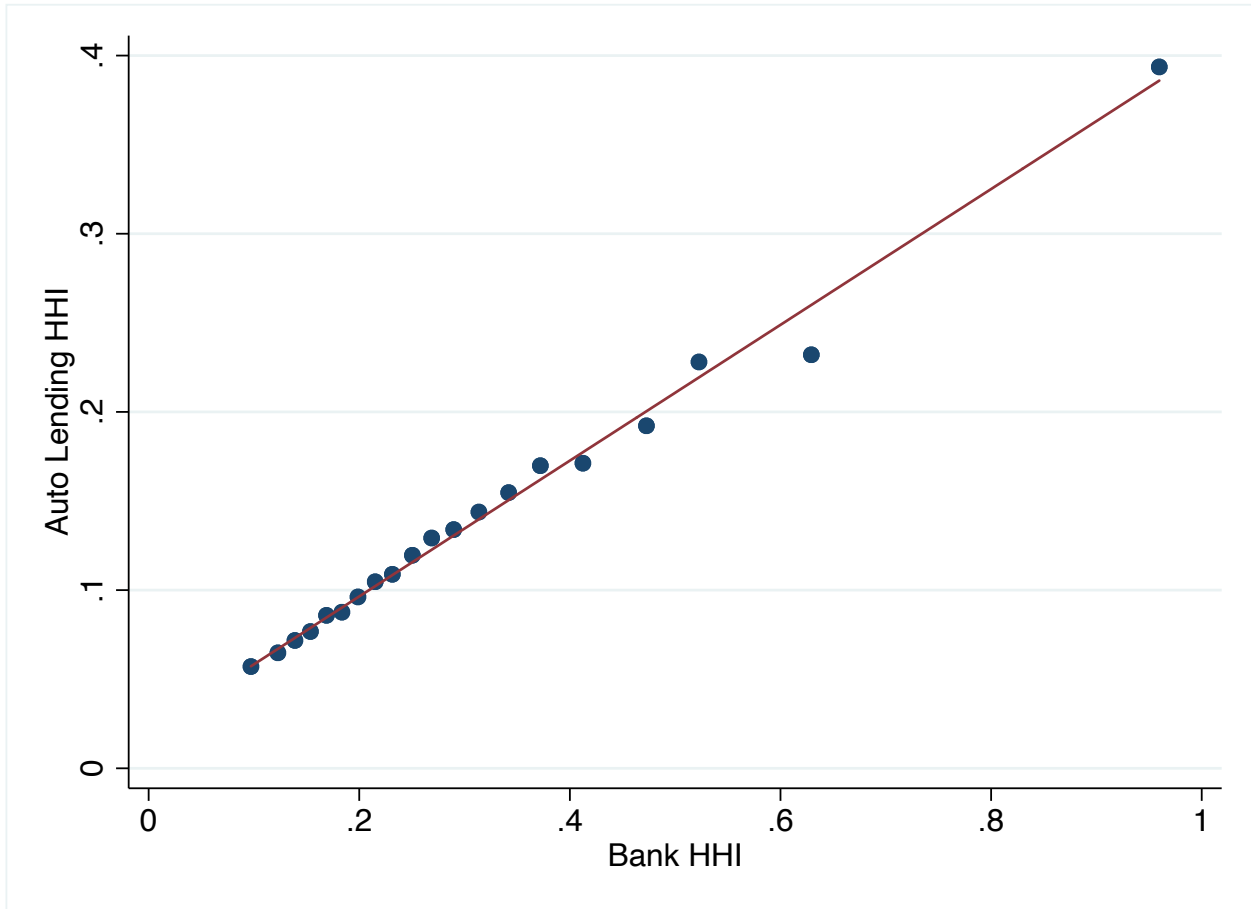
The above figure shows auto lending HHI, broken down by ventile of bank deposit market HHI. HHI is defined using the volume of auto loans or deposits, that is $HHI = \sum_i^N s_i^2$, where s_i is the share of auto lending or deposits in a county within a credit score range. Source: TransUnion and Federal Reserve Call Reports

