Competition and Selection in Credit Markets

Constantine Yannelis [†] Anthony Lee Zhang[‡]

March 12, 2023

Abstract

Screening in consumer credit markets is often associated with large fixed costs. We present both theory and evidence that, when lenders use fixed-cost technologies to screen borrowers, increased competition may increase rather than decrease interest rates in subprime consumer credit markets. In more competitive markets, lenders have lower market shares, and thus lower incentives to invest in screening. Thus, when markets are competitive, all lenders face a riskier pool of borrowers, which can lead interest rates to be higher. We provide evidence for the model's predictions in the auto loan market using administrative credit panel data.

We are grateful to Bronson Argyle, Mark Egan, Zhiguo He, Sasha Indarte, Mark Jansen, Jack Liebersohn, Holger Mueller, Taylor Nadauld, Scott Nelson, Amir Sufi and Toni Whited for helpful comments, as well as seminar participants at the NBER Household Finance Summer Institute, Utah Winter Finance Conference, AFA, NFA, ESEM, MFA, Philadelphia Fed, Fourth Biennial Conference on Auto Lending, 2021 NASMES, Barcelona GSE Summer Forum, Northwestern University Kellogg School of Management, UCLA Andersen School of Management, Chicago Corporate and Household Lending Conference and the University of Chicago Booth School of Business. Yannelis and Zhang gratefully acknowledge financial support from the Booth School of Business at the University of Chicago. Livia Amato, Baiyang Han, Peyman Shahidi, Greg Tracey, Julien Weber, Xinkai Wu, and Zichen Zhao provided superb research assistance. TransUnion (the data provider) has the right to review the research before dissemination to ensure it accurately describes TransUnion data, does not disclose confidential information, and does not contain material it deems to be misleading or false regarding TransUnion, TransUnion's partners, affiliates or customer base, or the consumer lending industry.

[†]University of Chicago, Booth School of Business, constantine.yannelis@chicagobooth.edu

[‡]University of Chicago, Booth School of Business, 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 (Mahoney and Weyl, 2017). We propose and test a new channel through which competition can have adverse effects on consumer credit markets.

Screening in consumer credit markets is often characterized by large fixed costs and small variable costs. Lending decisions are usually made through automated processes, based on default rate prediction models.[1] Building these models requires large upfront costs, but the marginal cost of applying these models to additional borrowers are low. Large fixed screening costs can reverse classical intuitions about how competition affects market outcomes, since they imply that there are returns to scale from having large market shares. When markets are concentrated, each lender has a large share of customers, so lenders have high incentives to screen. In competitive markets, lenders invest less in screening; the population of borrowers is thus riskier, and equilibrium interest rates can actually increase.

Our framework makes a simple and surprising empirical prediction: loan rates and concentration should be positively associated in low-risk market segments, and negatively associated in high-risk segments. 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 as of 2020. 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] For example, lenders can invest in better predictive analytics, using machine learning or artificial intelligence, specialized predictive tools or platforms. They can also purchase data, tools to verify information or analytical tools which can be used to screen borrowers. It is also important to note that if screening costs are nonlinear, with decreasing marginal costs – for example, if the marginal cost of purchasing credit reports declines in the number of reports purchased – this is essentially equivalent to having fixed costs in our setting.

We analyze a simple model of imperfect lender competition in consumer credit markets, with fixed screening costs. Competition has two opposing effects on interest rates. The first is a markup effect: in more competitive markets, lenders set smaller markups over their break-even rates. The second is a screening effect: in more competitive markets, each lender's market share is lower, so lenders have lower incentives to invest in screening out borrowers likely to default. This occurs because, in our baseline model, screening is fully fixed-cost: the lender's cost depends on the desired default rate, but not on the number of loans made. Lenders are also differentiated, so they are able to charge prices higher than their marginal costs. In equilibrium, consumers' interest rates are equal to the break-even interest rate – which depends on consumers' default rates – plus a markup which depends on lenders' market power.

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. In low-risk market segments, where baseline default rates are low, the markup effect dominates, and interest rates will tend to be lower in more competitive markets. In high-risk market segments, the screening force can dominate, so equilibrium interest rates can actually increase as markets become more competitive. The model also makes two auxiliary predictions. First, concentrated markets should always have lower default rates, since lenders have higher incentives to screen. Second, 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 their 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-risk 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. A 1% increase in HHI is associated with around a 0.08% to 0.27% increase in interest rates for high-credit score borrowers, compared to a 0.05% to 0.08% *decrease* in interest rates for low-credit score borrowers. These effects are sizable: a 2SD increase in HHI is associated with approximately a 1 percentage point increase in interest rates for low-risk borrowers, compared to a 1 percentage point decrease in interest rates for high-risk borrowers. This counterintuitive 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 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.

Next, we analyze differential predictions of our theory for large and small lenders. Local concentration should also matter more for screening decisions of smaller lenders, since large lenders that are active across multiple counties should tend to make some components of screening decisions at a geographically aggregated level. On the other hand, as long as price-setting decisions are made locally, the relationship of markups with local concentration should be similar for large and small lenders. As a result, local market concentration should be uniformly more negatively associated with interest rates for small lenders, since they increase screening most in response to increases in local market shares. We verify this prediction in the data. However, concentration is negatively associated with subprime interest rates even for large lenders, suggesting that even these lenders make some investments in screening which are location-specific rather than aggregated.

We also find direct evidence that is suggestive that lenders engage in more screening in concentrated markets. An important input into building default rate prediction models is alternative datasets on borrowers beyond their credit scores, offered by credit bureaus and other providers, which contains variables such as borrowers' education, address histories, and other features useful for predicting default. We obtain data on the fraction of lenders which purchase CreditVision, an alternative data product sold by TransUnion. We find that, in the cross-section, lenders in more concentrated markets are more likely to purchase CreditVision.

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) and Crawford et al. (2018), have difficulty explaining our results. In the model of Mahoney and Weyl (2017), market power increases prices regardless of whether selection is adverse or advantageous. In the model of Crawford, Pavanini and Schivardi (2018), competition can decrease interest rates for low-credit-score consumers, but the model has difficulty rationalizing our results on loan quantities and the relationship between competition and lenders' direct investments in screening technologies. Theories of screening through contract characteristics such as down payments, for example Veiga and Weyl (2016), also cannot explain our results. 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. Our results 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 paper fits into a large literature analyzing nonclassical, and potentially adverse, effects of competition in credit markets. We analyze a simple but novel channel, which is relevant in consumer lending markets: in the presence of fixed screening costs, increased competition may decrease screening incentives, and thus increase equilibrium interest rates and default rates.[2] Fixed-cost screening technologies can be thought of as a specific case of endogeneous sunk costs and resultant increasing returns to scale, where marginal costs decrease when firms' output increases. In this sense, our paper joins a broader literature considering competition and consumer welfare in industries with increasing returns to scale. Sutton (1991) provides a general overview of how endogeneous sunk costs affect equilibrium concentration and market structure.

Our work relates to a theoretical and empirical literature demonstrating many distinct channels through which competition can have adverse effects in credit markets. Petersen and Rajan (1995) show that, in relationship banking contexts, competition limits the ability of banks to lend to young and distressed firms, since banks cannot capture the upside if firms' condition

[2]A related paper involving fixed costs is Livshits et al. (2016), which assumes lenders have a fixed costs of contract design, and shows that innovations which reduce these fixed costs can have large extensive margin effects from bringing new borrowers into the market. The difference to our setting is that Livshits et al. analyze fixed costs of contract design rather than screening, and also do not analyze the interaction of fixed costs with market structure and competition.

improves. Other studies, including Broecker (1990) and Hauswald and Marquez (2006), show that competition can exacerbate adverse selection, potentially leading to higher average interest rates.[3] He, Huang and Zhou (2020) study 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. The main difference between our paper and this literature is that we assume information acquisition is a fixed cost, rather than a variable cost, motivated by the prevalence of fixed-cost screening technologies in consumer credit markets; this introduces a new channel through which increased competition can lead to adverse effects.

Our paper highlights data analysis and model-building as the likely foundations of fixed screening costs; we thus relate to a recent literature on the effects of data and data-related technologies on financial markets. This literature analyzes a number of different facets of the influence of data on market outcomes. Some papers look at how changing data technologies influence firms' data collection decisions and equilibrium industry structure. For example, Farboodi and Veldkamp (2020) analyze how technological progress affects investors' decisions to gather data about firm fundamentals, versus other investors' demand. Begenau et al. (2018) and Farboodi et al. (2019) analyze how the rise of big data technologies influences firm dynamics. He, Huang and Zhou (2020) analyze a theoretical model showing how "open data" influences competition between banks. Other papers analyze the welfare and distributional consequences of data for consumers: whether overall welfare increases, and which consumers are better off. Tang (2019) uses data from a fintech to estimate the value of privacy. Jansen et al. (2022) show how to quantify the effects of increased data availability on social welfare and the distribution of surplus in lending markets. Di Maggio et al. (2022) analyze data from a fintech platform, showing that fintechs are able to identify "invisible primes", or borrowers

[3] A number of other papers analyze interactions between competition and adverse selection. Parlour and Rajan (2001) analyze a model of menu competition, showing that outcomes may not converge to perfect competition even as the number of lenders becomes large. 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 et al. (2019) analyze competition and adverse selection in a search-theoretic model, finding that increasing competition can decrease welfare. Crawford et al. (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. We largely disregard adverse selection in the main text to focus on the fixed-cost screening channel; however, we show in Appendix A.8 that our results are robust to accounting for adverse selection.

with low credit scores who are very unlikely to default. We contribute to both these streams of literature, by pointing out that the new screening technologies analyzed by papers such as Di Maggio et al. (2022) involve fixed costs. This leads firms to screen less aggressively when markets are more competitive, potentially making consumers worse off in more competitive markets, which is a new force to consider in analyzing how firm dynamics and imperfect competition are influenced by new data technologies.

More broadly, our paper also relates to a large empirical literature on how competition affects outcomes in financial markets, presenting a new theoretical channel as well as empirical evidence that this channel is relevant. For example, Drechsler, Savov and Schnabl (2017), Egan, Hortaçsu and Matvos (2017), Whited, Wu and Xiao (2021), and Wang, Whited, Wu and Xiao (2022) 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), Agarwal, Song and Yao (2019) and Agarwal et al. (2022) study the effects of competition on lending and discrimination in the mortgage market. Beyhaghi et al. (2020) study the effects of competition in the business lending market. Choi, Kargar, Tian and Wu (2021) study how financial market power affects the misallocation of capital to firms. 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). Bank competition is also known to affect local industry structure (Cetorelli and Strahan, 2006). 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.

In the most narrow sense, this paper also joins a body of work on auto lending. 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. Agarwal et al. (2009) study auto and other lending in a lifecycle framework and Einav, Jenkins and Levin

(2012) present a model of subprime auto lending under imperfect competition with different risk types. Agarwal, Ambrose and Chomsisengphet (2008) provide a discussion of many of the institutional details in this market. We provide new evidence on how competition differentially affects prime and subprime segments of the market.

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 set loan rates for lending to the borrowers. There are N lenders, indexed by j.[4] Lenders simultaneously choose how much to invest in screening out bad-type consumers, and post interest rates at which they are willing to lend to consumers.

Consumers. Each consumer 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.5. In Appendix A.7, 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 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

[4] In Appendix A.6, 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.

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.9, we show that our main results also hold if consumers' preferences over lenders are instead described by a logit model. In Appendix A.11, we show our results also hold in a CES demand model. We also do an extended back-of-envelope quantification of consumers' gains from increased variety in the CES model. In the context of the model, gains from variety are not enough to offset the increase in interest rates, in more competitive markets for the lowest credit score groups. Note, however, that it is difficult to quantify gains from variety precisely, since the quantitative gains from variety are sensitive to the functional form chosen for preferences.

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 this as the cost of hiring data scientists and conducting analysis to build default rate prediction models and automated decision-making software. These systems have large upfront costs to build, but the variable cost of scaling a model to more consumers is low.

In the main text, we assume screening is entirely fixed-cost for expositional simplicity. However, our main results hold when screening has both fixed and variable cost components: as we show in Appendix A.5, any variable cost of lending would simply increase interest rates uniformly. Moreover, if screening has nonlinear variable costs, and the marginal cost is declining in the number of consumers that the screening technology is applied to, the effect is similar to having fixed costs, since the larger the customer base, the cheaper the average cost per consumer of attaining a given average default rate. The basic theoretical force is also robust to the case in which some lenders have fixed-cost screening technologies, and some have marginal-cost technologies. In such a setting, an increase in competition decreases all lenders' market shares; this has no effect on the screening incentives and marginal costs of variable-cost screening technologies, but decreases screening incentives, and increases default rates and marginal costs, for lenders with fixed-cost screening technologies.

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 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 $\tilde{c}\left(\alpha_j\right)$. We assume $\tilde{c}\left(\alpha_j\right)$ is strictly decreasing in α_j: stronger signals (smaller α_j) are more expensive. Aggregating across borrowers, a measure $q\left(1-\alpha_j\right)$ 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 $\tilde{c}\left(\alpha_j\right)$ 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 $q(1-\alpha)$ of bad-type 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

[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.

consumers who are not detected are treated identically by all firms.[6]

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

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

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

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

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:

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

Since $\tilde{c}$ is decreasing, the function $c_q\left(\delta_j\right)$ is increasing in q, fixing δ_j. The intuition behind the $c_q\left(\delta_j\right)$ function is due to the fact that, when the fraction of type-B consumers in the population is higher, lenders need to invest more in screening to achieve a given default rate. Suppose the measure of bad types in the population is, say, $q = 0.03$; then the lender can attain a default rate of $\delta_j \approx 0.03$ without investing anything into screening. In contrast, if the measure of bad types in the population is, say, $q = 0.10$, the lender has to acquire a fairly strong signal to remove enough bad types from the lending pool to decrease default rates to $\delta_j = 0.03$. Thus, when there are more bad types in the population, it is more costly for the lender to attain any fixed default rate, because the lender must invest in acquiring a strong signal, which identifies and removes more bad types from the population, in order to attain any fixed default rate.

[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.8, and show it does not affect our main results.

2.1 Equilibrium

We analyze symmetric equilibria of the model, in which lenders make optimal screening and price-setting decisions. First, note that in any symmetric equilibrium, lenders set identical interest rates, and each lender has market share:

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

If the default rate is δ_j, the total mass of consumers who receive good signals in the market is:

$$\frac{1}{1-\delta_j}$$

This is the measure 1 of good types, plus the measure $\frac{\delta_j}{1-\delta_j}$ of bad types who receive good signals. Lenders' total profits can thus be written as:

$$\Pi = \underbrace{\left(\frac{1}{1-\delta_j}\right) s_j \left(r_j\left(1-\delta_j\right)-\delta_j\right)}_{Lending\ profits} - \underbrace{c_q\left(\delta_j\right)}_{Screening\ costs} \tag{4}$$

Lenders choose δ_j and r_j to maximize (4). In words, j lends to a share s_j of the measure $\frac{1}{1-\delta_j}$ of consumers who pass the screening technology. Among these consumers, with probability $\left(1-\delta_j\right)$, 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 also pay the fixed screening cost of $c_q\left(\delta_j\right)$.

Price setting. Interest rates r_j do not enter into screening costs directly, so $c_q\left(\delta_j\right)$ does not directly affect lenders' price-setting incentives. In the main text, we assume lenders can borrow at zero interest rates; we relax this in Appendix A.5. Rearranging lending profits in (4) somewhat, lender j chooses r_j to maximize:

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

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. In Appendix A.2 and A.3, we characterize first- and second-order conditions for lenders' choice of δ_j. The first-order condition is:

$$\frac{s_j}{\left(1-\delta_j\right)^2} = -c_q'\left(\delta_j\right) \tag{7}$$

Intuitively, the left-hand side of (7) 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. 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 (7), the following proposition states conditions on r, s, δ which characterize symmetric equilibria.

Proposition 1. *Necessary conditions for a symmetric equilibrium are that all lenders' market shares s, default rates δ, and interest rates r are as follows. Market shares of each lender are:*

$$s = \frac{1}{N} \tag{8}$$

Lenders must set prices optimally:

$$r - \frac{\delta}{1-\delta} = \frac{\theta}{N} \tag{9}$$

All lenders must make optimal screening decisions:

$$\frac{1}{N(1-\delta)^2} = -c_q'(\delta) \tag{10}$$

$$c_q''(\delta) > \frac{1}{2\theta(1-\delta)^4} - \frac{2}{N(1-\delta)^3} \tag{11}$$

Total loan quantity is:

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

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

$$\mu - r - \frac{\theta}{4N} \tag{12}$$

Note that expression (12) 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.[7]

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

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

Plugging into (3), this implies that:

$$c_q\left(\delta_j\right) = \frac{kq\left(1-\delta_j\right)}{\delta_j} = \frac{kq}{\delta_j} - kq \tag{13}$$

Expression (13) 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

[7] 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.

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 these 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.[8]

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

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

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.[9] 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.[10]

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 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 incur costs if borrowers default. In 2020, only 14% of auto loans were securitized according to SPG Global.[11] A slightly higher fraction of subprime loans are securitized, but the vast majority – three quarters – of subprime auto loans are not securitized.

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

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

[11] 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.

In consumer lending settings, lenders' screening and monitoring costs have large components that are either fully fixed, or nonlinear, with decreasing marginal costs in the number of consumers they are applied to. Lending decisions are generally made using standardized processes applied to consumers, either automated decision software, or rate charts used manually by loan officers. Lenders improve these automated processes by investing in predictive analytics or using machine learning and artificial intelligence. Other methods involve modeling prepayment risk can also allow lenders to go beyond traditional credit scoring, or investing in software to verify a car's condition. Hiring analytics personnel and developing default prediction models is expensive, but the marginal cost of applying a model to one additional borrower is low. Thus, when a lender has a larger market share, her fixed costs of screening are amortized over a larger customer base, so she has a higher incentive to invest in improving her default rate prediction models.

