Market Power in Small Business Lending: A Two Dimensional Bunching Approach[*]

Natalie Bachas[†] and Ernest Liu[‡]

August 30, 2019

PRELIMINARY AND INCOMPLETE
NOT FOR CIRCULATION OR CITATION

Abstract

While bank lending is an important financing channel for small firms, banks in the U.S. have substantial market power. What are the efficiency implications and policy remedies to bank concentration? We build a model of bank competition with endogenous interest rates, loan size, and take-up. We estimate the model using the universe of loans made through the Small Business Administration (SBA). Our novel identification strategy builds on and extends the "bunching" literature that uses kinks and notches to identify key elasticities, utilizing a discontinuity in SBA's interest rate cap. We find banks capture at least 30% of the surplus in a majority of lending markets. Imposing a uniform interest rate cap of 5% would increase borrower welfare by 9%, but also cause substantial rationing. While the guarantee subsidy program used by the SBA raises borrower surplus by 17%, we find that banks capture the majority of increase in surplus.

[*] We thank Vivek Bhattacharya, Olivier Darmouni, Yiming Ma, David Sraer, Adam Szeidl, Kairong Xiao, Constantine Yannelis, Anthony Zhang, and seminar and conference participants at UNC, Princeton, and Columbia for helpful comments and discussions. We are grateful to Brian Headd at the SBA for helpful discussions and comments about institutional details. Christian Kontz and Daniel Morrison provided excellent research assistance. This draft is preliminary, please do not cite or circulate without the permission of the authors. Comments are welcome.

[†] Princeton University. Email: nbachas@princeton.edu.

[‡] Princeton University. Email: ernestliu@princeton.edu.

1 Introduction

Bank lending is an important financing channel for young and small firms and is therefore critically important for the aggregate economy (Kaplan and Zingales (1997)). Yet, reliance on geographic proximity between borrowers and lenders (Petersen and Rajan (1994)) can give banks substantial market power and potentially cause under-provision of credit (Dreschler, Savov, and Schnabl (2017)). Several federal programs in the United States exist to regulate pricing and encourage bank lending to small businesses. How does market power affects the terms of bank lending? How to estimate market power in lending? What are the effects of the existing regulations, and is there room for better policy? Despite broad academic and policy interests, these remain open questions.

In this paper, we build and estimate a model of imperfect competition in bank lending with endogenous interest rates, loan size, and take-up. In the model, a finite number of banks compete for borrowers by offering loan contracts. Each contract specifies both the interest rate and the loan size. Banks are differentially preferred by borrowers with idiosyncratic taste shocks over banking services. Taste heterogeneity, together with the finiteness of competing banks, grants banks market power.

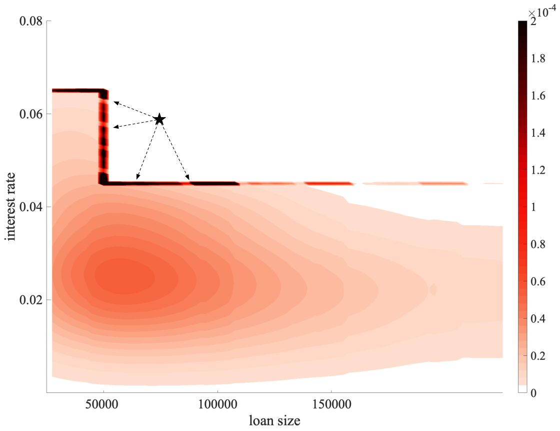
The model generates a mapping from bank concentration to lending outcomes and clarifies the implications of bank market power on both the intensive and extensive margins of bank lending. Despite market power, the intensive margin of loan size is always efficient conditioning on loan issuance, as banks choose the optimal loan size to maximize joint bank-firm surplus and only use interest rates to optimally extract surplus. However, market power distorts the extensive margin of lending, as high interest rates in concentrated local markets discourages firms from taking out loans. Our model is highly tractable: it yields analytic solutions and is amenable to theoretical normative policy analysis.

We estimate the model using the universe of small business loans made through a major federal loan subsidy program in the United States—the Small Business Administration (SBA) Express program. The SBA guarantees loans made by commercial lenders to in-need small businesses that are otherwise rejected from all other sources of external financing. It therefore relies on the existing banking infrastructure to pass the subsidy through to targeted firms.

Using a novel identification strategy that combines geographic variations in bank concentration and discontinuities in the regulatory specifications of the SBA program, we quantify the impact of market power on small business lending. We find quantitatively substantial market power and inefficiencies: on average, banks capture 20-30% of surplus from lending relationships. We estimate the efficacy of loan subsidy through the SBA program and perform a wide range of policy counterfactuals.

Our identification strategy builds on and extends the "bunching" literature that uses kinks and notches to identify key elasticities (Kleven (2016)). Broadly speaking, this approach uses discontinuities in economic agents' choice set and the consequent distortions in the equilibrium outcome distribution to infer structural parameters that govern economic behaviors. To the best of our knowledge, existing papers study one-dimensional bunching, meaning they study environments with distortions in a single choice variable.[1]

Figure 1: Loan-Size-Depedent Interest Rate Cap and the Empirical Distribution of Contracts

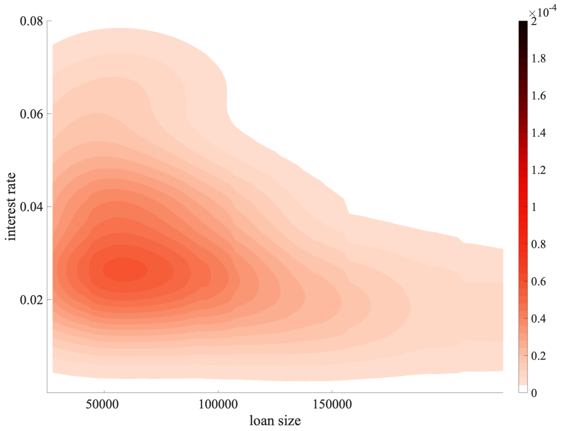


Our methodological contribution is to advance the bunching approach to multi-dimensional behavioral response. In our setting, loans made through the SBA are subject to a loan-size dependent interest rate cap as shown in Figure 1: loans smaller or equal to $50,000 are capped at a rate of 6.5% above a reference rate (the "Prime rate"), while loans larger than $50,000 are limited to Prime rate + 4.5%. Banks compete on two-dimensional loan contracts—loan size and interest rates; hence, banks respond in both dimensions to the interest cap. Consider in Figure 1 a borrower who, in the absence of the rate cap, would have been offered contract ($\star$), which is infeasible under the rate cap. When the rate cap is imposed, multiple contracts along the rate cap could have possibly been offered to the borrower. Banks could lower the interest rate and stay unconstrained on loan size; alternatively, banks can scale back loan size and, in exchange, charge a relatively higher interest rate. These contracts generate different profits for the banks, as loan size affects the surplus generated by each loan, whereas the interest rate determines the division of that surplus between the lender and the borrower.