Figure A.14: HHI Following Acquisition
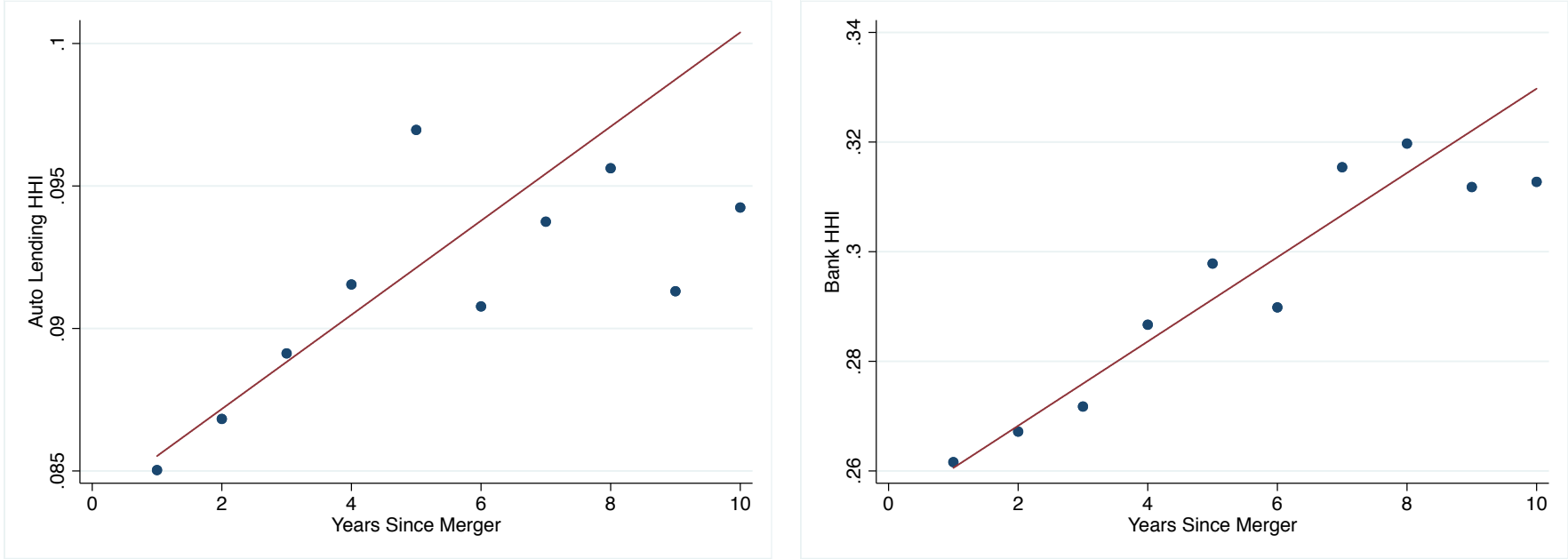
The above figure shows average HHI, by the numbers of years since a bank merger occurred. HHI is defined using the volume of auto loans or deposits, that is $HHI = \sum_i^N s_i^2$, where s_i is the share of auto lending or deposits in a county within a credit score range. Source: TransUnion and Federal Reserve Call Reports