A natural question is at what level screening decisions are made, and costs incurred: whether lenders develop a model for the entire US, or develop models for specific local regions. At least some component of screening costs is region-specific. 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 default risks than others; detecting these relationships and using them to price loans may require location-specific data acquisition and analysis.[12] This distinction is important for us because we use county-year HHI as our primary dependent variable; county-year concentration is only valid for measuring screening incentives if a large component of screening costs is region-specific.[13]

[12]Address history and property values appear to be important components of some alternative credit data products. See, for example, LexisNexis RiskView. These data investments must often be purchased in bulk.

[13]We provide additional evidence on this point in subsection 4.2, where we show that rates for large lenders, who are active across multiple counties, are more positively associated with local HHI. This suggests that these lenders' screening decisions are made at a more geographically aggregated level; local market concentration increasing gives them more market power, but does not lead them to screen more, so tends to increase interest rates more for these lenders.

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 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.

Table 1 shows summary statistics of our data. 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 auto loans from different datasets in Appendix Figure A.5. Appendix Figure A.6 shows the distribution of consumers by credit score in our data. The average interest rate in our sample is 7.8%, 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),

[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.

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

$$HHI_{ct} = \sum_{l=1}^{N} s_{clt}^2 \tag{14}$$

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.7 shows the geographic distribution of HHI.[17]

Transunion (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 auto loan is 3.65% and the average rate for a used auto 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 .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.

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 percentage point decrease in interest rates for high-risk borrowers, and a 1 percentage point increase in interest rates for low-risk borrowers.[18]

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.

[18] While the model predicts a linear relationship between interest rates and competition, we see a non-linear effect that is declining in concentration for high credit score borrowers. This is likely due to bias from smaller rural counties having fewer lenders and lower interest rates. The figure exploits cross-sectional variation, and does not include geography fixed effects. Appendix Figure A.8 repeats the analysis, weighting bins by the number of loans originated. The non-linearity is less pronounced, which is consistent with the aforementioned explanation.

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

$$ln(r_{ct}) = \alpha_c + \alpha_t + \beta ln(HHI_{ct}) + \varepsilon_{ct} \tag{15}$$

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 (15) separately by different credit score buckets. The sample is collapsed at the county level, taking the average interest rate and HHI in a county for each credit score group.

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 next three columns of Table 2 present an alternative and more granular specification, interacting HHI_{ct} with credit score bins. In these columns, the sample is again collapsed at the county level but we have a more granular sample for which we take the average interest rate and HHI in a county for each credit score group. The columns show estimates from the specification:

$$ln(r_{cst}) = \alpha_c + \alpha_t + \alpha_s + \sum_{s \in S} \beta_s ln(HHI_{ct}) \times \mathbb{1}[s \in S] + \zeta_{cst} \tag{16}$$

where $s \in S$ are six finer credit score buckets of approximately equal size: 300-550, 550-600, 600-650, 650-700, 700-750 and 750+. The coefficient β_s captures the effect of market concentration on interest rates for a given credit score group s. Here, the data are again collapsed at the county by year by credit score bucket level, albeit at a finer level of granularity.

The first six columns 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 column 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 .05 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 .08 to .27 percent *increase* in interest rates. The relationship is slightly weaker for high-credit score borrowers, losing significance in the specification with saturated fixed effects. The standard deviation of log HHI in our data is approximately 0.49, and the average interest rates are approximately 14% for low credit score borrowers and 5% for high credit score borrowers, implying that a 2SD increase in log HHI is associated with a roughly 0.69 to 1.1 percentage point decrease in interest rates for low credit score individuals, and a 0.34 to 1.32 percentage point increase in interest rates for high credit score individuals.

The next three columns of Table 2 splits the sample into finer credit score bins and presents estimates of (16), gradually adding in county and year fixed effects. In most specifications, 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. Figure 4 presents the same information graphically, showing estimates of column (8) of Table 2 along with a 95% confidence interval. The pattern is monotonic, with the relationship between HHI and rates moving from negative to positive as credit scores increase past 600.

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

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

$\mathbb{1}[CreditScore_{Low}]$ 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.10 and A.11 repeat the binscatters of Figure 2 with these different specifications; the stylized facts from the baseline specification hold in all cases.

4.2 Lender Size

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 for these lenders.

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 8.1 percentage points lower for subprime borrowers for a lender operating in a single county, but only 4.9 percentage points lower for a lender operating in all counties in the United States. Notice, however, that the relationship between local HHIs and interest rates is negative for subprime borrowers even for large lenders; this suggests that, even for large lenders, local factors may play some role in screening.[19]

In principle, we could test our results by regressing interest rates on lender size directly, rather than on local market concentration. We do not pursue this approach because lender size is possibly related to many other factors which may affect interest rates, such as funding costs and general operational costs, so it is hard to isolate effects of competition and screening in particular by looking at lender size.

4.3 Direct Evidence of Screening

Next, in Figure 5, we provide suggestive evidence that lenders engage in more screening in more concentrated markets. The figure shows that in the cross-section, lenders in more concentrated markets purchase more information on borrowers. The figure shows 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

[19]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.

regression using the number of loans based measure of HHI yields a coefficient of 0.881 with a standard error of 0.335.

A natural question is how the average credit score of borrowers is associated with market concentration. We show this relationship in Appendix Figure A.9. 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. Both panels show that credit scores of borrowers are higher in more concentrated markets: lenders in more concentrated markets tend to restrict credit somewhat to observably high-risk borrowers, lending comparatively more to observably low-risk borrowers. Our theory does not make a sharp prediction on how concentration should be associated with observable features of borrower creditworthiness. These results may be driven by lenders being partially capacity-constrained, and preferentially serving low-risk borrowers, who may be easier to lend to, when markets are concentrated and their capacity constraints are more binding.[20]

4.4 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. The first three columns of 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. We find a significant negative relationship between delinquency rates and HHI for borrowers in all credit score groups except for the highest, where the coefficient is insignificant.

We have thus shown that the qualitative prediction of our theory, that delinquency rates are lower in more concentrated markets, holds in the data. In Appendix B.1, we compare the relative magnitudes of the effects of concentration on interest rates and delinquency rates. We show quantitatively that the decrease in delinquency rates we find for subprime borrowers is

[20]These results are also consistent with Buchak and Jørring (2021), who find that increases in market concentration are associated with a restriction of credit towards low credit score, high-risk borrowers.

large enough to explain the decrease in subprime interest rates.

4.5 Loan Quantities

Next, we test Prediction 3, regarding the relationship between concentration, interest rates, and loan quantities in low credit-score buckets. In Table 4, we again present estimates of equation (16), in which we regress the log of loan quantity on the log of HHI. 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.

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 quantities. 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.

We highlight that our loan quantities results can also be interpreted as that, in concentrated markets, credit availability to subprime borrowers is limited. In our model, screening and credit restriction always increases social welfare, because all borrowers screened out have willingness-to-pay lower than the cost of providing credit to them. However, in reality, some borrowers who are screened out in concentrated markets may in fact have willingness-to-pay higher than their costs. Thus, there may be distributional implications of increases in market concentration: borrowers who pass lenders' screens get loans at lower rates, but some higher-risk borrowers may also be excluded from credit markets.

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:

$$ln(r_{clst}) = \alpha_c + \alpha_t + \alpha_s + \alpha_l + \sum_{s \in S} \beta_s ln(HHI_{ct}) \times \mathbb{1}[s \in S] + \zeta_{cst} \tag{18}$$

where r_{clst} is the average interest rate charged by lender l in county c in credit score bucket s at time t, and HHI_{ct} is the HHI for county c, time t. In addition to county, time, and credit score bucket fixed effects, specification (18) 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 qualitative results continue to hold. 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 5-10, we divide lenders into two subsamples: lenders who only make auto loans (columns 5-7) and lenders who make multiple kinds of loans (columns 8-10). In both cases, we include county, year, and lender fixed effects. For both subsamples, the results are quantitatively similar, though in some specifications the coefficients are not significant. Appendix Figure A.12 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%,[21] 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.[22]

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.14 and Table A.4 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.[23]

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. These arguments are summarized in Table 5.

[21] See S&P Global.

[22] 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 auto loan 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.

[23] In online appendix Table A.5, we also show that non-monotonic effects persist, but diminish over time. This is consist with increased use of data and technology in screening borrowers, and increased lending by Fintechs leading to less regional integration in lending markets.

Adverse selection. In our model, the primary channel driving our results is costly information 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 endogeneously. 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] Moreover, pure adverse selection in the framework of Mahoney and Weyl has ambiguous implications for how default rates are associated with competition, and cannot explain why prices and quantities can simultaneously decrease: if lenders only use prices to screen, lower prices should always associate with higher quantities of loans demanded.

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. The theory in Crawford et al. (2018) also cannot explain why prices and quantities simultaneously decrease in concentrated markets for low-credit score consumers.

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 (Favara and Imbs (2015), Mahoney and Weyl (2017)). Mian and Sufi (2009) construct a related model in which changes in mortgage interest rates are driven by variation in the risk premium charged

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

on mortgages. Our channel is related to these models, but distinct. As we discussed above, the model of Mahoney and Weyl (2017) predicts that increased competition should always be associated with *lower* interest rates, counter to what we observe in the data. 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).

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 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 6 and Table 4. Moreover, it is unclear how screening through down payments could explain why quantities are higher, 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. 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.7 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.

Dealer markups. Another hypothesis is that vertical competition is driving the effects we observe. 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. In its most basic form, 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.

Moreover, the dealer markups story does not seem to explain the patterns we find in default rates. If loan interest rates are higher in more competitive markets for subprime consumers because dealer markups are higher, under some assumptions, we should observe *lower* default rates in more competitive markets. For a given interest rate received by consumers, when dealers charge higher loan markups, the interest rate received by lenders is lower. Lenders would thus have to be more conservative in their lending decisions, lending only to consumers who are good credit risks. This should lead to lower default rates among consumers in more competitive subprime markets, if interest rate increases in more competitive markets are driven by increased dealer markups. Empirically, we observe that default rates are always higher in more competitive markets, providing some evidence against the hypothesis that outcomes are driven by dealer markups.

Appendix Table A.2 also 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 integrated lenders, who represent both the dealer and the lender in any given transaction; they would thus not have an incentive to charge themselves markups on loans made to customers, so the dealer markup channel should not apply to these lenders. Our stylized facts hold within the group of auto-loan-only lenders, providing suggestive evidence supporting the fact that dealer markups do not explain away our effects. However, this evidence is imperfect, since there may be auto-loan-only lenders whose lending is also mediated by dealers.

The dealer markups channel could play a role if the effects of bundling are heterogenous for low and high credit score borrowers. One possibility is that bundling between dealers and lenders is particularly relevant for low-sophistication consumers, who tend to have lower credit scores. Some variants of this story could also generate predictions on quantities, for example, if due to higher demand dealers could charge more to lenders, lowering bank profits perhaps by as much as higher interest rates increase profits. To address this concern, we show that our results are robust to at least noisily controlling for dealers' loan markups using aggregated data from Jansen et al. (2021). In online appendix Table A.6 we repeat our main specification controlling for average dealer markups on loans at the state by year level, using data from a large national automobile dealer; our main results continue to hold.

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. Collections technologies seem less likely to fit our framework because most such technologies, such as GPS tracking and remote-start, are more similar to marginal costs than fixed costs. With pure marginal cost technologies, lenders do not have higher incentives to invest when they have larger market shares, so interest rates should always decrease when competition increases, and the predictions for default rates are ambiguous.

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. The presence of fixed screening costs has interesting implications for equilibrium market structure in consumer lending markets.[25] One possible explanation is that regulatory constraints on lending markets cause concentration to be lower than what would persist in unconstrained free-entry equilibrium. Banks and credit unions, who are major players in the auto lending market, are unusually unconcentrated, due to factors such as historical branching regulation. Moreover, banks and credit unions may receive implicit subsidies to lending rates from access to low-cost sources of funds through insured deposits, limiting the possibility for pure auto lenders to fully displace their role in these markets. Thus, one possibility is that auto lending (and possibly other consumer lending markets) would be much more concentrated in free-entry equilibrium, but that regulation in financial markets maintain a fairly low level of concentration in these industries. Our results also suggest, that there is some degree of inefficiency in the industrial organization of these markets. Lenders appear to be developing screening technologies largely independently, though 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. The theory may also apply to other consumer lending settings, where screening technologies appear to have large fixed cost components. Some examples include mortgage markets, credit card markets, and markets for unsecured personal loans. 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.

[25] In a classic study of the determinants of equilibrium market structure, Sutton (1991) argues that the presence of endogeneous sunk costs – fixed-cost investments firms can choose to make, which decrease marginal costs or increase demand – can lead to high levels of industry concentration even as market size becomes large. Fixed screening costs are an example of endogeneous sunk costs in Sutton's framework. An interesting question for future research is why equilibrium market structure in the auto lending industry remains fairly unconcentrated despite the presence of these endogeneous sunk costs.

References

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

Agarwal, Sumit, Brent W Ambrose, and Souphala Chomsisengphet, "Determinants of Automobile Loan Default and Prepayment," *Economic Perspectives*, 2008, *32* (3).

—, **Changcheng Song, and Vincent Yao**, "Banking Competition and Shrouded Attributes: Evidence from the US Mortgage Market," *Available at SSRN 2900287*, 2019.

—, **Gene Amromin, Souphala Chomsisengphet, Tim Landvoigt, Tomasz Piskorski, Amit Seru, and Vincent Yao**, "Mortgage Refinancing, Consumer Spending, and Competition: Evidence from the Home Affordable Refinance Program," *Review of Economic Studies*, 2022.

—, **John C Driscoll, Xavier Gabaix, and David Laibson**, "The Age of Reason: Financial Decisions Over the Life Cycle and Implications for Regulation," *Brookings Papers on Economic Activity*, 2009, *2009* (2), 51–117.

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.

Begenau, Juliane, Maryam Farboodi, and Laura Veldkamp, "Big Data in Finance and the Growth of Large Firms," *Journal of Monetary Economics*, 2018, *97*, 71–87.

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.

Beyhaghi, Mehdi, Cesare Fracassi, and Gregory Weitzner, "Bank Loan Markups and Adverse Selection," *Available at SSRN 3733932*, 2020.

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.

Choi, Jaewon, Mahyar Kargar, Xu Tian, and Yufeng Wu, "Investor Demand, Financial Market Power, and Capital Misallocation," *Financial Market Power, and Capital Misallocation (November 19, 2021)*, 2021.

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.

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.

Farboodi, Maryam and Laura Veldkamp, "Long-run Growth of Financial Data Technology," *American Economic Review*, 2020, *110* (8), 2485–2523.

_ , Roxana Mihet, Thomas Philippon, and Laura Veldkamp, "Big Data and Firm Dynamics," in "AEA papers and proceedings," Vol. 109 2019, pp. 38–42.

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

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.

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).

Jansen, Mark, Fabian Nagel, Constantine Yannelis, and Anthony Lee Zhang, "Data and Welfare in Credit Markets," *Available at SSRN 4015958*, 2022.

_ , Sam Kruger, and Gonzalo Maturano, "Dealer Financing in the Subprime Auto Market: Markups and Implicit Subsidies," *Available at SSRN 3902847*, 2021.

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.

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.

Maggio, Marco Di, Dimuthu Ratnadiwakara, and Don Carmichael, "Invisible primes: Fintech lending with alternative data," Technical Report, National Bureau of Economic Research 2022.

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

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.

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

Sutton, John, *Sunk costs and market structure: Price competition, advertising, and the evolution of concentration*, MIT press, 1991.

Tang, Huan, "The Value of Privacy: Evidence from Online Borrowers," 2019.

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

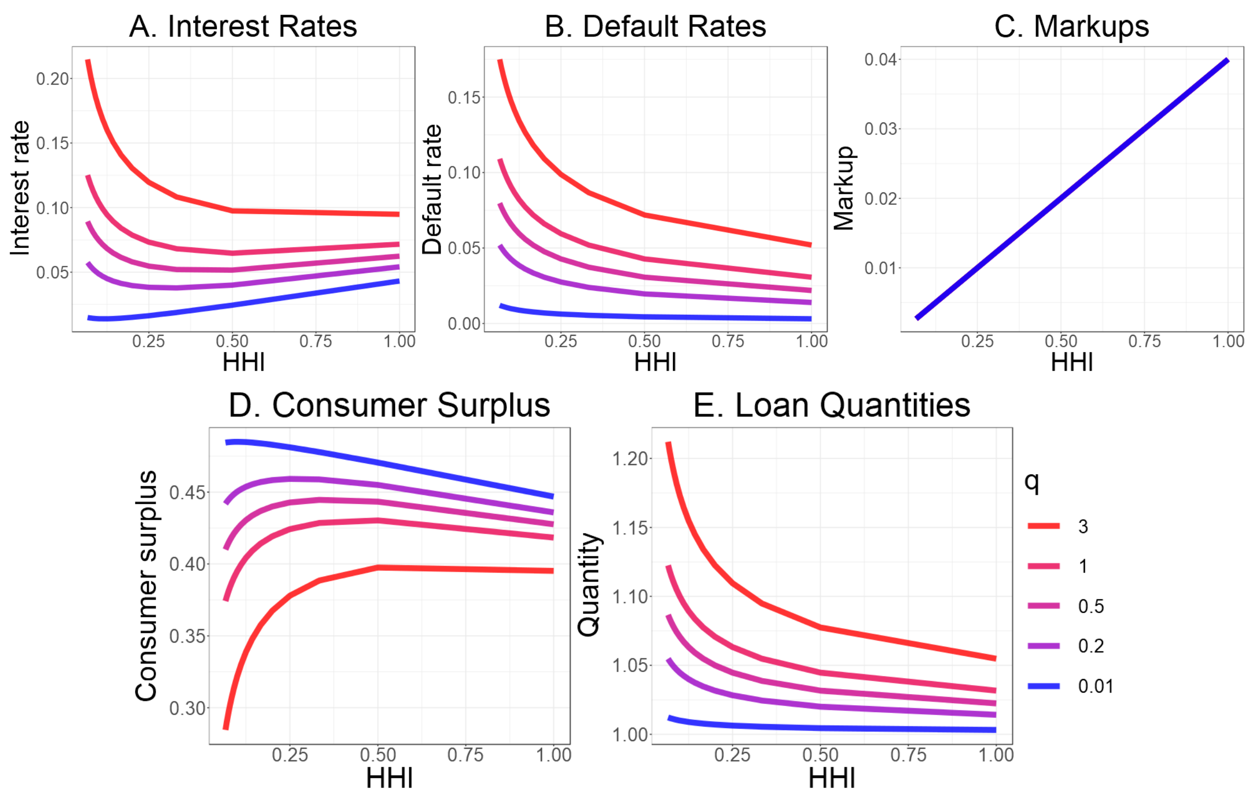
Vayanos, Dimitri and Jiang Wang, "Liquidity and Asset Returns Under Asymmetric Information and Imperfect Competition," *The Review of Financial Studies*, 2012, *25* (5), 1339–1365.