[1] For example Best and Kleven (2018) study how the UK mortgage market responds to transaction taxes, and study how labor supply responds to tax notches in Pakistan (Kleven and Waseem (2013)) and tax rate kinks in the US Saez (2010).

How a profit-maximizing bank respond to the size-dependent interest rate cap is indicative of their market power and the elasticity of lending surplus to loan size. For contract ($\star$), reducing the interest rate to the lower cap (Prime rate + 4.5%) significantly lowers the revenue-per-dollar-lent that banks are able to extract and may have relatively small impact on the surplus of lending because loan size remain unconstrained. In contrast, while reducing loan size down to $50,000 allows higher interest rates, doing so could significantly reduce the surplus of lending. We argue that a bank with no market power will always choose to scale back loan size. This is because the interest rate in ($\star$) fully reflects lending cost, and interest rates that are any lower can only lead to losses for the bank. The argument extends to non-competitive cases, that a bank with more market power is more likely to lower rates and avoid the constraint on loan size.

Figure 2: Counterfactual Distribution of Contracts



Our two-dimensional bunching approach operationalizes the argument above to identify bank market power. First, we start with the empirical joint distribution of loan size and interest rate, as shown in Figure 1. Second, we use a statistical procedure to recover the counterfactual distribution of contracts (i.e. the distribution that would have prevailed absent the rate cap), as shown in Figure 2. To do so, we assume that the distribution of contracts strictly below the policy cap is unaffected by the policy, and we extrapolate the distribution above the cap using the empirical distribution below. We compute the difference between the empirical and the counterfactual distributions of contracts, and we repeat the procedure across markets. Lastly, based on the argument above, we form moment conditions that identify two key model parameters, which respectively govern 1) the elasticity of bank's market power to market concentration and 2) the elasticity of lending surplus to loan size.

We begin with an exposition of the model in section 2. Section 3 discusses the empirical setting, data, and relevant policy variation, while section 4 discusses the identification strategy. Section 5 describes the estimation procedure and empirical findings. Section 6 concludes.

2 Model

2.1 Setup

Consider a market with finite K banks and a continuum of borrowers of finite measure. Both parties are risk neutral. Let k index for banks and i index for borrowers.

Investment Technology Each borrower i has a stochastic investment opportunity that produces output $f(L;\omega)$ as a function of investment size L in state ω. The state realization $\omega \in \mathscr{W}$ is idiosyncratic across borrowers and is drawn from a borrower-specific distribution $G_i(\omega;L)$ which may also depend on investment L. The expected output from investment i is

$$\mathbb{E}_i[f(L)] \equiv \int_{\omega\in\mathscr{W}} f(L;\omega) \;\; dG_i(\omega;L).$$

Loan Contract Borrowers may obtain investments from bank loans. A loan contract is a duplet of interest rate and loan size, (r,L). If the contract (r,L) offered by bank k is accepted by borrower i, it generates contractual value $v_i(r,L)$ to borrower i and expected profit $\pi_{ik}(r,L)$ for bank k:

$$v_i(r,L) \equiv \int_{\omega\in\mathscr{W}} \max\{f(L;\omega)-(1+r)L,0\} \;\; dG_i(\omega;L), \tag{1}$$

$$\pi_{ik}(r,L) \equiv \int_{\omega\in\mathscr{W}} \min\{(1+r)L,\; f(L;\omega)\} \;\; dG_i(\omega;L) \;\; - \;\; c_k L. \tag{2}$$

Note that loan contracts are equivalent to debt: the lender captures the investment payoff upto the specified repayment $(1+r)L$, and the borrower is the residual claimant. The term c_k represents the opportunity cost of funds to bank k.

The expected utility that borrower i obtains from picking contract (r,L) from bank k is

$$u_{ik}(r,L) \equiv \xi_{ik} \times v_i(r,L). \tag{3}$$

The term $\xi_{ik} \geq 0$ is a random taste shock and is i.i.d. across borrowers and banks. We refer to $v_i(r,L)$ as the *contractual* value, and $u_{ik}(r,L)$ as the *expected utility,* of loan (r,L) to borrower i. The taste shock ξ_{ik} represents idiosyncratic heterogeneity, such as borrowers' differential preferences for the services provided by differentiated banks.

When the context is clear, we abuse notations and use u_{ik} for $u_{ik}(r_{ik},L_{ik})$, v_{ik} for $v_i(r_{ik},L_{ik})$, and π_{ik} for $\pi_{ik}(r_{ik},L_{ik})$.

Bank Competition and Equilibrium Banks $k = 1, \ldots, K$ compete for borrowers by simultaneously offering contracts. Each bank k offers one contract (r_{ik}, L_{ik}) to each borrower i. Borrowers accepts the contract that generates the highest expected utility. The probability that borrower i chooses the contract offered by bank k is

$$q_{ik} \equiv Pr\left(i \text{ chooses } k\right) = Pr\left(u_{ik} \geq u_{ik'} \text{ for all } k'\right), \tag{4}$$

The randomness in borrower's choice of contract originates from the idiosyncratic taste shocks. Note that q_{ik} is increasing in the contractual utility v_{ik} offered by bank k and is decreasing in $v_{ik'}$ for all $k' \neq k$. We normalize

When competing for borrowers, banks observe borrower's production technology $G_i\left(\omega; L\right)$ but do not observe the idiosyncratic shocks. Each bank k offers the contract that maximizes expected profit:

$$(r^*_{ik}, L^*_{ik}) \equiv \arg\max_{r_{ik}, L_{ik}} q_{ik} \times \pi_{ik}. \tag{5}$$

Definition 1. A Laissez-faire equilibrium is the set of contracts $\left\{\left(r^*_{ik}, L^*_{ik}\right)\right\}_{k=1}^{K}$ that solves the profit maximization problem (5).

Default happens in the model when the output is below the required loan repayment ($f\left(L; \omega\right) < \left(1+r\right)L$). Note that default is always involuntary in the model: borrower repays as much as the output allows, and there is no strategic decision regarding default. Another feature of the model is that default generates no deadweight loss and simply represents a transfer between the borrower and the lender under the contingency that output is low. This can be seen by noting that the sum of bank profit and the contractual value to the borrower is a function of only loan size and is invariant to the interest rate:

$$v_i\left(r, L\right) + \pi_{ik}\left(r, L\right) = \mathbb{E}_i\left[f\left(L\right)\right] - c_k L.$$

We model this feature in order to abstract away from inefficient default; the only source of potential inefficiency in the model is market power.