Figure A.15: Interest Rates and Competition in the Mortgage Market
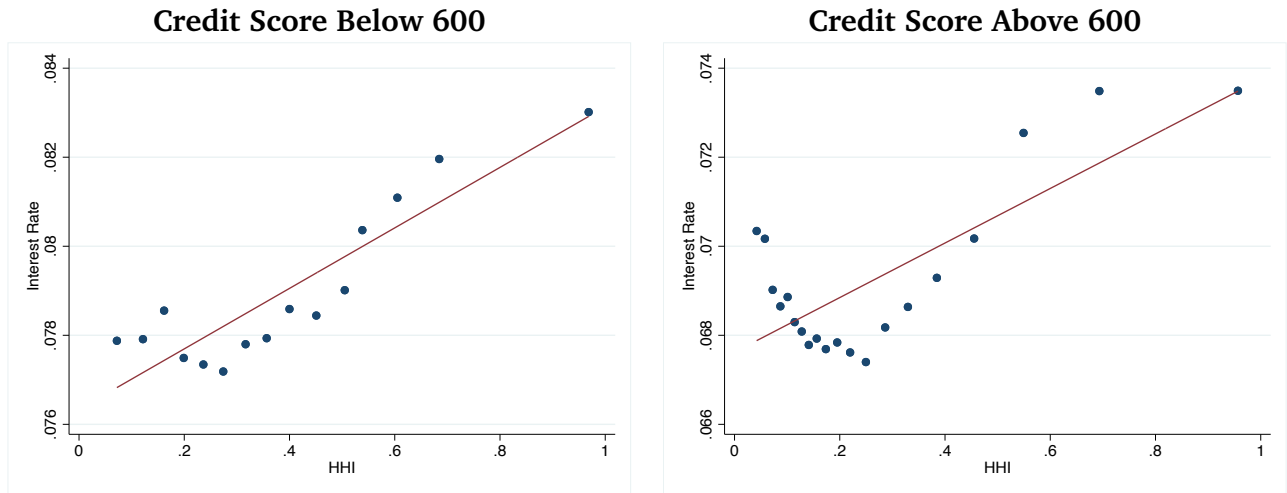
Credit Score Below 600   Credit Score Above 600


The above figure shows median interest rates in a county for given credit score ranges, broken down by ventile of HHI. HHI is defined using the volume of mortgage loans, that is $HHI = \sum_i^N s_i^2$, where s_i is a lender's share of auto lending in a county within a credit score range. Credit score ranges are denoted above each panel. Credit scores are given by VantageScore ratings. Source: TransUnion.

Table A.1: Variable Definitions

This table describes the main analysis variables used. All variables are constructed using a dataset provided to the University of Chicago Booth School of Business by TransUnion. Observations are at the county-year level.

Name	Description
Loan information	
Principal	The original balance on the loan.
Monthly Payment	The payment due on a loan each month.
Maturity	The total number of months that the borrower is scheduled to be making payments on the loan.
Interest Rate	The annual percentage rate on the loan. Calculated by authors using data on principal, payments, and maturity available in the TransUnion data.
90-day Delinquency	Indicates whether the payment on a loan has ever been delinquent by 90 days.
Credit Score	VantageScore 3.0 taken from January of each year for each borrower.
Market information	
HHI (Deals)	The HHI calculated for a given county-year, based on the number of auto loan originations.
HHI (Volume)	The HHI calculated for a given county-year, based on the volume of the auto loan originations.
Number of lenders	The number of lenders that are active in a given county-year.
Share of lenders only in auto market	The percentage of lenders in each county-year that are only active in the auto loan market, and no others.

Table A.2: Concentration and Interest Rates by Lender

The table shows the relationship between the number of loans rates and HHI, split by credit score at the county, lender by year level. The first four columns vary the inclusion of lender fixed effects. The second four columns split the sample by whether a lender offers only auto loans, or other types of consumer credit. The inclusion of fixed effects is denoted beneath each column. Standard errors are clustered at the county level. Credit scores are given by VantageScore ratings. Source: TransUnion. $^*p < .1$, $^{**} p < .05$, $^{***} p < .01$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Ln(Interest Rate)				
	Full Sample				Auto Lenders		All Lenders	
	Credit Score		Credit Score		Credit Score	Credit Score	Credit Score	Credit Score
	300-600		600-850		300-600	600-850	300-600	600-850
Ln(HHI)	-0.0737***	-0.0727***	0.110**	0.105**	-0.0689***	0.173**	-0.0785***	0.0282
	(0.0103)	(0.00996)	(0.0517)	(0.0502)	(0.0114)	(0.0764)	(0.00921)	(0.0245)
County X Credit Score Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lender Fixed Effects	No	Yes	No	Yes	Yes	Yes	Yes	Yes
Observations	1,866,741	1,865,935	5,213,296	5,212,710	1,107,222	2,401,999	758,663	2,810,678
R^2	0.678	0.689	0.802	0.805	0.728	0.813	0.622	0.760