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

Wang, Yifei, Toni M Whited, Yufeng Wu, and Kairong Xiao, "Bank market power and monetary policy transmission: Evidence from a structural estimation," *The Journal of Finance*, 2022, *77* (4), 2093–2141.

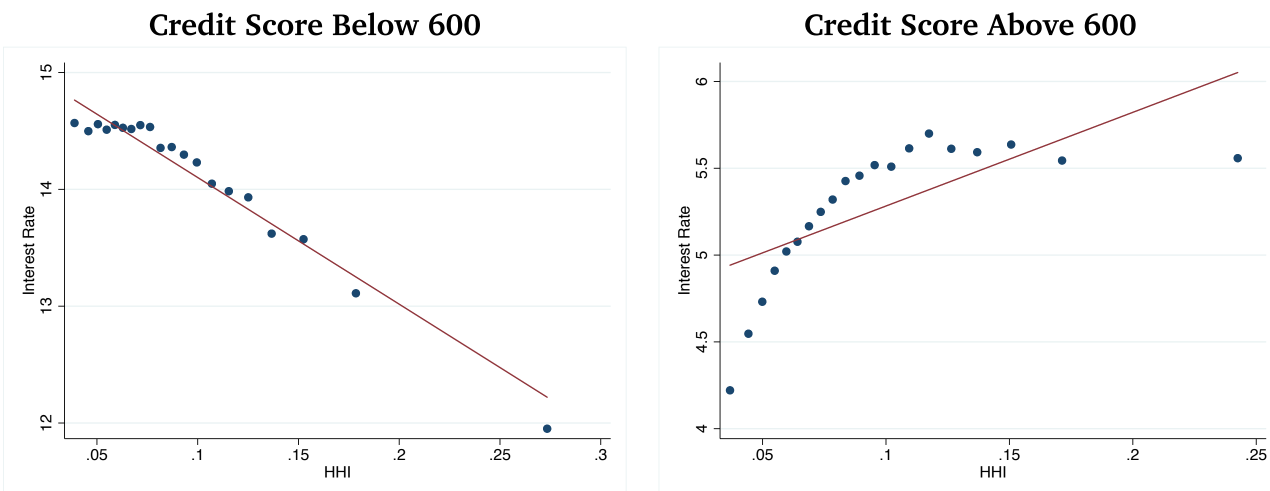
Whited, Toni M, Yufeng Wu, and Kairong Xiao, "Low interest rates and risk incentives for banks with market power," *Journal of Monetary Economics*, 2021, *121*, 155–174.

Figure 1: Model Simulations



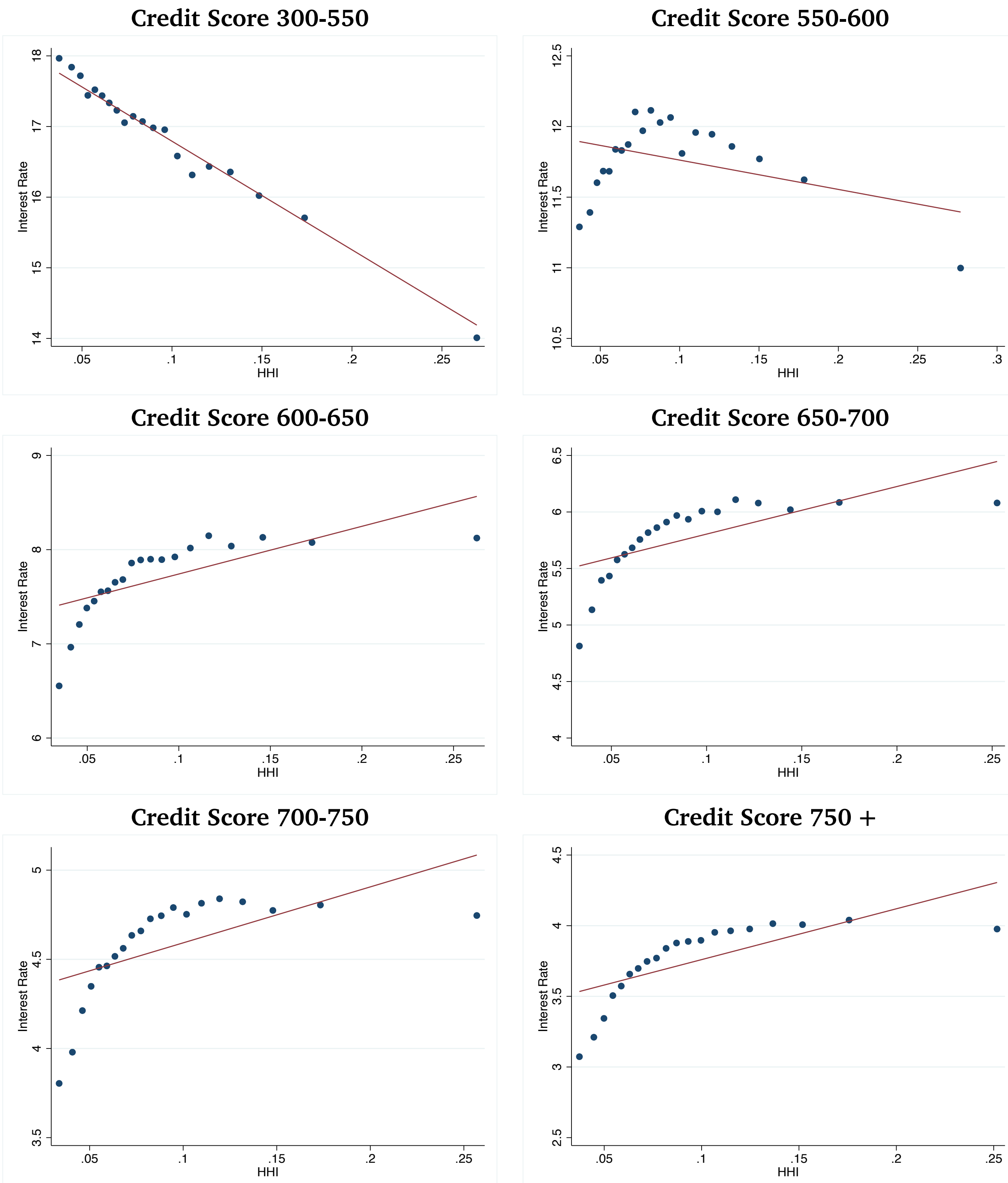
The above figure shows model outcomes, for different values of the Herfindahl-Hirschman index, HHI, and the measure of bad-type consumers, q. In all plots, bluer lines represent lower values of q (that is, lower baseline default rates in the consumer population) and redder liners represent higher values of q. Panel A shows equilibrium interest rates, r. Panel B shows default rates δ. Panel C shows markups over the break-even interest rate, $r - \frac{\delta}{1-\delta}$. Panel D shows consumer surplus, which is (12) in the text. Panel E shows total loan quantities. All simulations use $k = 0.001, \theta = 0.04, \mu = 0.5$.

Figure 2: Interest Rates and Competition



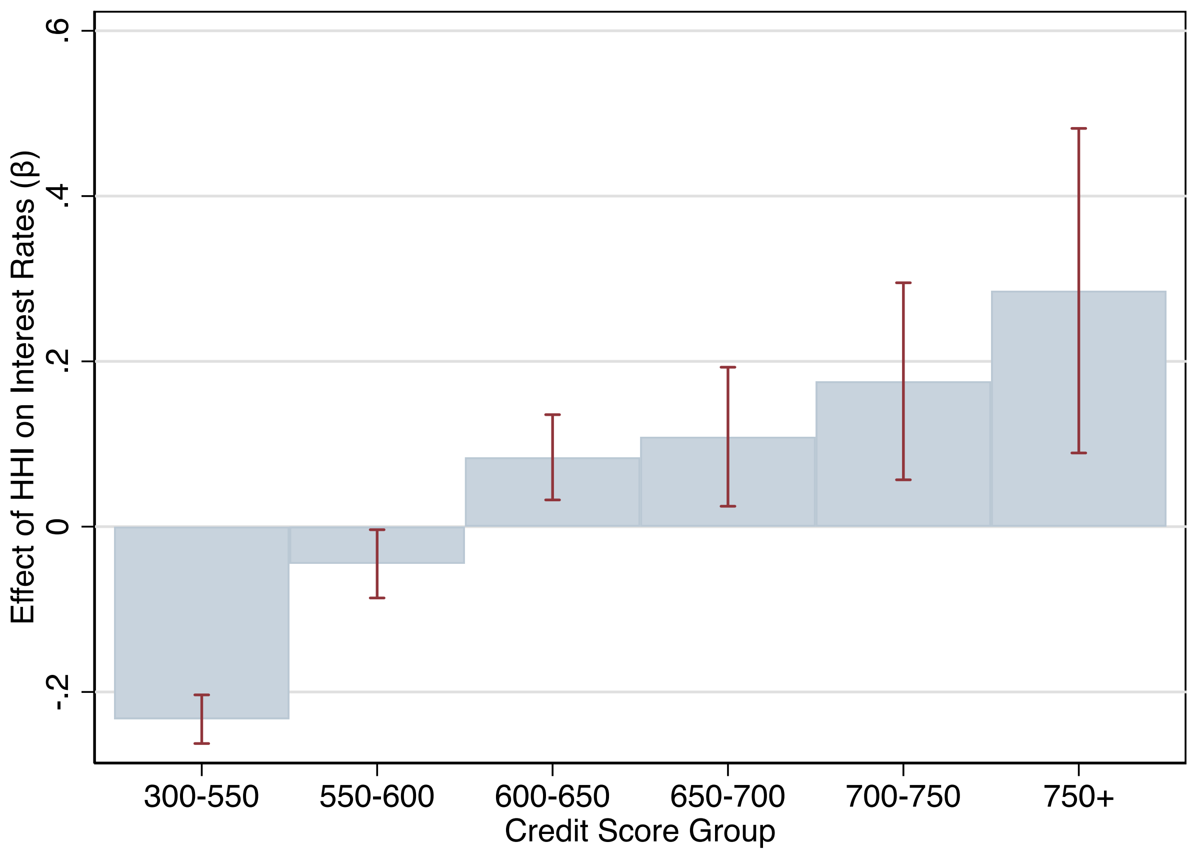
The above figure shows median interest rates in a county for given credit score ranges, broken down by ventile of HHI. HHI is defined using the volume of auto loans, that is $HHI = \sum_i^N s_i^2$, where s_i is a lender's share of auto lending in a county within a credit score range. Credit score ranges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

Figure 3: Interest Rates and Competition



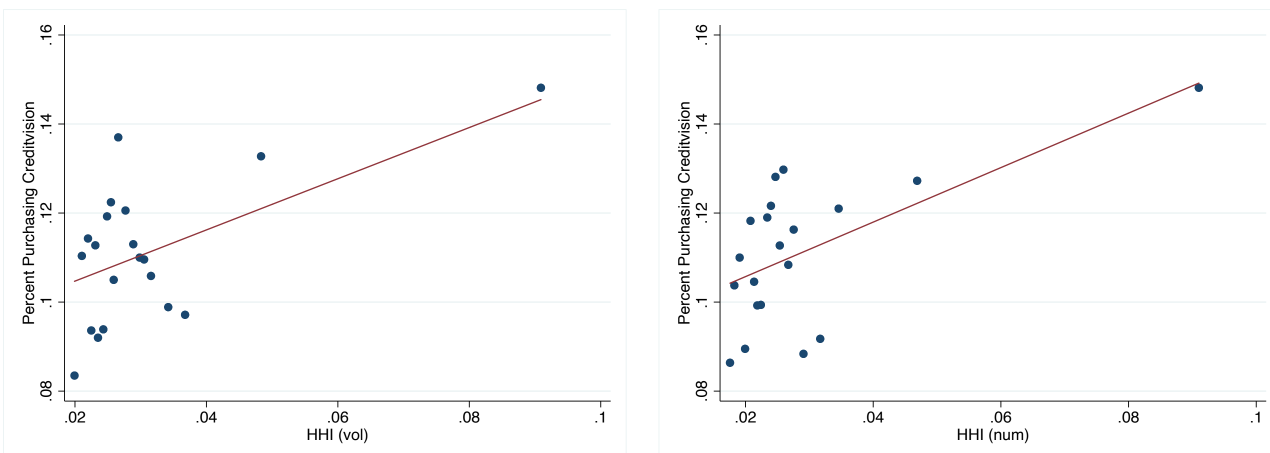
The above figure shows median interest rates in a county for given credit score ranges, broken down by ventile of HHI. HHI is defined using the volume of auto loans, that is $HHI = \sum_i^N s_i^2$, where s_i is a lender's share of auto lending in a county within a credit score range. Credit score ranges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

Figure 4: Market Concentration and Interest Rates



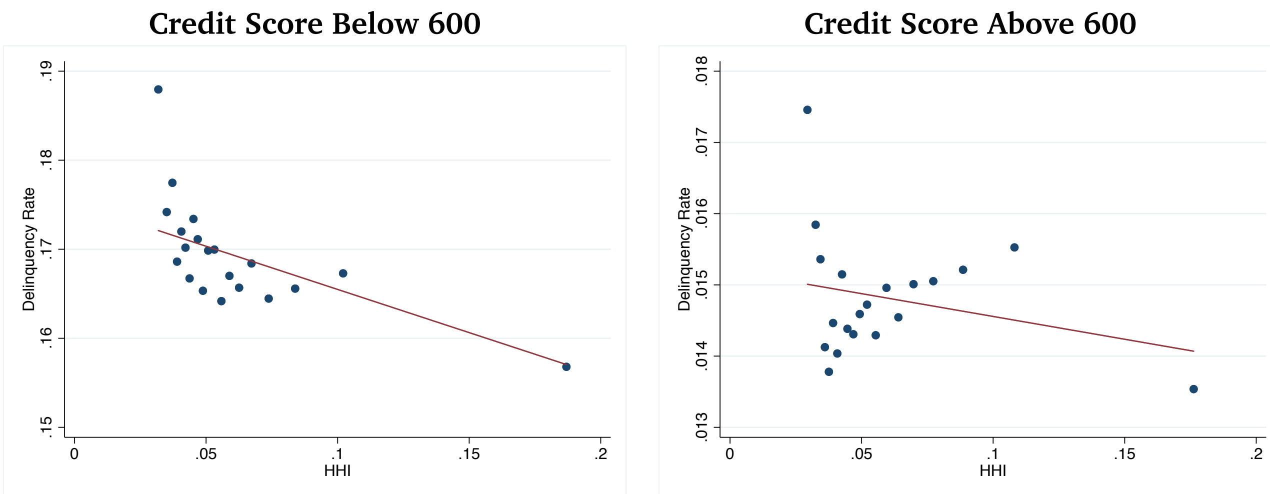
The above figure shows coefficients and a 95% confidence intervals from regressions of logged interest rates on interactions of logged HHI with indicators of credit score buckets. The specification includes year and score bucket fixed effects. Credit scores are given by VantageScore ratings. Source: TransUnion.

Figure 5: Screening and Competition



The above figure shows mean utilization of Creditvision, 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 state within a credit score range. Take-up of Creditvision is at the state level. Source: TransUnion.

Figure 6: Delinquency Rates and Competition



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.