2.2 The Laissez-faire Equilibrium

Let $\varepsilon_{ik} \equiv \partial \ln q_{ik} \big/ \partial \ln v_{ik}$ denote the elasticity of the choice probability q_{ik} (that borrower i chooses bank k) with respect to the contractual utility v_{ik}, holding contracts offered by all other banks constant. We refer to ε_{ik} simply as the "choice elasticity". The choice elasticity is always non-negative, as higher contractual utility v_{ik} always raises the likelihood for borrower i to accept the contract offered by bank k.

Proposition 1. *In the Laissez-faire equilibrium, loan terms $\left(r_{ik}^{*}, L_{ik}^{*}\right)$ satisfy*

$$L_{ik}^{*} = \arg\max_{L} \left\{\mathbb{E}_{i}\left[f\left(L\right)\right] - c_{k}L\right\}, \tag{6}$$

$$\frac{\pi_{ik}\left(r_{ik}^{*}, L_{ik}^{*}\right)}{v_{i}\left(r_{ik}^{*}, L_{ik}^{*}\right)} = \frac{1}{\varepsilon_{ik}}. \tag{7}$$

The proposition characterizes equilibrium loan terms offered by each bank as a function of lending cost c_k. Equation (6) shows that equilibrium loan size is efficient—it maximizes the expected investment output net of lending cost. This result may come as a surprise: banks do have market power and are profit-maximizing entities that choose contractual terms. Why is there is no distortion over loan terms? To understand this, note that each bank always offers the loan size that maximizes output net of lending cost and default costs; the bank can always extract rents by charging a high interest rate. Because default does not generate deadweight losses, the equilibrium loan size is always efficient.

The equilibrium interest rate is implicitly characterized by equation (7), which states that the ratio between bank profits and the contractual value captured by the borrower is inversely related to the choice elasticity ε_{ik}. When ε_{ik} is high, the choice probability q_{ik} is sensitive to contractual value v_{ik}; consequently, banks choose low interest rates, extract little surplus, and leave more surplus to the borrower (low π_{ik} and high v_{ik}). Conversely, an inelastic choice probability implies high interest rates and bank profits, and low contractual value to the borrower. In what follows, we say bank k has low market power to borrower i if ε_{ik} is high.

Proposition 1 clarifies the potential impact of market power in lending. Absent inefficient default, market power translates into high interest rates and efficient loan size offered; consequently, market power does not generate investment distortions along the intensive margin, i.e., investments are efficient as long as investments are made. However, because high market power translates into high interest rates and low contractual value, market power may lead to under-investment along the extensive margin, as entrepreneurs may choose not to borrow at all. This can be formalized by allowing the borrower to choose an outside option in our model.

In order to take our model to data, we now turn to a parametrized version of our model with functional form assumptions on the investment technology.

2.3 Parametrized Model

We parametrize the model so that it is sufficiently rich to be taken to the data, yet it remains simple and tractable. We maintain these parametric assumptions throughout the rest of the paper.

Investment Technology We assume the investment technology is isoelastic and, for a given investment L, has two random output levels:

$$f(L) = \begin{cases} z_i L^{\alpha} + \delta_i L & \text{(succeeds) with probability } p_i, \\ \delta_i L & \text{(fails) with probability } (1 - p_i). \end{cases}$$

Each borrower i can be summarized by its characteristic (z_i, δ_i, p_i). The borrower succeeds with probability p_i. The term $\delta_i L$ can be seen as the undepreciated investment after output is realized; this is also the amount that can be recovered by banks in case of investment failure. If the investment succeeds, it generates an additional output $z_i L^{\alpha}$. The term z_i is a Hicks-neutral productivity shifter, and the parameter α captures the concavity of the production function.

When the project fails, borrower gets paid zero and the bank gets paid $\delta_i L$. When the project succeeds, borrower gets paid $z_i L^{\alpha} - (1 + r - \delta_i) L$ and the bank gets paid $(1 + r) L$.

Let $F(\cdot)$ denote the distribution of borrower characteristics in the market.

Distribution of Taste Shock We assume the idiosyncratic taste shocks ξ_{ik} are drawn from a Fréchet distribution, with CDF $G(\xi; \sigma) = e^{-(\gamma\xi)^{-\sigma}}$, where $\gamma \equiv \Gamma(1 - 1/\sigma)$ is a normalizing constant and Γ is the Gamma function (Johnson and Kotz (1970)). This assumption enables us to analytically solve for equilibrium loan contracts as a function of the market structure; it has the implication that, ceteris paribus, the choice elasticity ε_{ik} is higher in more competitive markets.

Under the distributional assumption, the choice probability for any given bank becomes

$$q_{ik}\left(\{v_{ik'}\}_{k'=1}^{K}\right) = \frac{v_{ik}^{\sigma}}{\sum_{k'=1}^{K} v_{ik'}^{\sigma}}. \tag{8}$$

The demand elasticity is

$$\varepsilon_{ik} = \sigma(1 - q_{ik}).$$

The expected utility of borrower i is

$$EU_i \equiv \mathbb{E}\left[\max_k \xi_{ik} v_{ik}\right] = \left(\sum_{k=1}^{K} v_{ik}^{\sigma}\right)^{\frac{1}{\sigma}}.$$

$\sigma > 0$ is an important parameter. It captures the substitutability of loans across banks and it relates inversely to the variance of the idiosyncratic taste shocks. Banks are more substitutable when σ is high. As we show below, in the limit as $\sigma \to \infty$, banks become perfect substitutes. Conversely, as $\sigma \to 0$, the choice probability for any given bank converges to $\frac{1}{K}$ regardless of the contractual utilities $\{v_{ik'}\}$.

Each bank's profit maximization problem in this parametrized model can be written as

$$\max_{r,L} \underbrace{\left[p_i(1+r)+(1-p_i)\delta_i - c_k\right]L}_{\substack{\text{expected profit conditioninal} \\ \text{on contract being accepted}}} \times \underbrace{\frac{v_{ik}^{\sigma}}{\sum_{k'=1}^{K} v_{ik'}^{\sigma}}}_{\text{choice probability}} \quad \text{s.t. } v_{ik} = p_i\left(z_i L^{\alpha} - (1+r-\delta_i)L\right). \tag{9}$$

We define bank k's profit margin as

$$\mu_{ik}(r,L) \equiv \frac{\pi_{ik}(r,L)}{(c_k-\delta_i)L} = \frac{p_i(1+r-\delta_i)-(c_k-\delta_i)}{c_k-\delta_i}.$$

Because borrowers always repay δ_i fraction of the investment, $(c_k-\delta_i)$ can be interpreted as the effective marginal cost of lending. The profit margin μ_{ik} is therefore the ratio between expected bank profit and effective marginal cost, conditioning on the loan being accepted.