Table A.3: Concentration and Interest Rates by Lender Size

This table shows the relationship between interest rates and HHI, interacted with the total volume of outstanding loans (in billions of dollars) and the number of counties in which a lender is active. The first pair of columns interacts with volume, while the second pair interacts with the number of counties. The first columns in each pair shows estimates for borrowers with VantageScore scores below 600, while the second column in each pair shows estimates for borrowers with VantageScore scores above 600. The inclusion of fixed effects is denoted beneath each column. Standard errors are clustered at the county level. Credit scores are given by VantageScore ratings. Source: TransUnion. $^*p < .1$, $^{**} p < .05$, $^{***} p < .01$.

	(1)	(2)	(3)	(4)
		Credit Score		
	300-600	600-850	300-600	600-850
HHI	-0.0806***	0.0998**	-0.0909***	0.0881*
	(0.0106)	(0.0485)	(0.0108)	(0.0456)
Volume	0.000244***	0.000367**		
	(0.0000254)	(0.000146)		
HHI X Volume	0.0000800***	0.000113**		
	(0.00000833)	(0.0000453)		
# Counties			0.0000327***	0.0000369***
			(0.00000368)	(0.0000117)
HHI X # Counties			0.0000105***	0.0000132***
			(0.00000121)	(0.00000415)
County X Credit Score Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	1,866,741	5,213,296	1,866,741	5,213,296
R^2	0.678	0.802	0.678	0.802

Table A.4: Number of Merger Events by Year

This table shows the annual number of bank mergers and acquisitions. Source: NIC

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Total	55	52	66	109	134	162	189	173	153	184	165

Table A.5: Top 50 List of Bank Mergers

This table lists the top 50 bank mergers ranked by number of counties covered by the acquired bank during the merger year. Source: NIC, Federal Reserve Call Reports.

Event Date	Acquired Bank Name	# Counties
2009-5	PROVIDENT BANKSHARES CORPORATION	23
2010-10	SOUTH FINANCIAL GROUP, THE	58
2010-12	J. R. MONTGOMERY BANCORPORATION	42
2011-7	MARSHALL & ILSLEY CORPORATION	87
2012-7	AMERICAN STATE FINANCIAL CORPORATION	18
2013-2	BANCTRUST FINANCIAL GROUP, INC.	18
2013-4	CITIZENS REPUBLIC BANCORP, INC.	73
2013-4	WEST COAST BANCORP	13
2013-9	FIRST M & F CORPORATION	17
2013-10	LIBERTY BANCSHARES, INC	15
2014-1	STELLARONE CORPORATION	34
2014-4	STERLING FINANCIAL CORPORATION	62
2014-7	PIEDMONT COMMUNITY BANK HOLDINGS, INC.	21
2014-10	FIRST CITIZENS BANCORPORATION, INC.	50
2015-2	COMMUNITY FIRST BANCSHARES, INC.	15
2015-5	CENTRAL BANCSHARES, INC.	13
2015-7	Heritage Bank of the South	22
2015-8	SUSQUEHANNA BANCSHARES, INC.	40
2015-11	Hudson City Savings Bank	23
2015-11	CITY NATIONAL CORPORATION	18
2016-4	NATIONAL PENN BANCSHARES, INC.	18
2016-5	AnchorBank, fsb	18
2016-7	COMMUNITY & SOUTHERN HOLDINGS, INC.	28
2016-8	FIRSTMERIT CORPORATION	97
2016-8	FIRST NIAGARA FINANCIAL GROUP, INC.	59
2016-9	TALMER BANCORP INC.	25
2016-10	FNB UNITED CORP.	17
2017-3	YADKIN VALLEY FINANCIAL CORPORATION	36
2017-4	CARLILE BANCSHARES, INC.	17
2017-5	CASCADE BANCORP	21
2017-5	MERCHANTS BANCSHARES, INC.	13
2017-6	BANK OF THE OZARKS INC	101
2017-11	BANCORPSOUTH, INC.	97
2017-11	FIRST SOUTH BANCORP, INC.	18
2017-12	NORTH AMERICAN FINANCIAL HOLDINGS, INC.	61
2017-12	PARK STERLING CORPORATION	27
2018-1	HCBF HOLDING COMPANY, INC.	15
2018-1	CENTRAL COMMUNITY CORPORATION	13
2018-2	Bank Mutual	29
2018-4	MAINSOURCE FINANCIAL GROUP, INC.	42
2018-9	ZIONS BANCORPORATION	113
2019-1	FCB FINANCIAL HOLDINGS, INC.	17
2019-1	STATE BANK FINANCIAL CORPORATION	15
2019-4	AMERICUS FINANCIAL SERVICES, INC.	17
2019-7	FIDELITY SOUTHERN CORPORATION	23
2019-8	TCF FINANCIAL CORPORATION	38
2019-9	MIDSOUTH BANCORP, INC.	23
2019-9	LIBERTY SHARES, INC.	14
2019-11	LANDRUM COMPANY	14
2019-12	SUNTRUST BANKS, INC.	224