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 by credit score. Each observation is a county by year by credit score bucket. Columns (1)-(6) split each county into loans above and below a 600 score, and regress logged interest rates on logged HHI. Columns (7)-(9) split the analysis sample into finer county by year by credit score bins, and regresses logged interest rates on interactions of logged HHI with indicators of credit score buckets. The inclusion of fixed effects is denoted beneath each column. Regressions are weighted by the number of loans. Note that interest rates are determined for each county, within each credit score bucket. 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)	(9)
	Ln(Interest Rate)								
	Credit Score 300-600			Credit Score 600-850			Full Sample		
Ln(HHI)	-0.0710***	-0.0816***	-0.0576***	0.275***	0.209***	0.0786			
	(0.0159)	(0.0164)	(0.0126)	(0.0674)	(0.0691)	(0.0492)			
Ln(HHI) × Credit Score 300-550							-0.134***	-0.233***	-0.182*
							(0.0113)	(0.0150)	(0.0943)
Ln(HHI) × Credit Score 550-600							0.0336*	-0.0450**	-0.151***
							(0.0192)	(0.0211)	(0.0369)
Ln(HHI) × Credit Score 600-650							0.144***	0.0839***	-0.0393
							(0.0256)	(0.0263)	(0.0301)
Ln(HHI) × Credit Score 650-700							0.156***	0.109**	-0.0202
							(0.0424)	(0.0429)	(0.0237)
Ln(HHI) × Credit Score 700-750							0.222***	0.176***	0.0277
							(0.0607)	(0.0608)	(0.0329)
Ln(HHI) × Credit Score 750+							0.346***	0.286***	0.140***
							(0.0989)	(0.100)	(0.0537)
Year Fixed Effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
County Fixed Effects	No	No	Yes	No	No	Yes	No	No	Yes
Score Bucket Fixed Effects	No	No	No	No	No	No	Yes	Yes	Yes
Observations	27,887	27,887	27,826	31,773	31,773	31,733	172,428	172,428	172,428
R^2	0.014	0.052	0.611	0.018	0.063	0.862	0.395	0.420	0.760

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, Loans and Market Competition

The table shows the relationship between delinquency or the number of loans and HHI, split by credit score. Columns (1)-(3) show delinquency rates, while columns (4)-(6) show loan quantities. The first two columns of each group split each county into loans above and below a 600 score, and regress logged outcomes on logged HHI. The third column of each pair splits the sample into finer county by year by credit score bins, and regresses logged outcomes on interactions of logged HHI with indicators of credit score buckets. 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(Delinquency)			Ln(Loans)		
	Credit Score			Credit Score		
	300-600	600-850	All	300-600	600-850	All
Ln(HHI)	-0.0718***	-0.101***		-0.164***	-0.103***	
	(0.0211)	(0.0156)		(0.0109)	(0.0126)	
Ln(HHI) × Credit Score 300-550			-0.194***			-0.251***
			(0.0116)			(0.0114)
Ln(HHI) × Credit Score 550-600			-0.236***			-0.198***
			(0.0121)			(0.00920)
Ln(HHI) × Credit Score 600-650			-0.202***			-0.198***
			(0.0103)			(0.00843)
Ln(HHI) × Credit Score 650-700			-0.138***			-0.217***
			(0.00948)			(0.00991)
Ln(HHI) × Credit Score 700-750			-0.0623***			-0.226***
			(0.00881)			(0.0114)
Ln(HHI) × Credit Score 750+			-0.0114			-0.419***
			(0.00909)			(0.0168)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Score Bucket Fixed Effects	No	No	Yes	No	No	Yes
Observations	27,826	31,733	172,428	27,826	31,733	172,428
R^2	0.613	0.527	0.665	0.972	0.983	0.952

Table 5: Alternative Explanations

The table shows which of the stylized facts various alternative theories are able to explain. Check marks mean that the theory matches the prediction, crosses mean that the theory makes the opposite prediction, and question marks mean that the theory is ambiguous.

More competition leads to:	High score $r \downarrow$	Low score $r \uparrow$	$\delta \uparrow$	$r \uparrow$ and $Q \uparrow$	Screening investments $\downarrow$	Additional Comments
Our theory	✓	✓	✓	✓	✓	
Adverse selection (Mahoney and Weyl, 2017)	✓	×	?	×	×	Prices always decrease as competition increases
Adverse selection (Crawford et al., 2018)	?	?	?	×	×	Price-competition relationship ambiguous, price and quantity can't simultaneously decrease
Relaxed loan standards	✓	×	✓	×	×	Prices always decrease as competition increases
Down payment screening	?	?	×	?	×	Predicts lower default rates in more competitive markets
Contract characteristics and search (Lester et al., 2019)	✓	?	?	?	×	Price-competition relationship ambiguous
Moral hazard	?	?	✓	×	×	Explains higher default conditional on higher rates, but doesn't explain higher rates
Dealer markups (Grunewald et al., 2020)	?	?	?	?	×	Doesn't explain dealer/non-dealer split-sample results
Heterogeneous funding costs	?	?	?	?	×	Doesn't explain dealer/non-dealer split-sample results
Improved collections technology	✓	×	?	×	×	Likely mostly marginal costs rather than fixed, so hard to explain rate increases with competition

Internet Appendix

A Proofs and supplementary material for Section 2

A.1 Optimal price setting

Differentiating (5) with respect to r_j, lenders' first-order condition for optimal price-setting is:

$$s_j + \frac{\partial s_j}{\partial r_j}\left(r_j - \frac{\delta_j}{1-\delta_j}\right) = 0 \tag{19}$$

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:

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

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,

$$\frac{\partial s_j}{\partial r_j} = -\frac{1}{\theta} \tag{20}$$

Thus, expression (19) becomes:

$$r_j - \frac{\delta_j}{1-\delta_j} = \frac{\theta}{N} \tag{21}$$

A.2 Optimal information acquisition

Differentiating (4) with respect to δ_j, we have:

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

this is (7).

A.3 Second-order condition

Next, we check negative definiteness of the Hessian matrix of profits (4), with respect to r_j and δ_j. The elements of the Hessian are as follows:

$$\frac{\partial^2 \Pi}{\partial \delta_j^2} = \frac{\partial}{\partial \delta_j}\left[-\frac{s_j\left(r_j\right)}{\left(1-\delta_j\right)^2} - c_q'\left(\delta_j\right)\right] = -\frac{2s_j\left(r_j\right)}{\left(1-\delta_j\right)^3} - c_q''\left(\delta_j\right) \tag{22}$$

$$\frac{\partial^2 \Pi}{\partial r_j^2} = \frac{\partial}{\partial r_j}\left[s_j\left(r_j\right) + \left(-\frac{1}{\theta}\right)\left(r_j - \frac{\delta_j}{1-\delta_j}\right)\right] = -\frac{2}{\theta} \tag{23}$$

$$\frac{\partial^2 \Pi}{\partial \delta_j \partial r_j} = \frac{\partial}{\partial \delta_j}\left[s_j\left(r_j\right) + \left(-\frac{1}{\theta}\right)\left(r_j - \frac{\delta_j}{1-\delta_j}\right)\right] = \frac{1}{\theta\left(1-\delta_j\right)^2} \tag{24}$$

The Hessian is negative definite if and only if $\frac{\partial^2 \Pi}{\partial r_j^2} < 0$, which is always true from (23), and the determinant is positive, so:

$$\frac{\partial^2 \Pi}{\partial r_j^2}\frac{\partial^2 \Pi}{\partial \delta_j^2} - \left(\frac{\partial^2 \Pi}{\partial \delta_j \partial r_j}\right)^2 > 0$$

Plugging in (22), (23), and (24), we have:

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

Rearranging, we have:

$$c_q''\left(\delta_j\right) > \frac{1}{2\theta\left(1-\delta_j\right)^4} - \frac{2}{N\left(1-\delta_j\right)^3}$$

Specializing to the case of symmetric equilibria, this is (11).

A.4 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.5 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:

$$\Pi = \left(\frac{1}{1-\delta_j}\right) s_j \left(r_j\left(1-\delta_j\right) - \delta_j \phi - \rho\right) \tag{25}$$

In words, (25) 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:

$$r_j - \frac{\delta_j \phi + \rho}{1-\delta_j} = \frac{\theta}{N} \tag{26}$$

Comparing (26) 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 similar to the main text; screening incentives are qualitatively unchanged, with the only quantitative difference being that screening incentives are slightly lower when recovery rates are higher.

A.6 Entry costs

In the baseline model, we have taken the number of firms N as exogeneous. N can be microfounded 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 exogeneously 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:

$$\frac{\theta}{N^2} - c_q\left(\delta\left(N\right)\right)$$

where $\delta\left(N\right)$ is the solution to (10), 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 $N\left(C_e\right)$ satisfies:

$$\frac{\theta}{\left(N\left(C_e\right)\right)^2} - c_q\left(\delta\left(N\left(C_e\right)\right)\right) > C_e \tag{27}$$

$$\frac{\theta}{\left(N\left(C_e\right)+1\right)^2} - c_q\left(\delta\left(N\left(C_e\right)+1\right)\right) < C_e \tag{28}$$

That is, firms make profits greater than C_e with $N\left(C_e\right)$ entrants, but not with $N\left(C_e\right)+1$ entrants. Using expressions (27) and (28), 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.7 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 $\phi\left(r\right)$, 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:

$$\psi\left(r,\delta\right) \equiv \delta + \left(1-\delta\right)\phi\left(r\right) \tag{29}$$

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

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

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

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

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 (20) from Appendix A.1, we have $s_j'=-\frac{1}{\theta}$. Rearranging,

$$\left(r_j-\frac{\psi\left(r_j,\delta_j\right)}{1-\psi\left(r_j,\delta_j\right)}\right)\left(-\left(1-\psi\left(r_j,\delta_j\right)\right)\frac{1}{\theta}-\frac{d\psi}{dr_j}s_j\right)=$$
$$-\left[\left(1-\psi\left(r_j,\delta_j\right)\right)s_j-\left(1-\psi\left(r_j,\delta_j\right)\right)s_j\left(\frac{\partial}{\partial r_j}\left(\frac{\psi\left(r_j,\delta_j\right)}{1-\psi\left(r_j,\delta_j\right)}\right)\right)\right] \quad (30)$$

Now, we have:

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

Hence, (30) becomes:

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

Rearranging, we have:

$$r_j - \frac{\psi\left(r_j,\delta_j\right)}{1-\psi\left(r_j,\delta_j\right)} = \left(\left(1-\psi\left(r_j,\delta_j\right)\right)\frac{1}{\theta} + \left(1-\delta_j\right)\frac{d\phi}{dr_j}s_j\right)^{-1}\left[1-\frac{1-\delta_j}{\left(1-\psi\left(r_j,\delta_j\right)\right)^2}\frac{d\phi}{dr_j}\right]\left(1-\psi\left(r_j,\delta_j\right)\right)s_j \tag{31}$$

When there is no moral hazard, $\phi(r)=0$, so $\frac{d\phi}{dr_j}=0$, so (31) reduces to (6) in the main text. When there is moral hazard, the RHS of (31) 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:

$$\max_{\delta_j}\max_{r_j}\left(\frac{1}{1-\delta_j}\right)s_j\left(r_j\left(1-\psi\left(r_j,\delta_j\right)\right)-\psi\left(r_j,\delta_j\right)\right)-c_q\left(\delta_j\right) \tag{32}$$

Using (29), the objective in (32) becomes:

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

This rearranges to:

$$= s_j\left(r_j - \frac{\delta_j}{1-\delta_j}\right) + s_j\left(-\phi\left(1+r_j\right)\right) - c_q\left(\delta_j\right) \tag{33}$$

As in Appendix A.2, the envelope theorem applies, so lenders' investment incentives are determined by differentiating (33) with respect to δ_j, holding s_j and r_j fixed. This means that optimal investments are determined by the first-order-condition:

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

which is identical to (7) 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 constant. We will assume that moral hazard has a linear effect on type-G customers' default rates:

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

This gives:

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

Plugging into (31), we have:

$$r_j - \frac{\psi\left(r_j, \delta_j\right)}{1 - \psi\left(r_j, \delta_j\right)} = \left(\left(1 - \psi\left(r_j, \delta_j\right)\right)\frac{1}{\theta} + \left(1 - \delta_j\right)\chi s_j\right)^{-1}\left[1 - \frac{1 - \delta_j}{(1 - \psi)^2}\chi\right]\left(1 - \psi\left(r_j, \delta_j\right)\right)s_j \tag{34}$$

In any symmetric equilibrium, we have $s_j = \frac{1}{N}$. Using (34), 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.8 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 $s_j\left(r_j\right)$ 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:

$$\Pi_j = \underbrace{s_j\left(r_j\right)r_j}_{Good\ type\ profits} \underbrace{-s_j\left(r_j\right)q\alpha_j\alpha_{-j}}_{Group\ 1\ losses} \underbrace{-\frac{2}{N}q\alpha_j\left(1-\alpha_{-j}\right)}_{Group\ 2\ losses} -\tilde{c}\left(\alpha_j\right) \tag{35}$$

In words, for the measure $s_j\left(r_j\right)$ of price-elastic type-G consumers, the lender makes r_j. 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 $\frac{2q}{N}\alpha_j\left(1-\alpha_{-j}\right)$ 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 $\tilde{c}\left(\alpha_j\right)$.

Price-setting. Differentiating (35) with respect to r_j, we have:

$$s_j'\left(r_j - q\alpha_j\alpha_{-j}\right) + s_j = 0 \tag{36}$$

Using (20) from Appendix A.1, we have $s_j' = -\frac{1}{\theta}$. Hence, (36) rearranges to:

$$r_j - q\alpha_j\alpha_{-j} = \theta s_j \tag{37}$$

(37) 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 $q\alpha_j\alpha_{-j}$ type-B's. Hence, this is the relevant marginal cost for lenders' markups. In symmetric equilibrium, $s_j = \frac{1}{N}$, so markups over marginal costs still decrease as N increases.

Screening. To solve for optimal screening decisions, we differentiate (35) with respect to

α_j. Using the envelope theorem, we can ignore effects of changes in r_j. This gives:

$$\frac{d\Pi_j}{d\alpha_j} = -s_j\left(r_j\right)q\alpha_{-j} - \frac{2q\left(1-\alpha_{-j}\right)}{N} - \tilde{c}'\left(\alpha_j\right)$$

$$-\tilde{c}'\left(\alpha_j\right) = s_j\left(r_j\right)q\alpha_{-j} + \frac{2q\left(1-\alpha_{-j}\right)}{N}$$

In symmetric equilibrium, $s_j\left(r_j\right) = \frac{1}{N}$, hence we have:

$$-\tilde{c}'\left(\alpha_j\right) = \frac{\left(2-\alpha_{-j}\right)q}{N} \tag{38}$$

(38) 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, (38) 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.9 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:

$$\mu - \alpha r_j + \epsilon_{ij} \tag{39}$$

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 auto 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.[26] Given lenders' interest rates r_j, the

[26]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.

market share of lender j is:

$$s_j = \frac{\exp\left(\mu - \alpha r_j\right)}{1 + \sum_j \exp\left(\mu - \alpha r_j\right)} \tag{40}$$

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

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

Appendix A.10.1 shows that lenders' optimal prices satisfy:

$$r_j - \frac{\delta_j}{1-\delta_j} = \frac{1}{\alpha\left(1-s_j\right)} \tag{41}$$

The intuition for (41) 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.10.2 shows that lenders' FOC for optimal information acquisition is unchanged from the main text:

$$\frac{s_j}{\left(1-\delta_j\right)^2} = -c_q'\left(\delta_j\right) \tag{42}$$

Expressions (40) and (41) are thus necessary conditions for a symmetric equilibrium.[27] Equilibrium consumer surplus, for type G consumers, follows from the standard logit surplus formula, from, for example, Train (2009) :

$$\log\left(1 + \sum_{j=1}^{J} \exp\left(\mu - \alpha r\right)\right) + C \tag{43}$$

where C is a constant.

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

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

[27]We note that these conditions are necessary, but not sufficient, for equilibrium, and that the model does not always have unique equilibria, even when $c_q\left(\cdot\right)$ 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 $c_q\left(\cdot\right)$, there is no $\delta < 1$ which satisfies all three conditions simultaneously.

implying that:

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

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 q is lower and the population is low-risk, and negative when q 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 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 q is very large – the red line – increasing N and decreasing HHIs can still decrease consumer surplus in the logit model: note that the red consumer surplus line becomes flat and dips slightly downwards for HHIs close to 0.1.

A.10 Proofs for Appendix A.9

A.10.1 Optimal price setting

From (40), 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:

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

Differentiating (4) with respect to r_j, the first-order condition for optimal price-setting is:

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

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

This simplifies to:

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

This is (41).

A.10.2 Optimal information acquisition

Differentiating (4) with respect to δ_j, we have, identically to the baseline model:

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

This is (42), which is identical to the optimal screening condition (7) in the baseline model.

A.11 CES Model

In this appendix, we consider another alternative model of demand, the constant-elasticity-of-substitution model. Using the model, we do an extended back-of-envelope calculation, to approximately quantify how much the welfare gains from increased product variety may offset the welfare losses from increased prices.

As in the main text, we assume there is a unit measure of good type consumers who never default. There is a measure $\frac{\delta_j}{1-\delta_j}$ of bad type consumers who always default; δ_j is determined through lenders' fixed-cost screening decisions. To allow us to map the model to data, we also assume lenders have a cost of funds ρ for each unit of funds that they borrow, as in Appendix A.5. We assume both good and bad type consumers demand loans according to a CES aggregator, over loan quantity q_i and money y:

$$U\left(y, q_1 \ldots q_n\right) = \frac{\zeta}{\zeta - 1}\left(\left(\sum_{i=1}^{n} q_i^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}\right)^{\frac{\zeta-1}{\zeta}} - y \tag{44}$$

As is standard in nested CES models, we assume that $\sigma > \zeta > 1$. The parameter σ represents the elasticity of substitution between different lenders' loans, and the parameter ζ represents the elasticity of substitution between loans and money.