Proposition 2. *The Laissez-faire equilibrium of the parametrized model has the following features.*

1. *Every borrower chooses bank k with the same probability: $q_{ik} = s_k$ for all i, where s_k is the market share of bank k.*

2. *Loan terms satisfy*

$$L_{ik} = \left(\frac{\alpha p_i z_i}{c_i - \delta_i}\right)^{\frac{1}{1-\alpha}}, \tag{10}$$

$$\mu_{ik} = \frac{1-\alpha}{\alpha}\frac{1}{1+\sigma(1-s_k)}. \tag{11}$$

3. *Let $HHI \equiv \sum_{k=1}^{K} s_k^2$ denote the Herfendahl index in the lending market. Then the average profit margin ($\mu = \frac{\int \mu_{ik} s_k L_{ik} dF}{\int \sum_{k'=1}^{K} q_{ik'} s_{k'} dF}$) in the market can be written as*

$$\mu \quad \approx \quad \frac{1-\alpha}{\alpha(1+\sigma)} + \frac{(1-\alpha)}{\alpha(1+\sigma)}\frac{\sigma}{(1+\sigma)} \times HHI,$$

where the approximation error is $o\left(\max_k (s_k)^2\right)$, i.e., second-order in the market share of the largest bank.

The first part of the proposition states that each bank's market share captures the choice probability for any borrower in the market. Note that market share s_k is itself an endogenous outcome of bank competition. Banks may differ in their equilibrium market share due to heterogeneity in funding cost, c_k. Banks with lower funding costs have higher market share. When all banks have identical funding costs, they also have the same market share: $s_k = 1/K$ for all k, where K is the total number of banks.

The second part of Proposition 2 characterizes equilibrium loan terms. Equation (10) is an application of Proposition 1, that equilibrium loan size is efficient and solves $\max_L p_i z_i L^{\alpha} + \delta_i L - c_i L$. Equation (11), which is derived from (7), solves for the equilibrium profit margin and thus the interest rate. To understand this, note that the total surplus generated by a loan of size L, $p_i z_i L^{\alpha} - (c_k - \delta_i) L$, is equal to $\frac{1-\alpha}{\alpha}(c_k - \delta_i) L$ and depends on the concavity of borrower's production technology, α. The profit margin therefore depends on α. The remaining term, $\frac{1}{1+\sigma(1-s_k)}$ captures the fraction of surplus accrued to the bank and is a direct consequence of (7). Bank's share of surplus relates to its market power (recall $\sigma(1-s_k) = \varepsilon_{ik}$): banks have higher profit margins when they are less substitutable (lower σ). Banks with greater market shares (higher s_k) also have higher profit margins.

The last part of Proposition 2 shows that in equilibrium, the average profit margin in a given market is approximately linearly in the HHI index for bank loans. The HHI index is a summary statistics for market concentration and it ranges between zero and one. Higher HHI indicates greater market concentration. The index is equal to 1 when a single bank captures the entire market and is equal to $1/K$ when all K banks are symmetric. The HHI index can therefore also be seen as inversely related to the effective number of banks operating in the market. The proposition implies that, holding borrower characteristics constant, profit margins and interest rates are higher in more concentrated markets, i.e., when there are fewer banks or when banks have more asymmetric lending costs. The proposition also implies that profit margins and interest rates are higher when banks are less substitutable (lower σ). These results are intuitive: when there are fewer competing banks or when banks are less substitutable, demand for loans from a specific bank should become more inelastic, as a marginal increase in interest rate—and the consequent reduction in contractual utility—should lead to a smaller outflow of potential borrowers. Consequently, competition is weaker, and banks offer loan terms that are less favorable to borrowers.

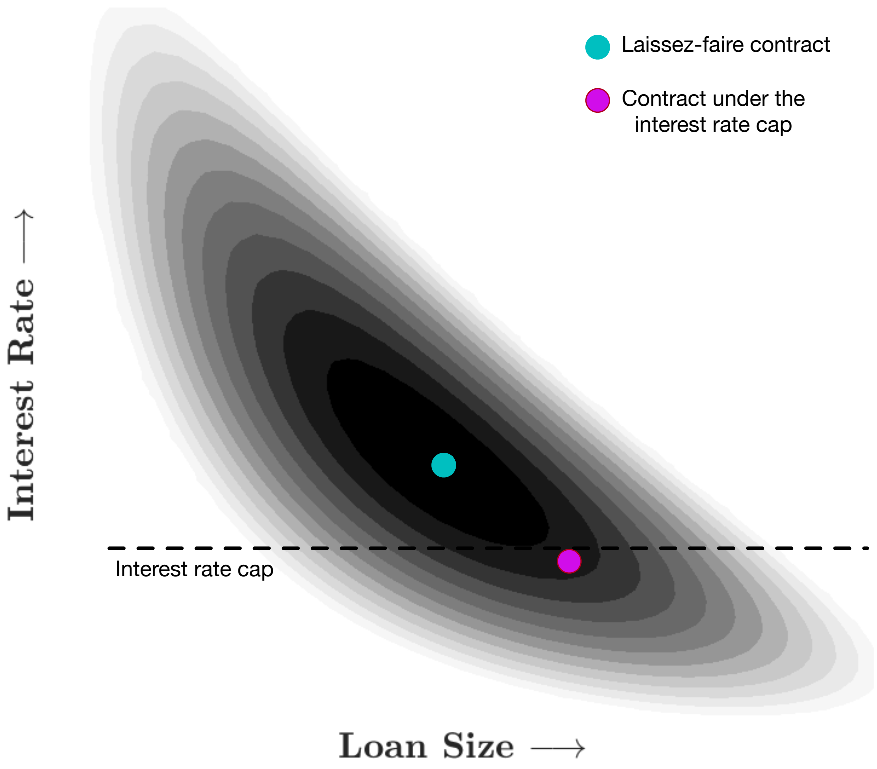
3 Banks' Response to Policy Interventions

We now analyze how banks respond to constraints in the contract space imposed by policy. We conduct this analysis for two reasons. First, when we estimate the model in the data, our identification strategy exploits banks' response to constraints in the contract space to recover model primitives. Second, interest rate caps are common policy tools; this section guides our analysis of these policies as we perform counterfactuals in section 6.