Table A.6: Concentration, Failures and Interest Rates

The first two columns show the relationship between ln(HHI) and the share of deposits in the three banks that failed during the financial crisis. The second two columns show the relationship between the log of interest rates and the share of deposits in the three banks that failed during the financial crisis. The last two columns show the relationship between ln(HHI) and the log of interest rates, instrumenting using the share of deposits in the three banks that failed during the financial crisis. In each pair, the first column shows the relationship for low credit score borrowers (VantageScore below 600) while the second column shows the relationship for high credit score borrower (VantageScore above 600.) The inclusion of fixed effects is denoted beneath each column. Standard errors are clustered at the county level. Credit scores are given by VantageScore ratings. Source: TransUnion. $^*p < .1$, $^{**} p < .05$, $^{***} p < .01$.

	(1)	(2)	(3)	(4)	(5)	(6)
			Ln(Interest Rate)			
	First Stage		Reduced Form		IV	
Share Failed Banks	-0.739***	-0.995***	0.180**	-0.241		
	(0.100)	(0.133)	(0.0826)	(0.253)		
ln(HHI)					-0.243**	0.242
					(0.0982)	(0.250)
Observations	27887	31773	27887	31773	27887	31773
R^2	0.406	0.418	0.014	0.012	.	.

Table A.7: Concentration, Interest Rates and Failures Split by Credit Score Group

The top panel regresses HHI on the deposit share in 2008 of the three largest failed banks, Wachovia, Countrywide and Washington Mutual. The bottom panel presents 2SLS estimates of interest rates of HHI, instrumenting for HHI with the deposit share in 2008 of the three largest failed banks. The sample is split by borrowers' credit scores groups, denoted above each column. The inclusion of fixed effects is denoted beneath each column. Standard errors are clustered at the county level. Credit scores are given by VantageScore ratings. Source: TransUnion. $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$.

Panel A: HHI Following Acquisition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Credit Score			
	300-550	550-600	600-650	650-700	700-750	750-800	800+
				Bank HHI			
Share Failed Banks	-0.650***	-0.836***	-1.031***	-1.104***	-1.137***	-1.104***	-0.902***
	(0.104)	(0.104)	(0.124)	(0.149)	(0.171)	(0.152)	(0.142)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Merger Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	26049	27094	29219	30154	29864	29310	28070
R^2	0.332	0.393	0.400	0.385	0.369	0.368	0.389

Panel B: Interest Rates Following Acquisition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Ln(Interest Rate)			
Years Since Merger	-0.281***	-0.325***	-0.0390	0.103*	0.195***	0.283***	0.349***
	(0.0742)	(0.0904)	(0.0583)	(0.0594)	(0.0582)	(0.0639)	(0.0790)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Merger Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	26049	27094	29219	30154	29864	29310	28070

Table A.8: Concentration, Interest Rates and Bank Acquisitions

The table shows estimates of HHI or log interest rates on the number of years since a bank merger or acquisition occurred, split by borrowers having credit scores below 600. The inclusion of fixed effects is denoted beneath each column. Standard errors are clustered at the county level. Credit scores are given by VantageScore ratings. Source: TransUnion. $^*p < .1$, $^{**}p < .05$, $^{***}p < .01$.