Consumers are initially endowed with income M. Thus, if lenders set rates $r_1 \ldots r_n$, consumers solve:

$$\max_{q_1 \ldots q_n} \frac{\zeta}{\zeta - 1}\left(\left(\sum_{i=1}^{n} q_i^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}\right)^{\frac{\zeta-1}{\zeta}} + M - \sum_{i=1}^{n} q_i r_i \tag{45}$$

Lenders' profit is:

$$\Pi = q_j\left(r_j\right)\left(\frac{1}{1-\delta_j}\right)\left(r_j\left(1-\delta_j\right) - \delta_j - \rho\right) - c_q\left(\delta_j\right) \tag{46}$$

Rearranging somewhat, we have:

$$\Pi = q_j\left(r_j\right)\left(r_j - \frac{\delta_j + \rho}{1-\delta_j}\right) - c_q\left(\delta_j\right) \tag{47}$$

To solve for lenders' optimal price-setting decisions, we differentiate (47), to get:

$$q_j'\left(r_j\right)\left(r_j - \frac{\delta_j + \rho}{1-\delta_j}\right) + q_j = 0$$

This gives a markup formula analogous to (9) in the main text:

$$r_j - \frac{\delta_j + \rho}{1-\delta_j} = \frac{-q_j}{q_j'\left(r_j\right)} \tag{48}$$

Optimal information acquisition is also analogous to the main text; differentiating (47) with respect to δ_j, we have:

$$\frac{\partial}{\partial \delta_j} q_j\left(r_j\right)\left(r_j - \frac{\delta_j + \rho}{1-\delta_j}\right) - c_q\left(\delta_j\right) = -\frac{q_j\left(r_j^*\right)(1+\rho)}{\left(1-\delta_j\right)^2} - c_q'\left(\delta_j\right)$$

Hence,

$$c_q'\left(\delta_j\right) = -\frac{q_j\left(r_j^*\right)(1+\rho)}{\left(1-\delta_j\right)^2}$$

analogous to the main text. The following claim characterizes analytical expressions for consumer demand, markups, and consumer welfare.

Claim 1. *Define the CES price index:*

$$R \equiv \left(\sum_{i=1}^{n} r_i^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \tag{49}$$

Consumers' demand for lender i is:

$$q_i\left(r_1 \ldots r_N\right) = r_i^{-\sigma} R^{\sigma-\zeta} \tag{50}$$

The demand elasticity is:

$$\frac{\partial q_i}{\partial r_i}\frac{r_i}{q_i} = -\left(s_i\zeta + \sigma\left(1-s_i\right)\right) \tag{51}$$

Interest rates are:

$$\frac{r_i - \frac{\delta_i+\rho}{1-\delta_i}}{r_i} = \frac{1}{s_i\zeta + \sigma\,(1-s_i)} \tag{52}$$

where s_i is the expenditure share of lender i:

$$s_i \equiv \frac{r_i q_i}{\sum_{j=1}^{n} r_j q_j} \tag{53}$$

In symmetric equilibrium, where all lenders are identical, we have $s_i = \frac{1}{n}$, hence we have:

$$\frac{r_i - \frac{\delta_i+\rho}{1-\delta_i}}{r_i} = \frac{1}{\frac{\zeta}{n} + \sigma\left(1-\frac{1}{n}\right)} \tag{54}$$

Consumer welfare is:

$$M + \frac{1}{(\zeta-1)R^{\zeta-1}} \tag{55}$$

The CES price index R is thus a sufficient statistic for consumer welfare; consumer welfare is strictly decreasing in R.

The core equation in Claim 1 is the markup formula, (54). The LHS of (54) is the Lerner index, that is, the fraction of the price r_j which consists of the lender's markup. The RHS of (54) is the inverse of the elasticity of demand. Consumer demand in this model is effectively a nested CES. The elasticity of substitution between different lenders is σ; the elasticity of substitution between aggregated loans and the outside option is the lower quantity ζ. Thus, when $n = 1$ and there is a monopolist lender, the RHS of (54) is the higher quantity $\frac{1}{\zeta}$. When n approaches ∞, the RHS of (54) approaches the lower quantity $\frac{1}{\sigma}$. As n increases, lenders' markups thus decrease from $\frac{1}{\zeta}$ to $\frac{1}{\sigma}$. Note that, since the CES is a differentiated-products model, lenders are able to charge markups even as the number of lenders becomes infinite. This contrasts with the Salop circle model in the main text; (9) of Proposition 1 shows that, in the Salop circle model, markups decline to 0 as the number of lenders grows to infinity.

Back-of-envelope quantification. Using the model, we then conduct an extended back-of-envelope calculation to approximate how much the welfare gains from increased product variety could offset increased interest rates. We assume that consumers in all credit score groups have the same elasticities ζ and σ. We then identify ζ, σ by focusing on the highest credit score group, of consumers with credit scores above 800. Default rates for this group are very low; the median 90-day delinquency rate across county-years is only 0.18%. We will thus assume that $\delta = 0$ for this group of consumers. Given this assumption, the markup formula

(54) simplifies to:

$$\frac{r_i - \rho}{r_i} = \frac{1}{\frac{\zeta}{n} + \sigma\left(1 - \frac{1}{n}\right)} \tag{56}$$

For a monopolistic market, where $n = 1$, (56) states that:

$$\frac{r_i - \rho}{r_i} = \frac{1}{\zeta} \tag{57}$$

that is, the markup depends only on the elasticity ζ between loans and the outside option. For very competitive markets, as $n \to \infty$, we have:

$$\frac{r_i - \rho}{r_i} \to \frac{1}{\sigma} \tag{58}$$

that is, the markup depends only on the elasticity of substitution between lenders, σ. We will identify ζ and σ simply by attempting to measure markups in very concentrated and very competitive markets, respectively. As a proxy for funding costs ρ, we use the 5-year high quality market corporate bond spot rate from FRED, which is series HQMCB5YR. We then consider the 5% of most and least concentrated county-years in our data, respectively; the average HHIs within these groups are respectively 0.69 and 0.04.

In each of the high- and low-concentration cases, we calculate average interest rates across counties in each year from 2009-2019, then take the average rate across all years as a representative observation for r_i. We then calculate the average 5-year corporate bond rate over 2009-2019 as an observation for ρ. This process gives an interest rate of $r_i = 4.43\%$ for very concentrated markets, and $r_i = 3.09\%$ for very competitive markets, and a funding cost of $\rho = 2.82\%$. Plugging these values into (57) and (58), we estimate $\zeta = 2.756$, and $\sigma = 11.42$.

Using these estimates, we can then evaluate whether consumers in different credit score groups are better or worse off when markets are more competitive. In low credit-score groups, we found empirically that interest rates are higher when markets are more competitive, since default rates are higher. However, despite higher interest rates in these groups, consumer welfare may still increase with higher competition, if consumers' preferences for variety are strong enough. In the context of the CES model, we can evaluate consumer welfare for any market simply by measuring the price index, (49), for markets with different levels of competition. When there are n lenders who set symmetric interest rates $r_i = r$, the price index simplifies to:

$$R = \left(\sum_{i=1}^{n} r^{1-\sigma}\right)^{\frac{1}{1-\sigma}} = n^{\frac{1}{1-\sigma}} r \tag{59}$$

We will use expression (59) to calculate the price index R from the HHI and average interest rates for each county-year in our data. If there are n symmetric lenders, the HHI is simply:

$$HHI = \sum_{i=1}^{n} \frac{1}{n^2} = \frac{1}{n}$$

We thus calculate an implied number of lenders in each county-year simply as:

$$\hat{n}_{ct} = \frac{1}{HHI_{ct}} \tag{60}$$

In each county-year, we then calculate a price index by plugging the observed rate r and $\hat{n}_{ct}$ into (59), using our estimate $\sigma = 11.42$. We then calculate consumer welfare for each county-year by plugging the price index into (55), ignoring the constant M. The intuition behind our calculation is as follows. Suppose interest rates increase as markets become more competitive, that is, the number of lenders n increases. There are thus two competing forces acting on the price index R in (59), and thus consumer welfare. On the one hand, the rate r increases for all lenders, which tends to increase R and decrease consumer welfare. On the other hand, n is larger, decreasing R and increasing consumer welfare.

The results of this exercise are shown in Figure A.4. We find that, for the lowest credit score group, consumer welfare is in fact higher in more concentrated markets: the decreases in interest rates are strong enough that consumers prefer concentrated markets, despite the fact that there is less surplus from lender variety. For all other credit score groups, welfare is mostly lower in more concentrated markets.

We caveat that this back-of-envelope calculation requires a number of assumptions: for example, funding costs must equal 5-year corporate bond rates, and elasticities of substitution must be equal across all credit score groups. The functional form used for consumer utility also has important effects on outcomes: in the logit model of Appendix A.9, gains from product variety also do not decrease to zero asymptotically as the number of lenders increases, whereas in the Salop circle model of the main text, gains from variety do decrease to 0 as the number of lenders increases. It is thus difficult to take a strong quantitative stance on the gains from variety in a model-free way. This exercise also does not represent a full social welfare analysis, since we have not taken into account lenders' entry and screening costs. However, our extended back-of-envelope calculation suggests that, under realistic parameter settings that are consistent with crude measures of markups in the data, the effect of increased interest rates can be strong enough in the lowest credit score group to offset the gains from variety from increased competition, making consumers worse off in more competitive markets.

A.11.1 Proof of Claim 1

Demand. First, to solve for demand, we take derivatives of (45) with respect to each q_j:

$$\frac{\partial}{\partial q_j}:\left(\left(\sum_{i=1}^{n} q_i^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}\right)^{-\frac{1}{\zeta}}\left(\sum_{i=1}^{n} q_i^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{\sigma-1}} q_j^{-\frac{1}{\sigma}}-r_j=0\ \forall j \tag{61}$$

Rearranging, we have:

$$\left(\frac{q_j}{q_i}\right)^{-\frac{1}{\sigma}}=\frac{r_j}{r_i}$$

for all i, j. That is,

$$q_i=q_j\left(\frac{r_j}{r_i}\right)^{\sigma} \tag{62}$$

Plugging (62) into (61), we have:

$$q_j^{-\frac{1}{\zeta}} r_j^{-\frac{\sigma}{\zeta}+1}\left(\left(\sum_{i=1}^{n}\left(\frac{1}{r_i}\right)^{\sigma-1}\right)^{\frac{\sigma}{\sigma-1}}\right)^{-\frac{1}{\zeta}}\left(\sum_{i=1}^{n}\left(\frac{1}{r_i}\right)^{\sigma-1}\right)^{\frac{1}{\sigma-1}}-r_j=0 \tag{63}$$

Using the definition of the price index in (49), (63) rearranges to:

$$q_i\left(p_1 \ldots p_N\right)=r_i^{-\sigma} R^{\sigma-\zeta} \tag{64}$$

This is (50) of Claim 1. Note also that total expenditures of the consumer on lender i are thus:

$$r_i q_i=r_i^{1-\sigma} R^{\sigma-\zeta}$$

This implies that the share of expenditure on lender i, defined in (53), becomes:

$$s_i \equiv \frac{r_i q_i}{\sum_{j=1}^{n} r_j q_j}=\frac{r_i^{1-\sigma}}{\sum_{j=1}^{n} r_j^{1-\sigma}}=\left(\frac{r_i}{R}\right)^{1-\sigma} \tag{65}$$

Demand elasticities. To calculate demand elasticities, take logs of the demand equation (64) to get:

$$\log q_i=-\sigma \log r_i+(\sigma-\zeta) \log R$$

Hence,

$$\frac{\partial \log q_i}{\partial \log r_i}=-\sigma+(\sigma-\zeta) \frac{\partial \log R}{\partial \log r_i}$$

From (49), we also have:

$$\log R = \frac{1}{1-\sigma}\log\left(\sum_{i=1}^{n} r_i^{1-\sigma}\right)$$

Thus,

$$\frac{\partial \log R}{\partial \log r_i} = \frac{\partial \log R}{\partial r_i} r_i = \left(\frac{1}{1-\sigma}\frac{1}{\sum_{j=1}^{n} r_j^{1-\sigma}}(1-\sigma) r_i^{-\sigma}\right) r_i = \frac{r_i^{1-\sigma}}{\sum_{j=1}^{n} r_j^{1-\sigma}} = \frac{r_i^{1-\sigma}}{R^{1-\sigma}}$$

Hence we have:

$$\frac{\partial \log q_i}{\partial \log r_i} = -\sigma + (\sigma - \zeta)\frac{r_i^{1-\sigma}}{R^{1-\sigma}} \tag{66}$$

Using the expenditure share equation (65), we can also write (66) as:

$$\frac{\partial \log q_i}{\partial \log r_i} = -\sigma + (\sigma - \zeta) s_i$$

This rearranges to (51) of Claim 1.

Markups. From (48), lenders' FOC is:

$$r_j - \frac{\delta_j + \rho}{1-\delta_j} = \frac{-q_j}{q_j'(r_j)}$$

Rearranging slightly,

$$\frac{r_j - \frac{\delta_j+\rho}{1-\delta_j}}{r_j} = -\frac{1}{\frac{\partial q_j}{\partial r_j}\frac{r_j}{q_j}} = -\frac{1}{\frac{\partial \log q_j}{\partial \log r_j}}$$

Plugging in for $\frac{\partial \log q_j}{\partial \log r_j}$ using (51), we have:

$$\frac{r_j - \frac{\delta_j+\rho}{1-\delta_j}}{r_j} = \frac{1}{s_j\zeta + \sigma(1-s_j)}$$

This is (52) of Claim 1.

Welfare. First, note that we can write consumer utility as:

$$U(y, q_1 \ldots q_n) = \frac{\zeta}{\zeta - 1} Q(q_1 \ldots q_n)^{\frac{\zeta-1}{\zeta}} - y$$

where:

$$Q(q_1 \ldots q_n) = \left(\sum_{i=1}^{n} q_i^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{67}$$

can be thought of as a composite good, constructed based on loan output of each lender. Now, repeating consumers' first-order condition in (62), we have:

$$q_i = q_j \left(\frac{r_i}{r_j}\right)^{-\sigma}$$

Multiplying both sides by r_i, we have:

$$q_i r_i = q_j r_i^{1-\sigma} r_j^{\sigma}$$

Summing over all i, we have:

$$\sum_{i=1}^{n} q_i r_i = \sum_{i=1}^{n} q_j r_i^{1-\sigma} r_j^{\sigma}$$

$$\sum_{i=1}^{n} q_i r_i = q_j r_j^{\sigma} \sum_{i=1}^{n} r_i^{1-\sigma}$$

Rearranging, we have:

$$q_j = \frac{\sum_{i=1}^{n} q_i r_i}{\left(\sum_{i=1}^{n} r_i^{1-\sigma}\right) r_j^{\sigma}} \tag{68}$$

Note that the numerator of (68) is the total expenditures the consumer spends on loans; let us define this as $I \equiv \sum_{i=1}^{n} q_i r_i$. The term in the denominator includes the *price index* of the CES demand system; let us define this as $R \equiv \left(\sum_{i=1}^{n} r^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$, as in (49). We can then write (68) as:

$$q_j = \frac{I}{R^{1-\sigma} r_j^{\sigma}} \tag{69}$$

Expression (69) can be thought of as the Marshallian demand function for lender j, if a total amount I is spent on all loans, given the price index R and the price r_j of lender j. Now, if we plug (69) into the production function (67) for the aggregate good, we have:

$$Q(q_1 \dots q_n) = \left(\sum_{i=1}^{n} q_i^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} = \left(\sum_{i=1}^{n} \left(\frac{I}{R^{1-\sigma} r_i^{\sigma}}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} = \left(\sum_{i=1}^{n} I^{\frac{\sigma-1}{\sigma}} R^{\frac{\sigma-1}{\sigma}(\sigma-1)} r_i^{1-\sigma}\right)^{\frac{\sigma}{\sigma-1}}$$

$$= \left(I^{\frac{\sigma-1}{\sigma}} R^{\frac{\sigma-1}{\sigma}(\sigma-1)} \sum_{i=1}^{n} r_i^{1-\sigma}\right)^{\frac{\sigma}{\sigma-1}} = IR^{(\sigma-1)} \left(R^{1-\sigma}\right)^{\frac{\sigma}{\sigma-1}} = IR^{(\sigma-1)} R^{-\sigma} = \frac{I}{R}$$

We have thus shown that, if income I is spent on loans, the price index is R, and borrowers' first-order conditions for optimal consumption hold, then the quantity of the composite good purchased is simply $\frac{I}{R}$. Hence, R is the optimized cost of producing a unit of the composite good.

Borrowers' optimal consumption problem can thus be formulated simply as:

$$\max_{Q} \frac{\zeta}{\zeta-1} Q^{\frac{\zeta-1}{\zeta}} + M - RQ$$

where we have ignored a constant representing consumers' initial wealth. Differentiating, the first-order condition is:

$$Q^{\frac{-1}{\zeta}} - R = 0$$

Optimal consumption is:

$$Q^{*}(R) = \frac{1}{R^{\zeta}}$$

The representative consumer's welfare is:

$$U\left(Q^{*}(R)\right) = \frac{\zeta}{\zeta-1}\left(\frac{1}{R^{\zeta}}\right)^{\frac{\zeta-1}{\zeta}} + M - R\left(\frac{1}{R^{\zeta}}\right)$$

$$= M + \frac{\zeta}{\zeta-1}\frac{1}{R^{\zeta-1}} - \left(\frac{1}{R^{\zeta-1}}\right) = M + \frac{1}{\zeta-1}\frac{1}{R^{\zeta-1}}$$

This proves (55). Thus, the price index R is a sufficient statistic for consumer welfare. Since we assumed $\zeta > 1$, consumer welfare is strictly decreasing in R.

B Supplementary material for Section 4

B.1 Relative Magnitudes of Coefficient Estimates

Our model attributes the decrease in interest rates for subprime customers in more concentrated markets to a decrease in deliquency rates. Our empirical results are qualitatively consistent with our model: in more concentrated markets, delinquency rates and interest rates are both lower for subprime consumers. A further test of our model is to ask quantatively whether the drop in delinquency rates for subprime consumers in concentrated markets is large enough, under reasonable assumptions about default rates, to drive the observed decrease in subprime interest rates. We can also ask how large the implied effects of concentration on interest rates and default rates are for prime consumers.