We first analyze how banks respond to simple, flat constraints on the interest rate and loan size. We then analyze how banks respond to interest rate caps that vary with loan size.

Interest Rate and Loan Size Are Strategic Substitutes Because contracts are two-dimensional, banks have two levers to extract profit from borrowers: either by raising the interest rate or by increasing the loan size. In equilibrium, a bank sets contractual terms to balance the trade off between extracting profits π_{ik} and leaving surplus to the borrower to raise the choice probability q_{ik}. Imposing any binding constraints over one of the choice variables r and L will intuitively cause banks to respond over the other choice variable. More importantly, because a higher value in either r or L raises π_{ik} and lowers q_{ik}, the two choice variables are strategic substitutes, meaning imposing a binding interest rate cap leads banks to over-lend, as loan size becomes larger than what's efficient; likewise, imposing a binding loan size cap leads banks to charge higher interest rates than what would have prevailed absent the constraints.

Figure 3: Contour plot of bank's profit as a function of contractual terms



To understand this better, Figure 3 shows a contour plot of bank's iso-profit curves as a function of the two choice variables r and L. Darker shades indicate higher profits. Because bank's maximization problem is concave, the profit function is single peaked: the blue dot indicates the contract that would have been offered if no policy constraints are imposed. Now imagine an interest rate cap (dotted horizontal line) is imposed so that the contract in blue is no longer feasible.

Which contract does the bank offer now? The bank would choose, among all feasible contracts, one with the highest profit—the darkest spot—in the constrained set. Because the decline in profits is least steep in the direction of higher L, the bank conforms to the interest rate cap and sets a larger loan size, as indicated by the red dot. Likewise, a binding loan size cap induces the bank to raise the interest rate.

The property, that r and L are strategic substitutes, holds with and without the parametric assumptions. Under the parametric forms, we are further able to provide analytic solutions to the contractual response under various policy constraints. For notational simplicity, we now assume all K banks are symmetric, and we characterize how contracts change in response to policy constraints. We drop the subscript k whenever it's unambiguous.

Simple Caps

Proposition 3. *Consider a bank's profit maximization problem (9) under additional constraints. Let (r_i^*, L_i^*) represent the Laissez-faire contract.*

1. *Consider the constraint $r_i \le \bar{r}$. If $p_i(1+\bar{r}-\delta_i) < c_k - \delta_i$, then the loan will be rationed as the bank is no longer able to recover lending cost under the constraint. Otherwise, the equilibrium contract is*

$$(r_i, L_i) = \left(\min\left\{\bar{r}, r_i^*\right\}, L_i^* \times \max\left\{1, \left(\frac{1+r_i^*-\delta_i}{1+\bar{r}-\delta_i}\right)^{\frac{1}{1-\alpha}}\right\}\right).$$

2. *The equilibrium contract under the constraint $L_i \le \bar{L}$ satisfies*

$$\left(\underbrace{\frac{p_i(1+r_i-\delta_i)-(c_k-\delta_i)}{c_k-\delta_i}}_{\textit{profit margin}}, L_i\right) = \left(\frac{\min\left\{\left(\bar{L}/L_i^*\right)^{\alpha-1}, 1\right\}/\alpha - 1}{1+(1-1/K)\,\sigma}, \min\left\{\bar{L}, L_i^*\right\}\right).$$

Loan Size-Dependent Interest Rate Caps Now consider size-dependent interest rate caps, i.e., an interest ceiling $\bar{r}^H$ for loans size below $\bar{L}$ and ceiling $\bar{r}^L < \bar{r}^H$ for loan size above $\bar{L}$. We continue to use (r_i^*, L_i^*) to represent the Laissez-faire contract and use (r_i, L_i) to represent the equilibrium contract under the policy.

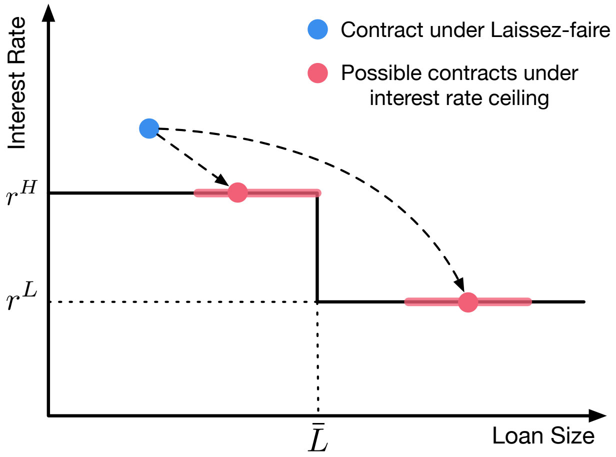
Because the two choice variables r and L are strategic substitutes, we can intuitively categorize each bank's response to the size-dependent interest rate cap into three scenarios, depending on borrower i's characteristics and the policy environment $\left(\bar{r}^L, \bar{r}^H, \bar{L}\right)$:

A. $r_i^* < \bar{r}^L$, or ($r_i^* < \bar{r}^H$ and $L_i^* < \bar{L}$): for these borrowers, the interest rate ceilings do not bind.

B. $r_i^* > \bar{r}^H$ and $L_i^* < \bar{L}$: for these borrowers, equilibrium loan terms have two possibilities other than being rationed:

   (a) $L_i \leq \bar{L}$ and $r_i = \bar{r}^H$;

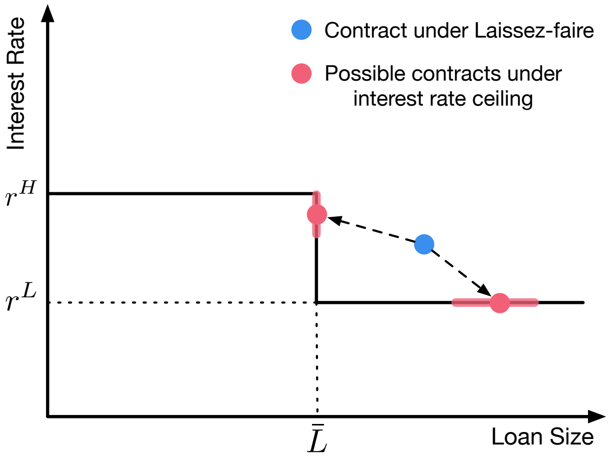
   (b) $L_i > \bar{L}$ and $r_i = \bar{r}^L$.