Panel A: HHI Following Acquisition

	(1)	(2)	(3)	(4)
		Bank HHI		
	Credit Score		Credit Score	
	300-600		600-850	
Years Since Merger	0.00186***	0.00186***	0.00173***	0.00173**
	(0.000691)	(0.000691)	(0.000672)	(0.000672)
County Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Merger Year Fixed Effects	No	Yes	No	Yes
Observations	27,481	27,481	31,298	31,298
R^2	0.958	0.958	0.959	0.959

Panel B: Interest Rates Following Acquisition

	(1)	(2)	(3)	(4)
		Ln(Interest Rate)		
	Credit Score		Credit Score	
	300-600		600-850	
Years Since Merger	-0.00835**	-0.00835**	0.00175	0.00175
	(0.00399)	(0.00399)	(0.00195)	(0.00195)
County Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Merger Year Fixed Effects	No	Yes	No	Yes
Observations	27,481	27,481	31,298	31,298
R^2	0.375	0.375	0.827	0.827

Table A.9: Concentration, Interest Rates and Bank Acquisitions Split by Credit Score Group

The table shows estimates of HHI or log interest rates on the number of years since a bank merger or acquisition occurred, split by borrowers' credit scores groups. The inclusion of fixed effects is denoted beneath each column. Standard errors are clustered at the county level. Credit scores are given by VantageScore ratings. Source: TransUnion. $^*p < .1$, $^{**} p < .05$, $^{***} p < .01$.

Panel A: HHI Following Acquisition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Credit Score			
	300-550	550-600	600-650	650-700	700-750	750-800	800+
				Bank HHI			
Years Since Merger	0.00181**	0.00190***	0.00172**	0.00180***	0.00177***	0.00204***	0.00206***
	(0.000706)	(0.000689)	(0.000687)	(0.000676)	(0.000666)	(0.000674)	(0.000666)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Merger Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25683	26689	28792	29706	29424	28877	27661
R^2	0.957	0.958	0.958	0.958	0.957	0.957	0.955

Panel B: Interest Rates Following Acquisition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Ln(Interest Rate)			
Years Since Merger	-0.00774**	-0.00442	-0.00174	-0.000360	0.00265	0.00562**	0.00397*
	(0.00368)	(0.00492)	(0.00261)	(0.00176)	(0.00192)	(0.00229)	(0.00235)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Merger Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25683	26689	28792	29706	29424	28877	27661
R^2	0.309	0.349	0.568	0.701	0.757	0.803	0.801

Table A.10: Interest Rates and Market Competition in the Mortgage Market

The table shows the relationship between interest rates and HHI for mortgage loans, split by credit scores being above 600. The inclusion of fixed effects is denoted beneath each column. Regressions are weighted by the number of loans. Standard errors are clustered at the county level. Credit scores are given by VantageScore ratings. Source: TransUnion. $^*p < .1$, $^{**}p < .05$, $^{***}p < .01$.

	(1)	(2)	(3)	(4)	(5)	(6)
			Ln(Interest Rate)			
		Credit Score 300-600			Credit Score 600-850	
Ln(HHI)	0.0110***	0.0126***	-0.00100	0.0139***	0.0164***	0.00777**
	(0.00425)	(0.00425)	(0.00401)	(0.00324)	(0.00332)	(0.00302)
Year Fixed Effects	No	Yes	Yes	No	Yes	Yes
County Fixed Effects	No	No	Yes	No	No	Yes
Observations	27,497	27,497	27,418	34,099	34,099	34,098
R^2	0.001	0.017	0.311	0.005	0.087	0.631