Since the specifications in Table 2 regress log interest rates on log HHIs, changing log HHIs by 1 is associated with an increase in interest rates of approximately:

$$\beta_{s}^{r,HHI}\bar{r}_{s} \tag{70}$$

where $\beta_s^{r,HHI}$ is the coefficient on $log(HHI_{cst})$ from Table 2 for score group s, and $\bar{r}_s$ is the mean interest rate for group s. The median interest rate for prime borrowers is 4.781%, and the median interest rate for subprime borrowers is 13.99%. We construct upper and lower estimates of $\beta_s^{r,HHI}$, the effect of HHI on interest rates, for prime (subprime) consumers by taking the coefficients in columns 4-6 (1-3) of Table 2, adding and subtracting two times the standard error of each coefficient, and taking the largest and smallest coefficients thus obtained. These upper and lower estimates thus incorporate uncertainty from different specifications, as well as statistical error within each specifications. Thus, we find that increasing log HHIs by 1 increases prime interest rates by −0.09% to 1.95%, and decreases subprime interest rates by 0.45% to 1.60%.

Next, we aim to estimate costs. We can estimate the effect of HHI on delinquency rates using specifications analogous to those in Table 2. These estimates imply that a change in log HHIs of 1 is associated with a change in delinquency rates of:

$$\beta_s^{r,del}\bar{\delta}_s \tag{71}$$

where $\beta_s^{r,del}$ is the coefficient on $log(HHI_{cst})$ for score group s, and $\bar{\delta}_s$ is the mean deliquency rate for group s. The mean 90-day delinquency rate for prime borrowers is around 1.42%, and it is 9.65% for subprime borrowers. We construct upper and lower estimates of $\beta_s^{r,del}$ analogously to the interest rate effects, adding and subtracting two times the standard error, and taking the maximum and minimum of the resultant numbers. We find that increasing log HHIs by 1 decreases prime delinquency rates by around −0.05% to 0.08%, and decreases subprime delinquency rates by −0.3% to 1.67%.

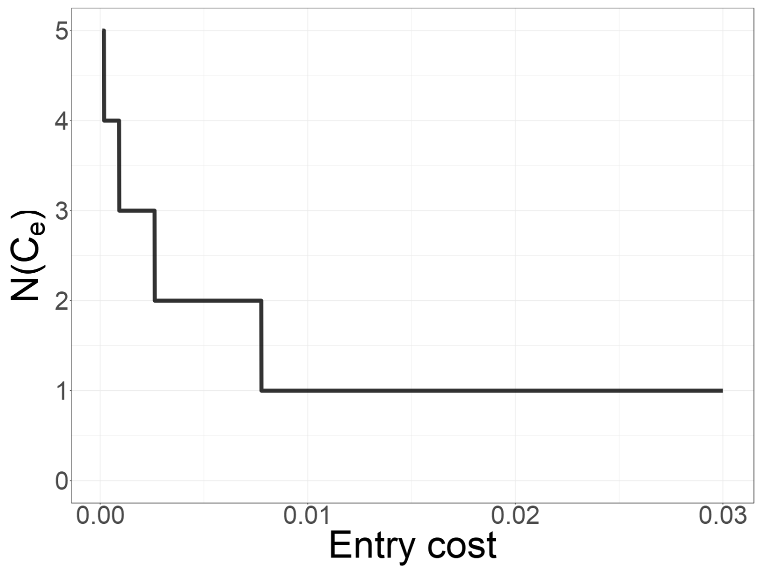
To convert delinquency rate changes into changes in the cost of providing credit, we multiply these changes by recovery rates, which are around 40% on auto loans on average.[28] This implies that increasing log HHIs by 1 decreases prime costs by −0.03% to 0.05%, and subprime costs by −0.18% to 1.00%.

Finally, we can use these to analyze markups. For each group, the high estimate of markups is the high effect on interest rates minus the low effect on costs, and vice versa for the low estimate. Using these, for the prime group, markups increase by roughly −0.12% to 2.01%; hence, most of our estimates suggest that markups increase for the high credit score group. On the other hand, for the low credit score group, we find that markups increase by roughly −1.78% to 0.55%. Thus, it is inconclusive whether markups increase. However, our estimates are consistent with markups staying constant or increasing slightly for the low credit score group, though we are unable to reject 0.

[28] See subprime recovery rates here.

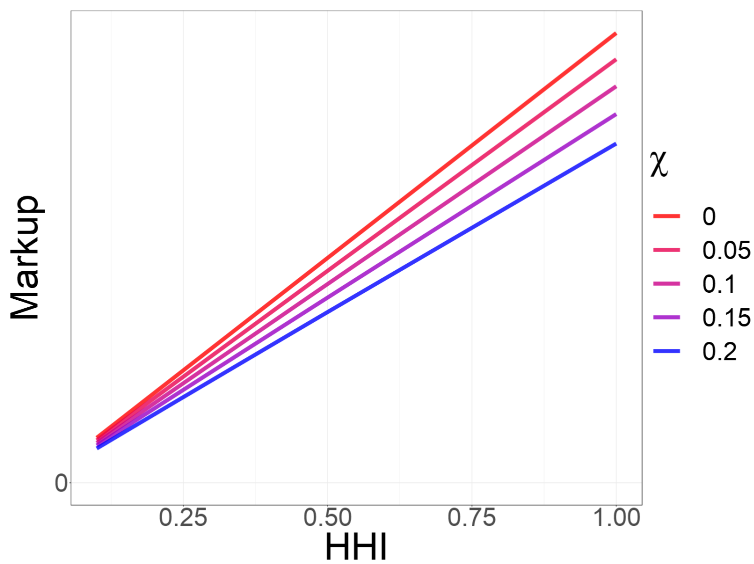
These calculations suggest that the magnitudes of the effects we estimate empirically are generally reasonable and consistent with each other. For prime consumers, increasing concentration is associated with a very small decrease in delinquency rates, and a sizable increase in interest rates in most specifications. We thus infer that markups are higher in concentrated markets for prime consumers, though a markup increase of 0 is within our confidence interval. For subprime consumers, our estimate for changes in markups contains zero. Thus, the size of the effects of concentration on delinquency rates is large enough to explain the drop in interest rates, while holding markups fixed or increasing them slightly.

Figure A.1: Equilibrium Entry



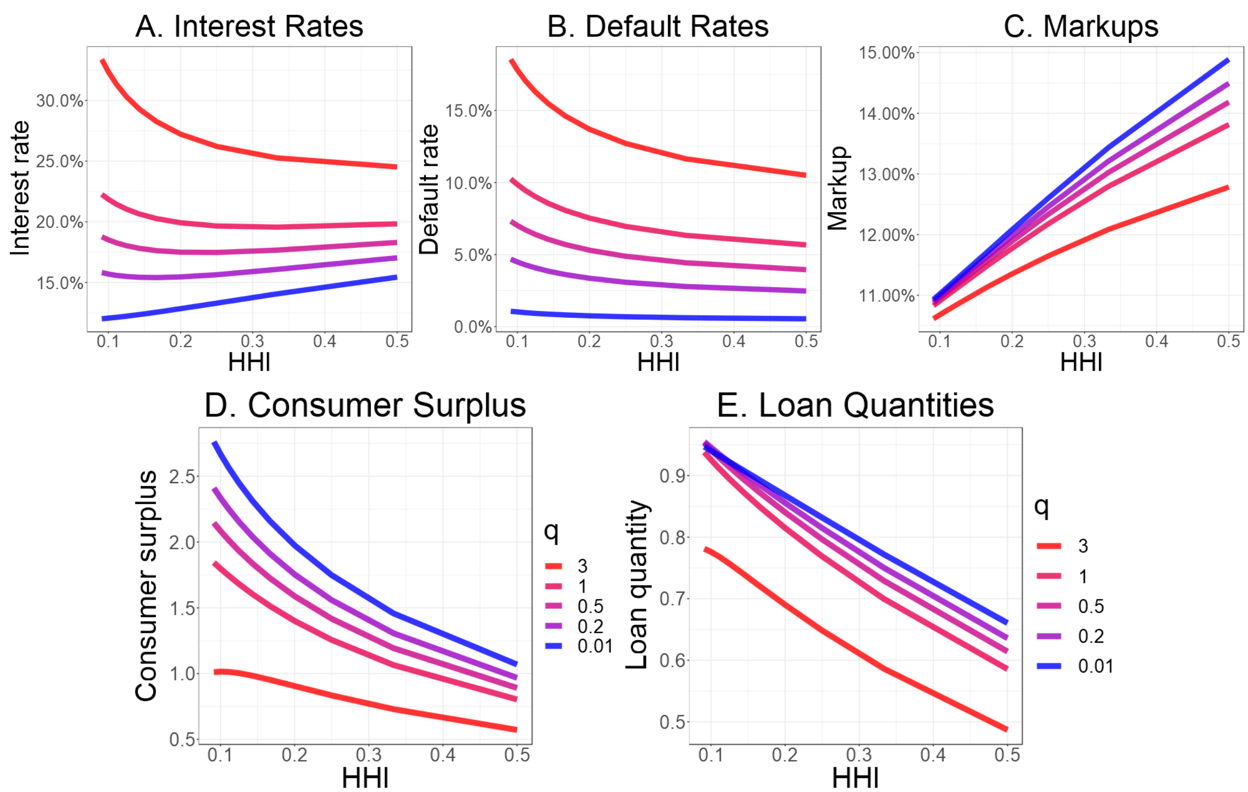
The above figure shows the equilibrium number of entrants as a function of the fixed entry cost C_e, in the entry model of Appendix A.6. The x-axis shows the fixed entry cost C_e. The y-axis shows the equilibrium number of entrants, $N(C_e)$, which we calculate by solving (27) and (28). We use the cost function (13), and we set $\theta = 0.04, \mu = 0.5, k = 0.001, q = 0.01$.

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 (34). We set $\delta = 0.1, \theta = 0.04$.

Figure A.3: Logit Model Simulations

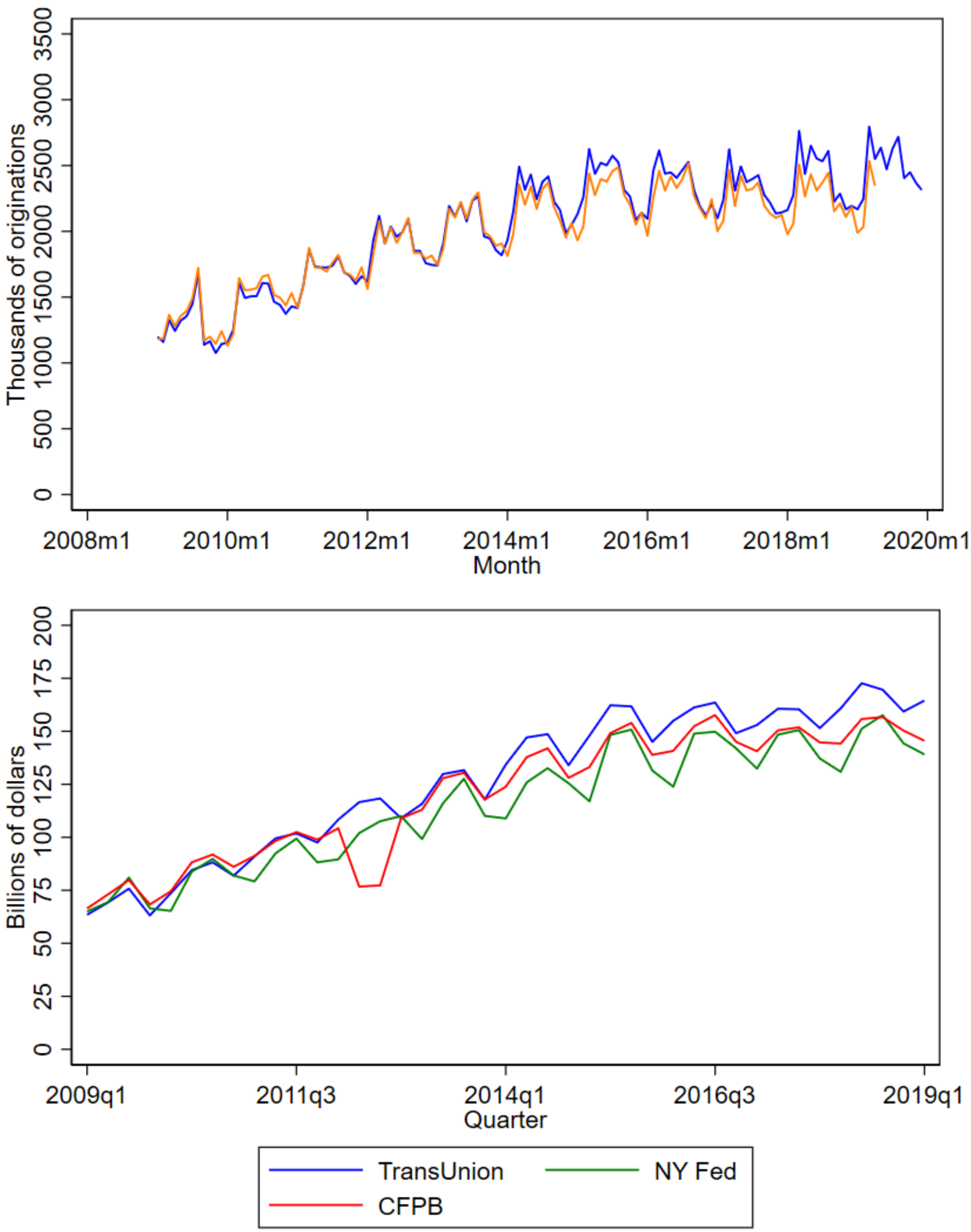


The above figure shows model outcomes, the Herfindahl-Hirschman index, HHI, and the measure of bad-type consumers, q, in the logit model described in Appendix A.9. In all plots, bluer lines represent lower values of q (that is, lower baseline default rates in the consumer population) and redder liners represent higher values of q. Panel A shows equilibrium interest rates, r. Panel B shows default rates δ. Panel C shows markups over the break-even interest rate, $r - \frac{\delta}{1-\delta}$. Panel D shows consumer surplus, (43). Panel E shows total loan quantities. All simulations use $\alpha = 10, \mu = 1.5, k = 0.001$.

Figure A.4: Welfare and Competition in the CES Demand Model

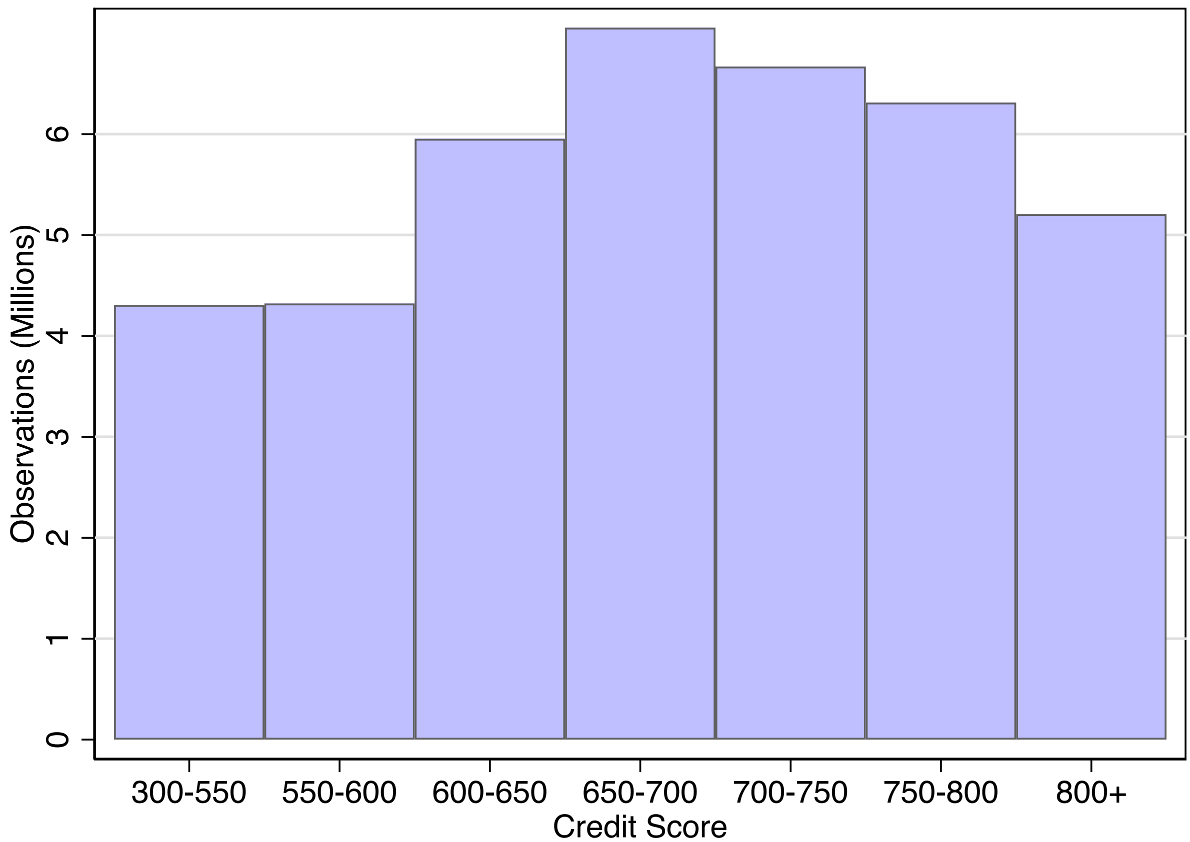
The above figure shows welfare in the CES demand system, broken down by ventiles of HHI, for different credit score groups. For each county-year, we compute welfare from the average interest rate and the HHI by calculating an implied number of lenders through (60), plugging average rates and the implied number of lenders into expression (59) for the price index R, and then plugging the price index into the welfare expression (55) of Claim 1. 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.5: Auto Loans in TU and Comparable Datasets