Contract under Laissez-faire
Possible contracts under interest rate ceiling
Interest Rate
r^H
r^L
$\bar{L}$
Loan Size

C. $r_i^* > \bar{r}^L$ and $L_i^* \geq \bar{L}$: for these borrowers, equilibrium loan terms have two possibilities other than being rationed:

   (a) $L_i > \bar{L}$ and $r_i = \bar{r}^L$;

   (b) $L_i = \bar{L}$ and $r_i \in (\bar{r}^L, \bar{r}^H]$.


Contract under Laissez-faire
Possible contracts under interest rate ceiling
Interest Rate
r^H
r^L
$\bar{L}$
Loan Size

The next proposition formally characterizes the equilibrium contract.

Proposition 4. *Suppose the Laissez-faire contract* (r_i^*, L_i^*) *is infeasible under the policy environ-*

ment with rate cap $\bar{r}^H$ for $L < \bar{L}$ and $\bar{r}^L$ for $L > \bar{L}$. Let $\left(r_i^L, L_i^L\right) \equiv \left(\bar{r}_i^L, L_i^* \left(\frac{1+r_i^*-\delta_i}{1+\bar{r}^L-\delta_i}\right)^{\frac{1}{1-\alpha}}\right)$, and let

$$\left(r_i^H, L_i^H\right) \equiv \begin{cases} \left(\bar{r}^H, \min\left\{\bar{L}, L_i^* \left(\frac{1+r_i^*-\delta_i}{1+\bar{r}^H-\delta_i}\right)^{\frac{1}{1-\alpha}}\right\}\right) & \text{if } L_i^* < \bar{L} \\ \left(\min\left\{\bar{r}^H, \frac{z_i\bar{L}^{\alpha-1}+\frac{c_k-\delta_i}{p_i}(1-1/K)\sigma}{1+(1-1/K)\sigma} - 1 + \delta_i\right\}, \bar{L}\right) & \text{if } L_i^* \geq \bar{L}. \end{cases}$$

If $p_i\left(1+\bar{r}^H-\delta_i\right) < c_k - \delta_i$, then the loan will be rationed. Otherwise, the equilibrium contract under rate ceilings is either $\left(r_i^L, L_i^L\right)$ or $\left(r_i^H, L_i^H\right)$. The equilibrium contract is $\left(r_i^H, L_i^H\right)$ if and only if

$$V_i \equiv \frac{\left(\left(1+r_i^H-\delta_i\right)-\left(c_k-\delta_i\right)/p_i\right)L_i^H}{\left(\left(1+r_i^L-\delta_i\right)-\left(c_k-\delta_i\right)/p_i\right)L_i^L} > 1.$$

For borrowers whose Laissez-faire contract is infeasible under the interest rate ceilings, bank either offer smaller loans with higher interest rates, $\left(r_i^H, L_i^H\right)$, or larger loans with lower interest rates, $\left(r_i^L, L_i^L\right)$. The object V_i represents the relative bank payoff between offering high interest rate (and small) loan and offering low interest rate (and large) loan. Marginal borrowers are those with $V_i = 1$.

4 Data and Empirical Setting: SBA Express Loan Program

We apply and estimate our model using the universe of small business loans made through the Small Business Administration (SBA) lending program from 2008-2017. In this section we describe the SBA guaranteed lending program and provide some descriptive statistics of the data. In the next section we discuss the identification strategy that allows us to estimate the model parameters using the empirical policy variation.

The SBA is an independent federal government agency. It provides commercial lenders with an indirect guarantee on loans made to small businesses who document that they have been turned down for alternative forms of credit. Lenders pay a fixed fee (typically 1 to 3% of loan principal) to the SBA in return for a guarantee that the SBA will reimburse a certain percentage of loan principal in the case of default. Loans made through the SBA guarantee program are subject to specific rules and regulations, including the interest rate cap studied here. Over 2,000 commercial lenders across the entire country participate in the program, and offer guaranteed SBA loans to clients who qualify.

The coverage, granularity, and policy variation contained within this dataset makes it the ideal laboratory to study market concentration. The dataset contains contract-level information on loan terms (interest rate, size) and repayment outcomes, borrower identity and characteristics and bank identity. We know the location of both borrowers and banks, which allows us to generate measure

Table 1: Summary Statistics for the SBA Express Loan Program

	Mean	Std Dev
Avg Interest Rate	6.86	.428
Loan Size	71,925	77,292
Charge-off Rate	0.03	.172
% at Cap	11.3	
Maturity	77.81	31.4
N	240,188	

This table displays summary statistics for loans used in our estimation sample from the SBA Express Loan program, 2008-2017. Average interest rate is expressed in percentage points, and is captured when the loan is first made. It includes both the base rate and the margin. Loan size is expressed in dollar units. The charge-off rate is calculated using a dummy variable for whether or not a loan has gone into default. % at Cap is calculated using a dummy variable that indicates whether a loan has an interest rate that is at either the lower or high interest rate cap. Maturity is expressed in months.

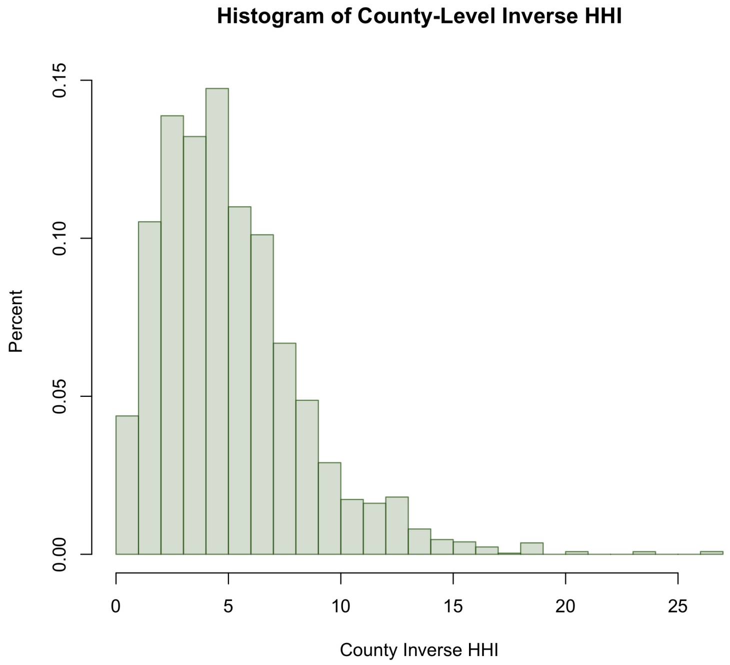
of market concentration both cross-sectionally and over time. Borrowers in the SBA market must also document that they have been turned down for other forms of credit; this creates a clearly defined market of banks (regional SBA lenders) for that particular borrower, and rules out the possibility that borrowers are "topping up" their SBA loans with additional sources of credit. For these borrowers, the relevant market of lenders are those we observe participating in the SBA program.