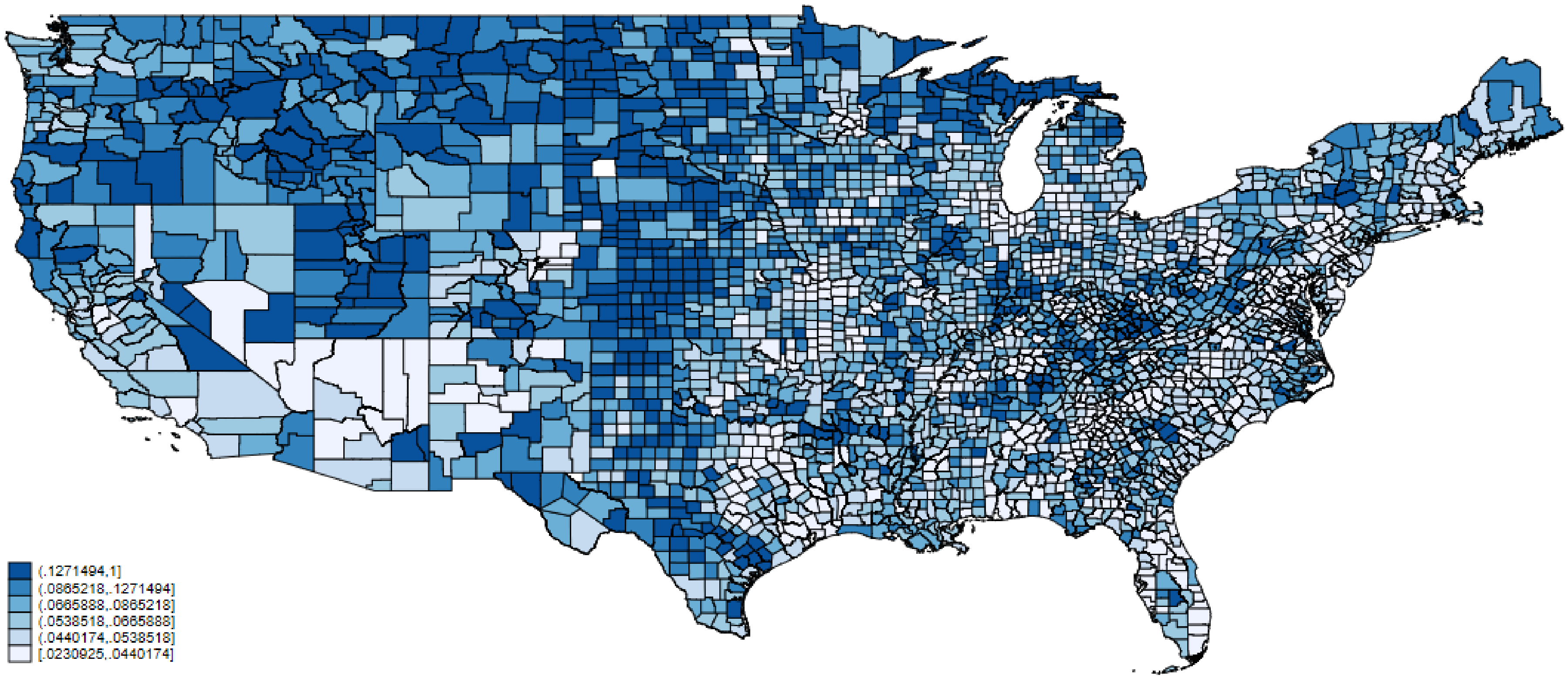
The above figure shows monthly auto loan originations. The top panel shows the number of loans, while the bottom panel shows the volume. The blue line shows the TU data series. In the top panel, the orange line shows a CFPB data series. In the bottom panel, the red line shows a CFPB data series and the green line shows an NY Fed data series. Source: TransUnion, CFPB, and NY Fed.

Figure A.6: Distribution of Credit Scores



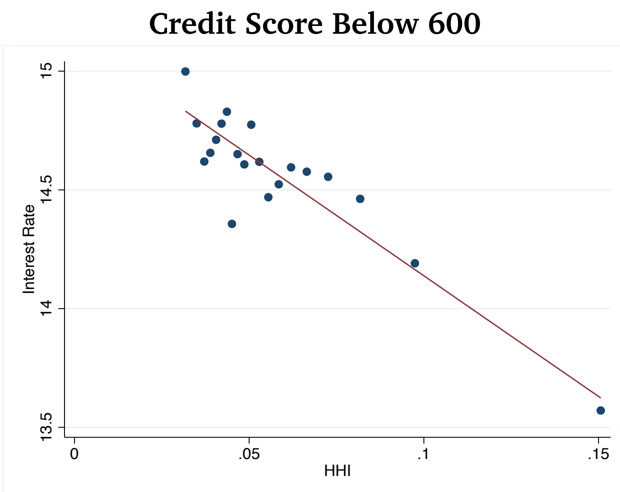
The above figure shows the total number of observations within each credit rating bin. Credit scores are given by VantageScore ratings. Source: TransUnion

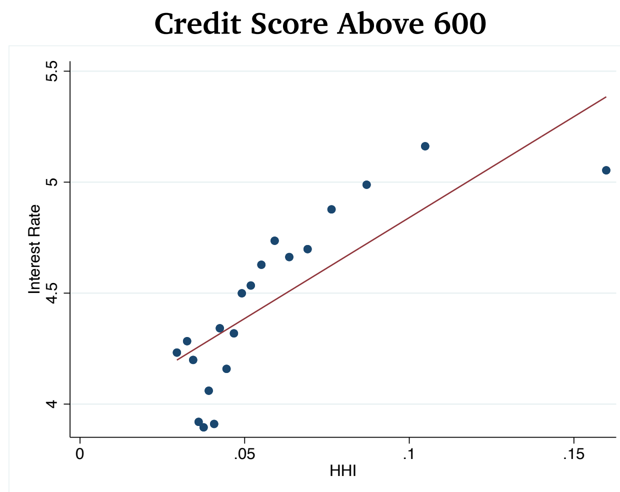
Figure A.7: Geographic Distribution of Market Competition



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.8: Interest Rates and Competition, Weighted

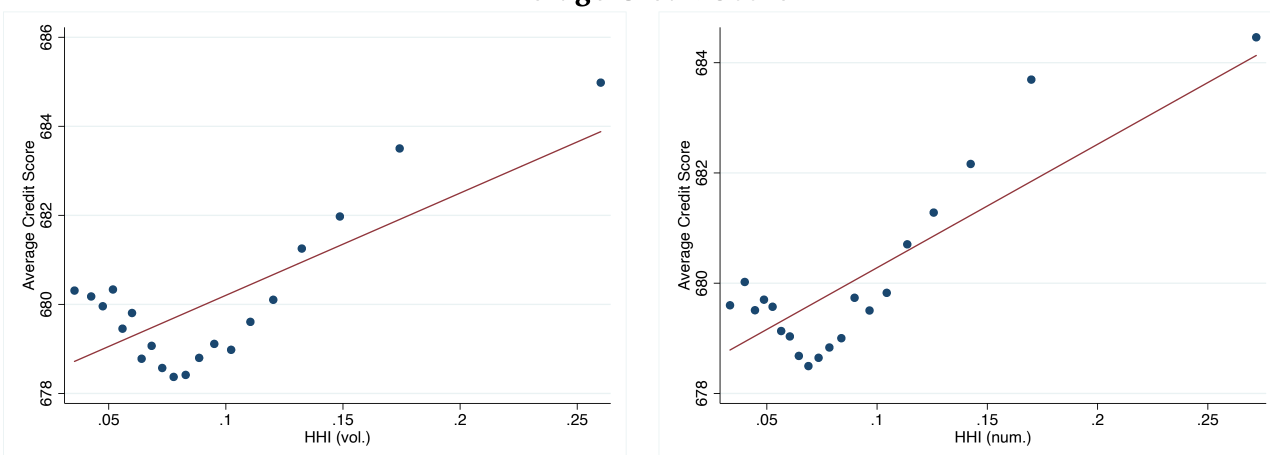




The above figure shows median interest rates in a county for given credit score ranges, broken down by ventile of HHI. HHI is defined using the volume of auto loans, that is $HHI = \sum_i^N s_i^2$, where s_i is a lender's share of auto lending in a county within a credit score range. Credit score ranges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

Figure A.9: Credit Scores and Competition

Average Credit Score

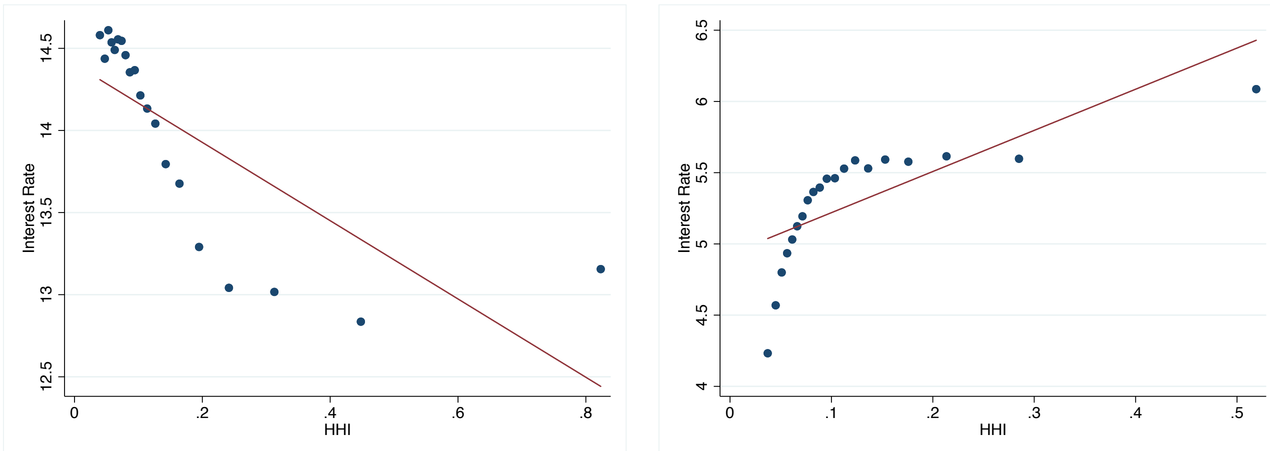


The above figure shows mean credit scores, broken down by ventile of HHI. In the left panels, HHI is defined using the volume of auto loans, and in the right panels HHI is defined using the number of auto loans; that is, $HHI = \sum_i^N s_i^2$, where s_i is a lender's share of auto lending in a county 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.

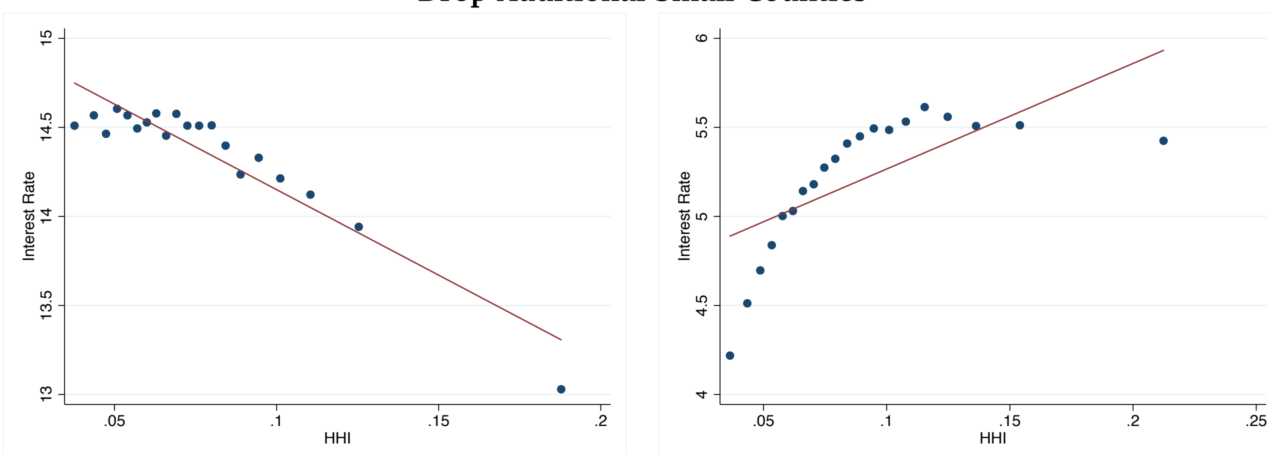
Figure A.10: Interest Rates and Competition: Alternative Measurement

Credit Score Below 600 **Credit Score Above 600**

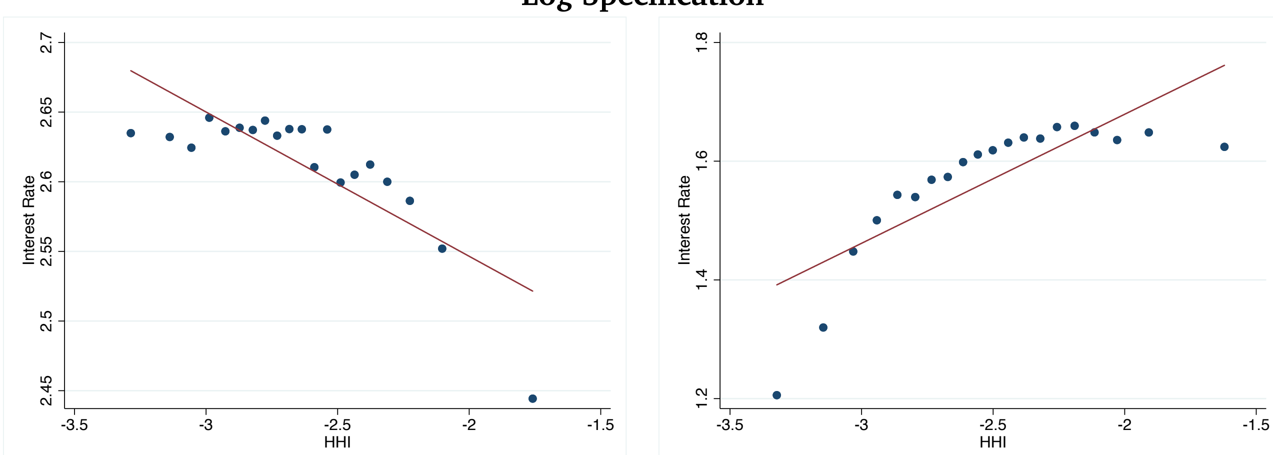
No Small Counties Dropped



Drop Additional Small Counties



Log Specification

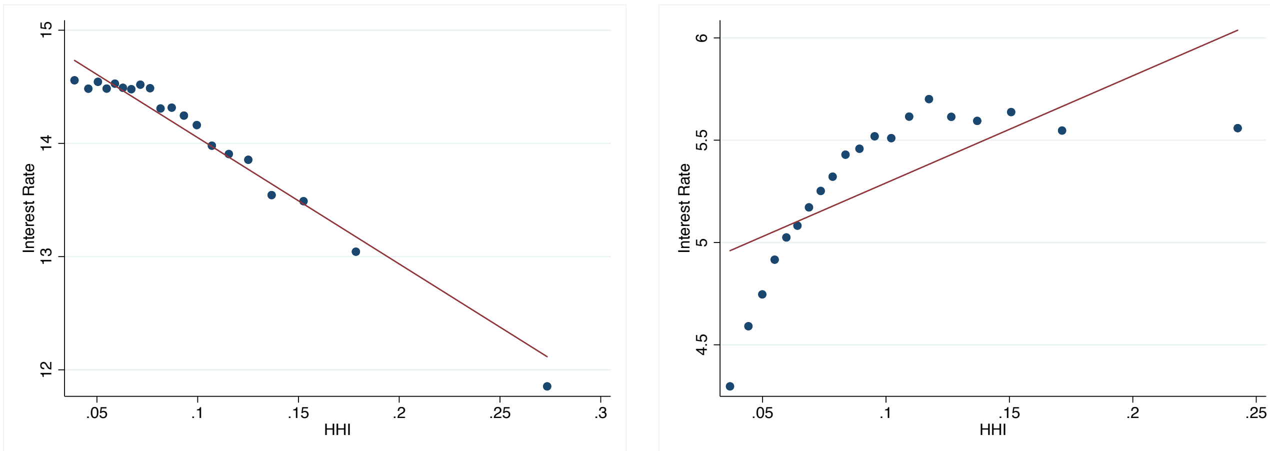


The above figure shows median interest rates in a county for given credit score ranges, broken down by ventile of HHI using alternative samples or measurement. The measurement or sample is noted above each panel. HHI is defined using the volume of auto loans, that is $HHI = \sum_i^N s_i^2$, where s_i 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.

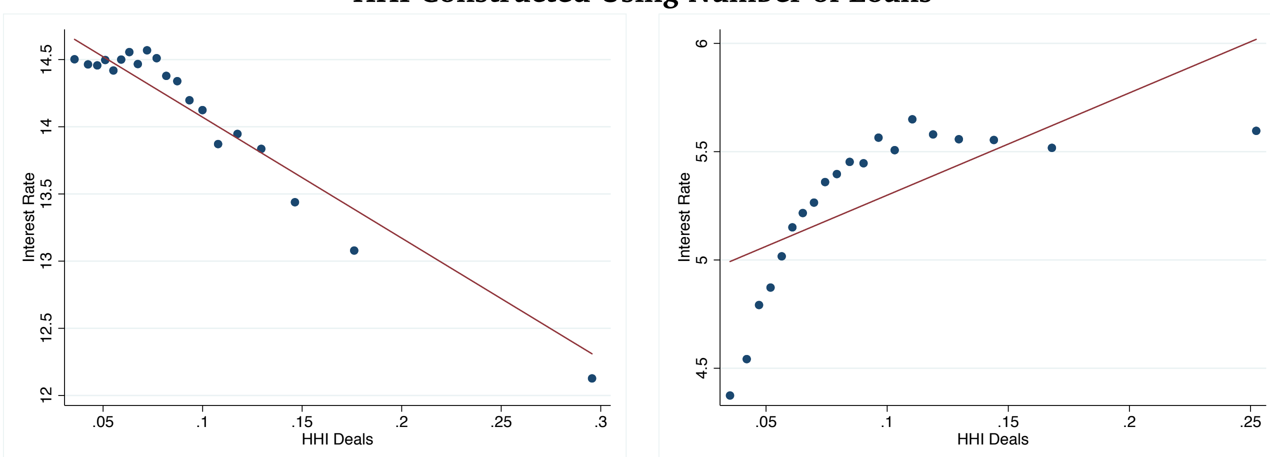
Figure A.11: Interest Rates and Competition: Alternative Measurement

Credit Score Below 600 | **Credit Score Above 600**

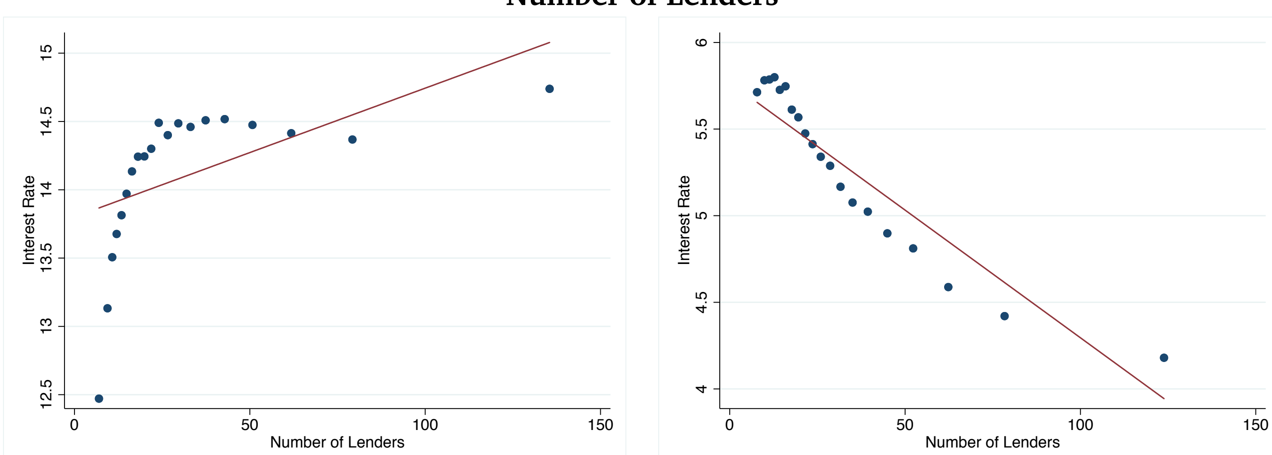
Winsorized



HHI Constructed Using Number of Loans

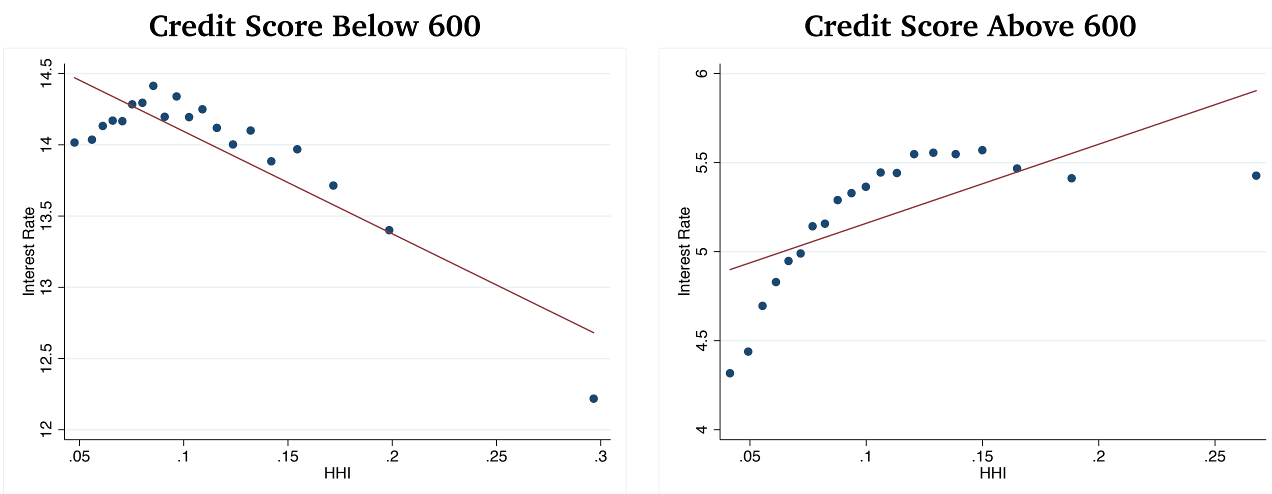


Number of Lenders



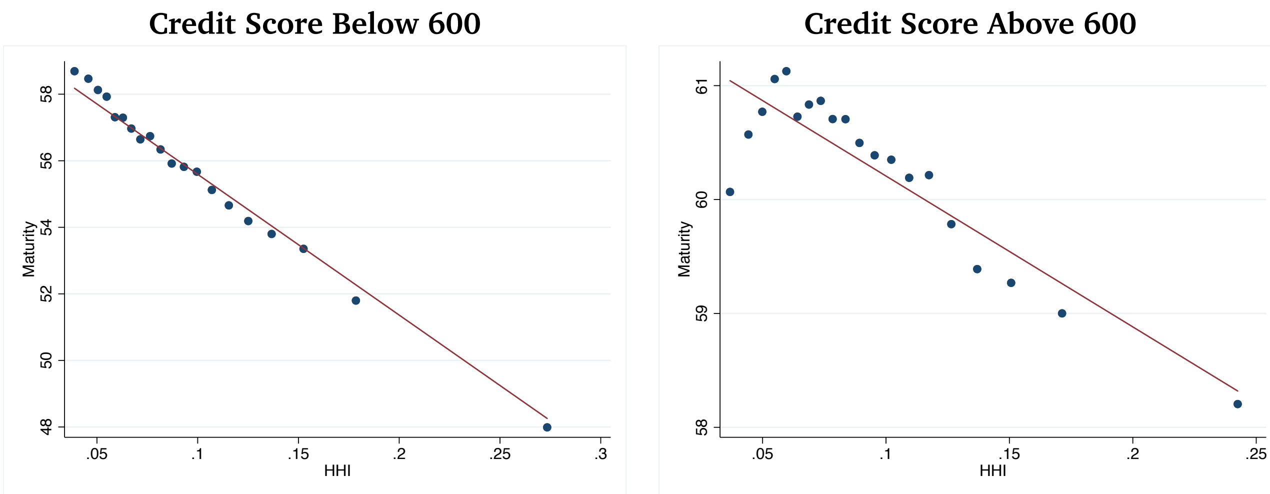
The above figure shows median interest rates in a county for given credit score ranges, broken down by ventile of HHI using alternative measurement. The measurement change is noted above each panel. For the top panel, HHI is defined using the volume of auto loans, that is $HHI = \sum_i^N s_i^2$, where 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.12: Interest Rates and Competition Excluding New Entrants



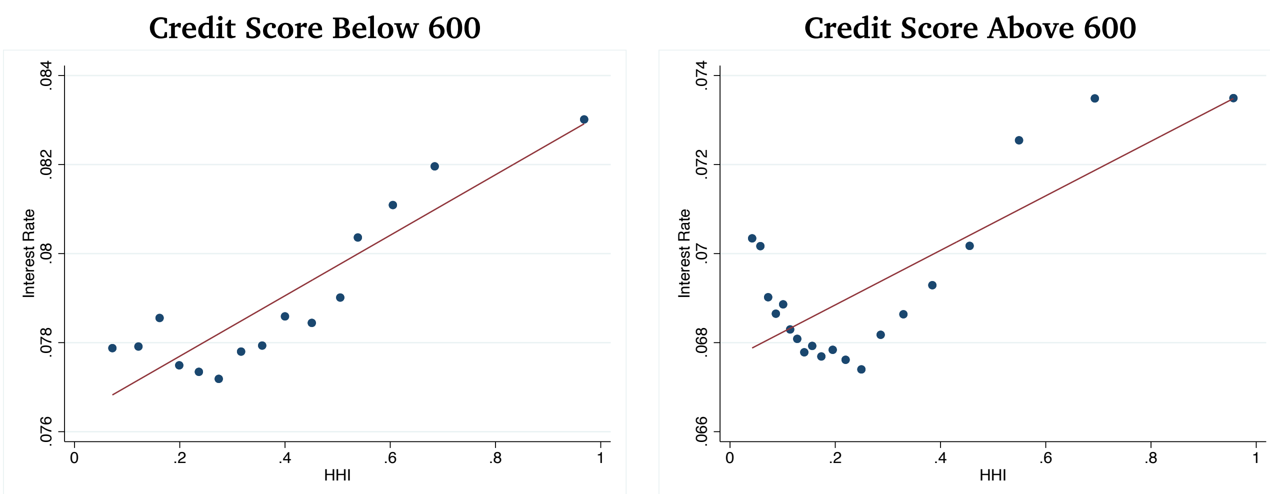
The above figure shows median interest rates in a county for given credit score ranges, broken down by ventile of HHI. The sample is restricted to lenders that operated since 2009, and thus excludes new entrants. HHI is defined using the volume of auto loans, that is $HHI = \sum_i^N s_i^2$, where s_i 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.13: 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.14: Interest Rates and Competition in the Mortgage Market



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 interest rates and HHI by credit score, at the lender by year level. Each observation is a county by year by lender by credit score bucket. Columns (1)-(4) regress logged outcomes on interactions of logged HHI with indicators of credit score buckets, varying the inclusion of lender and county fixed effects. Columns (5)-(10) split the sample by whether a lender offers only auto loans, or other types of consumer credit and regresses logged interest rates on interactions of logged HHI with indicators of credit score buckets. The inclusion of fixed effects is denoted beneath each column. Regressions are weighted by the number of loans. Note that interest rates are determined for each county, lender, and year, within each credit score bucket. 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)	(9)	(10)
	Ln(Interest Rate)									
	Full Sample				Auto Lenders			General Lenders		
Ln(HHI) ×	-0.245***	-0.223*	-0.243***	-0.227**	-0.103***	-0.210***	-0.139	-0.185***	-0.273***	-0.339***
Credit Score 300-550	(0.0158)	(0.116)	(0.0146)	(0.109)	(0.0130)	(0.0154)	(0.151)	(0.0163)	(0.0150)	(0.0373)
Ln(HHI) ×	-0.0583***	-0.221***	-0.0946***	-0.214***	0.0546**	-0.0372	-0.163*	-0.0331***	-0.103***	-0.209***
Credit Score 550-600	(0.0206)	(0.0593)	(0.0212)	(0.0591)	(0.0232)	(0.0247)	(0.0969)	(0.0117)	(0.0114)	(0.0197)
Ln(HHI) ×	0.0801***	-0.109*	0.0157	-0.108*	0.145***	0.0650**	-0.0789	0.0844***	0.0272***	-0.0751***
Credit Score 600-650	(0.0284)	(0.0567)	(0.0219)	(0.0566)	(0.0314)	(0.0323)	(0.0988)	(0.00999)	(0.00859)	(0.0195)
Ln(HHI) ×	0.115**	-0.0829**	0.0316	-0.0856**	0.127*	0.0529	-0.0862	0.0879***	0.0414***	-0.0541***
Credit Score 650-700	(0.0524)	(0.0403)	(0.0390)	(0.0405)	(0.0740)	(0.0767)	(0.0653)	(0.0126)	(0.0117)	(0.0161)
Ln(HHI) ×	0.199***	-0.0230	0.101*	-0.0295	0.210*	0.130	-0.0297	0.112***	0.0701***	-0.0272
Credit Score 700-750	(0.0769)	(0.0368)	(0.0589)	(0.0366)	(0.113)	(0.115)	(0.0559)	(0.0189)	(0.0179)	(0.0165)
Ln(HHI) ×	0.309***	0.0837**	0.199**	0.0730*	0.304**	0.209	0.0878	0.167***	0.123***	0.0212
Credit Score 750+	(0.101)	(0.0425)	(0.0784)	(0.0410)	(0.137)	(0.140)	(0.0618)	(0.0236)	(0.0224)	(0.0177)
Lender Fixed Effects	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes
County Fixed Effects	No	No	Yes	Yes	No	No	Yes	No	No	Yes
Score Bucket Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,080,050	7,080,050	7,080,050	7,080,050	3,509,639	3,509,639	3,509,639	3,570,411	3,570,411	3,570,411
R^2	0.302	0.687	0.362	0.692	0.320	0.333	0.695	0.471	0.507	0.695

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. Each observation is a county by year by lender by credit score bucket. 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: 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 or below 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

Table A.5: Interest Rates and Market Competition Over Time

The table shows the relationship between interest rates and HHI by credit score, by year. Each column splits the analysis sample into finer county by year by credit score bins, and interacts logged HHI with indicators of credit score buckets. The inclusion of fixed effects is denoted beneath each column. Regressions are weighted by the number of loans. Note that interest rates are determined for each county, within each credit score bucket. 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)	(9)	(10)	(11)
	Ln(Interest Rate)										
	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Ln(HHI) ×	-0.291***	-0.330***	-0.289***	-0.204***	-0.101***	-0.0898***	-0.0642***	-0.0314	-0.0391*	-0.0291	-0.0175
Credit Score 300-550	(0.0321)	(0.0333)	(0.0234)	(0.0166)	(0.0119)	(0.0125)	(0.0157)	(0.0221)	(0.0211)	(0.0196)	(0.0183)
Ln(HHI) ×	-0.114***	-0.0609***	-0.0430**	0.00435	0.0184	0.0227	0.0556**	0.0490*	0.0784***	0.0318	0.00850
Credit Score 550-600	(0.0174)	(0.0181)	(0.0181)	(0.0200)	(0.0203)	(0.0224)	(0.0253)	(0.0262)	(0.0285)	(0.0246)	(0.0209)
Ln(HHI) ×	-0.00647	0.0721***	0.0881***	0.110***	0.123***	0.0960***	0.135***	0.127***	0.138**	0.130***	0.0927***
Credit Score 600-650	(0.0101)	(0.0152)	(0.0165)	(0.0154)	(0.0249)	(0.0210)	(0.0302)	(0.0361)	(0.0553)	(0.0483)	(0.0335)
Ln(HHI) ×	0.0265***	0.124***	0.112***	0.104***	0.156***	0.124***	0.140**	0.0794	0.123*	0.0742	0.103**
Credit Score 650-700	(0.00960)	(0.0223)	(0.0203)	(0.0218)	(0.0398)	(0.0477)	(0.0630)	(0.0652)	(0.0657)	(0.0740)	(0.0455)
Ln(HHI) ×	0.0373***	0.231***	0.173***	0.183***	0.235***	0.178**	0.195***	0.165**	0.217***	0.0459	0.151***
Credit Score 700-750	(0.00949)	(0.0637)	(0.0384)	(0.0495)	(0.0610)	(0.0775)	(0.0737)	(0.0795)	(0.0763)	(0.0956)	(0.0564)
Ln(HHI) ×	0.0918***	0.379***	0.362***	0.255**	0.312**	0.224	0.305**	0.227*	0.313**	0.102	0.272***
Credit Score 750+	(0.0227)	(0.0932)	(0.112)	(0.101)	(0.141)	(0.139)	(0.123)	(0.127)	(0.127)	(0.140)	(0.0765)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Score Bucket Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14,002	14,698	15,249	15,732	15,903	16,094	16,194	16,141	16,048	16,238	16,129
R^2	0.645	0.394	0.518	0.581	0.519	0.458	0.440	0.393	0.346	0.325	0.358

Table A.6: Interest Rates Delinquency, Loans and Market Competition Controlling for Markups

The table shows the relationship between interest rates, delinquency or the number of loans and HHI, split by credit score, controlling for loan markups at the state level. Columns (1) shows interest rates, while column (2) shows delinquency rates, and columns (3) shows quantities. Each group of columns splits the sample into county by year by credit score bins, and regresses logged outcomes on interactions of logged HHI with indicators of credit score buckets. The inclusion of fixed effects is denoted beneath each column. Regressions are weighted by the number of loans. Note that interest rates are determined for each county, within each credit score bucket. Standard errors are clustered at the county level. Credit scores are given by VantageScore ratings. Source: TransUnion and Jansen et al. (2021). $^{*}p < .1$, ** $p < .05$, *** $p < .01$.

	(1)	(2)	(3)
	Ln(Interest Rate)	Ln(Delinquency)	Ln(Loans)
Ln(HHI) ×	-0.0146	-0.143***	-0.264***
Credit Score 300-550	(0.122)	(0.0137)	(0.0129)
Ln(HHI) ×	-0.0945**	-0.210***	-0.212***
Credit Score 550-600	(0.0456)	(0.0136)	(0.00981)
Ln(HHI) ×	-0.0195	-0.197***	-0.207***
Credit Score 600-650	(0.0361)	(0.0118)	(0.00885)
Ln(HHI) ×	-0.0213	-0.141***	-0.223***
Credit Score 650-700	(0.0267)	(0.0105)	(0.0105)
Ln(HHI) ×	0.0144	-0.0665***	-0.226***
Credit Score 700-750	(0.0364)	(0.00990)	(0.0120)
Ln(HHI) ×	0.105*	-0.0284***	-0.426***
Credit Score 750+	(0.0586)	(0.0102)	(0.0177)
County Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Score Bucket Fixed Effects	Yes	Yes	Yes
Observations	131,977	131,977	131,977
R^2	0.792	0.693	0.957