Table 1 presents summary statistics for our Express loan sample, which includes 240,188 loans made under the SBA Express program between 2008 and 2018. On average, these loans are \$71,925 in size, and have a maturity of 6.5 years. Interest rates for SBA Express loans can be fixed or variable; they are tied to base rates, with the maximum allowable interest rate ranging from 4.5 to 6.5 percent above the base rate, depending on loan size.[2] The average interest rate in our sample is 6.9%, well above typical rates for corporate loans.

Despite the fact that the SBA lending market is heavily regulated, we still observe strong suggestive evidence in the data of imperfect competition. We calculate an *inverse* Herfendahl-Hirschman Index (HHI) based on the dollar volume lending share (s_{kct}), of each bank k within a given county c and year t: $InverseHHI_{c,t} = \frac{1}{\sum_{k=1}^{K} s_{kct}^2}$. A value of 1 means that a single bank holds the entire market share, whereas larger values signal less market concentration. In a market where banks have equal market shares, the inverse HHI is simply the number of banks in the market. Figure 4, which plots the distribution of $InverseHHI_{c,t}$, suggests that a large portion of markets are monopolistic and only a small minority of markets have a substantial level of lender competition. The inverse

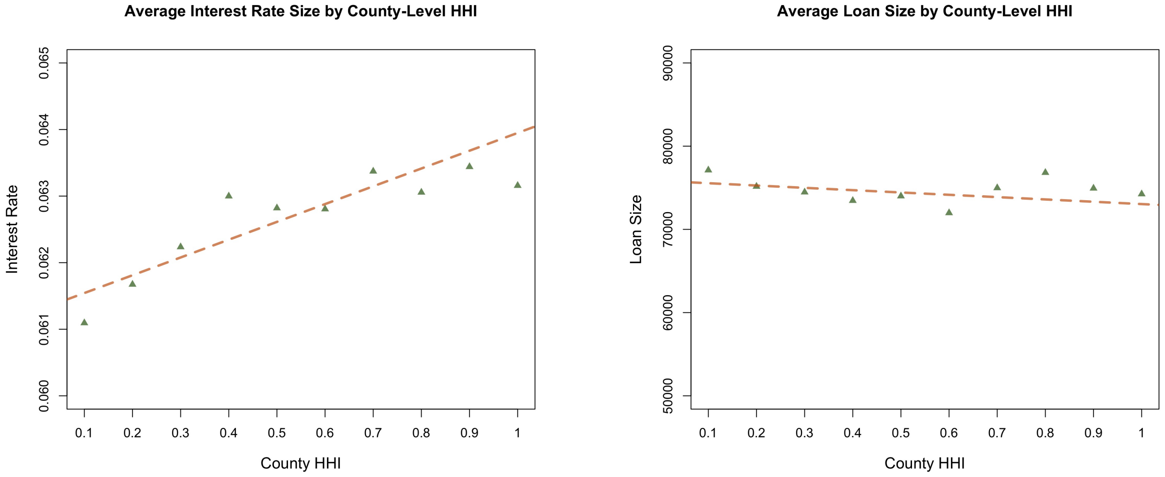
[2] These base rates are the prime rate, the LIBOR, and the PEG, which can fluctuate based on market conditions. For variable rates, the base rate for the computation of interest rate is the lender's choice, provided that the maximum interest rate the borrower is charged still does not exceed prime rate plus 4.5 percent to 6.5 percent. It should also be similar to the rates the lender charges for other, similarly-sized, non-SBA guaranteed loans.

Figure 4: Distribution of County-Year Inverse HHI



This figure plots the distribution of inverse HHI over all county-year observations in our data. We calculate an *inverse* Herfendahl-Hirschman Index (HHI) based on the dollar volume lending share (s_{ict}), of each bank within a given county-year: $InverseHHI_{c,t} = \frac{1}{\sum_{i=1}^{N} s_{ict}^2}$. A value of 1 means that a single bank holds the entire market share, whereas larger values signal less market concentration. The majority of counties in a given year in our dataset are dominated by less than 4 lenders.

Figure 5: Observed Average Interest Rate and Loan Size by Market Concentration



The left-hand figure plots the average interest rate charged in markets with differing levels of concentration, as measured by number of competing banks within a zip code. The interest rate measure controls for loan maturity, log size, business NAICS category, time fixed effects, ex-post performance, and bank brand. The plot suggests that higher interest rates are charged in more concentrated markets. Default (i.e. charge-off) rates do not increase in more concentrated markets; if anything, loans in more concentrated markets are less costly for lenders. Therefore risk-related costs cannot explain the downward sloping relationship between competition and interest rates. The right-hand figure plots the average loan size in each of these markets, which is relatively flat.

HHI of a median county-year is 4.5.

We also observe an impact of market concentration on loan pricing. Figure 5 documents a strong positive relationship between the average initial interest rate charged on observationally identical loans[3] within a county, and the HHI of that county. A similar relationship exists for other measures of market concentration (e.g. number of banks in the county or zipcode HHI, banks within an X-mile radius). This is not driven by changes in borrower risk across markets – we plot ex-post measures of default again market HHI, and find no relationship. The right hand panel plots average loan size across the same measure of market concentration.

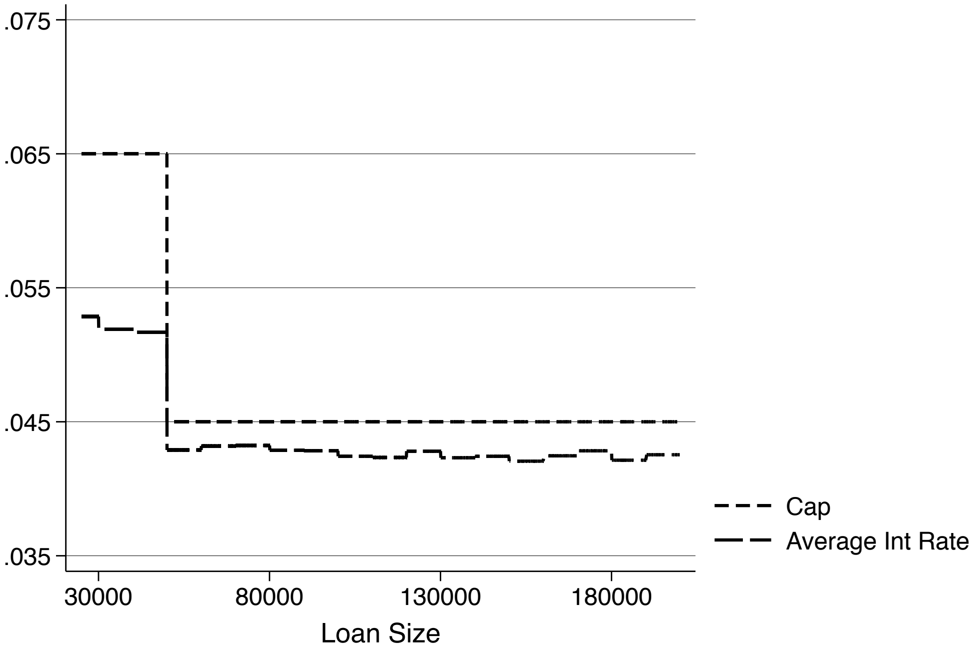
While the relationship between HHI and interest rates shown above motivates an analysis of market power, it remains suggestive; an exogenous shift or shock to lenders' maximization problem is required to separate and identify the relevant demand and supply parameters from our model. Loans made through the SBA Express program are subject to specific SBA rules and regulations that provide this identifying policy variation. Specifically, they face a loan-size dependent interest rate cap — loans smaller or equal to \$50,000 are capped at Prime + 6.5%, while loans larger than 50,000 are limited to Prime + 4.5%. This "notch" in the interest rate cap imposes a size-dependent constraint on banks' pricing problem, and generates specific lending and pricing responses under

[3] We control for bank brand (i.e. West America, Chase, etc), borrower business NAICS code, loan maturity, and time fixed effects.

our model.

Figure 6 shows the interest rate cap over loan size in our dataset, as well as the average interest rate by loan size; there is a clear impact of the cap on prices in the market. In total 11% of loans are constrained by the interest rate cap, 8.5% of loan above the threshold and 13% of loans below the threshold. SBA regulations do not allow lenders to originate multiple loans to the same borrower at the same time. Thus lenders cannot "piggyback" loans to take advantage of the notch.

Figure 6: Interest Rate Ceiling and Average Interest Rate (Minus the Base Rate) by Loan Size



Loans made through the SBA Express program are subject to specific SBA rules and regulations that provide identifying policy variation. Specifically, we use the fact that they face a loan-size dependent interest rate cap — loans smaller or equal to $50,000 are capped at Prime + 6.5%, while loans larger than 50,000 are limited to Prime + 4.5%. This figure plots the interest rate cap (minus the Prime base rate), and the average interest rate by loan size.

5 Two Dimensional Bunching: Identification

We use the empirical distribution of contractual terms under a loan-size-dependent interest rate caps policy to identify model parameters. Our identification strategy builds on and extends the “bunching” literature that uses kinks and notches to identify key elasticities (Kleven (2016)). Broadly speaking, this approach exploits discontinuities in economic agents’ choice sets and the consequent distortions in the equilibrium outcome distribution to infer structural parameters that govern economic behaviors. The bunching approach requires two steps: 1) recover a counterfactual distribution $H^0(\cdot)$ of equilibrium outcome absent the policy discontinuity; 2) use the difference between the counterfactual distribution and the observed, equilibrium distribution $H^P(\cdot)$ under policy to infer parameters.

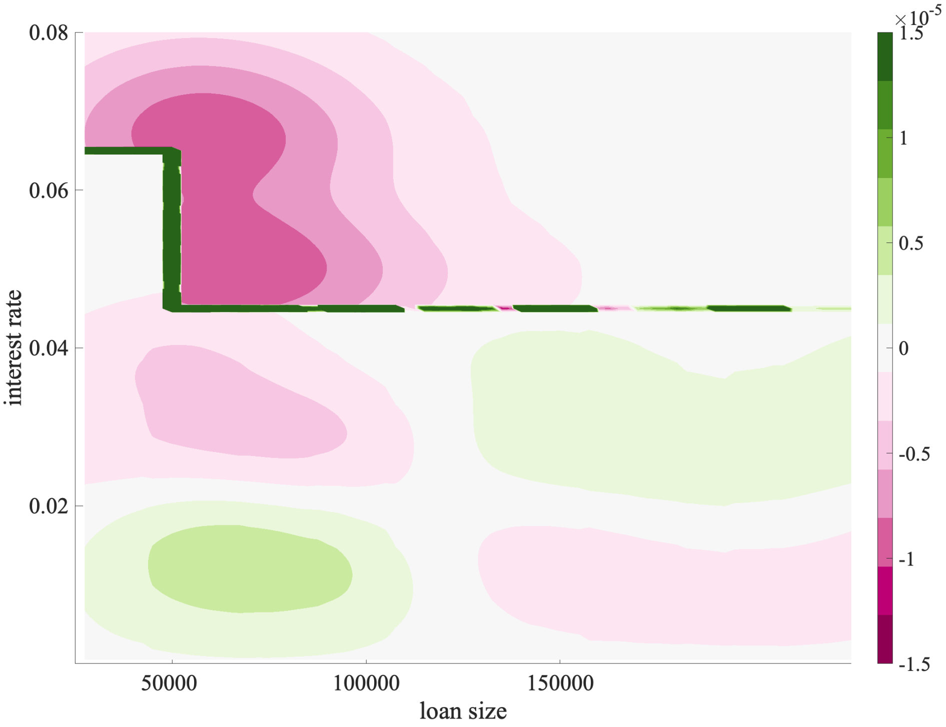
Our methodological contribution is to advance the bunching approach to a multi-dimensional behavioral response.

First, we recover the two-dimensional joint distribution $H^0(r,L)$ of Laissez-faire contracts from the observed distribution of contracts $H^P(r,L)$ in the data. We start from the subsample of loans with interest rates strictly below the cap, $S \equiv \left\{r,L | r < \bar{r}^L \text{ or } \left(r < \bar{r}^H \text{ and } L < \bar{L}\right)\right\}$, and we recover $H^0(r,L)$ from the conditional distribution of loans $H^P(r,L|S)$. This strategy is motivated by the fact that if it were optimal for banks to offer a contract in set S under Laissez-faire, then such a contract is still optimal and available even in the presence of the policy cap. Moreover, because each bank's profit maximization problem is concave, the policy cap does not move any Laissez-faire contract strictly outside of the set S to the interior region strictly below the policy cap. Therefore, the distribution of contracts in set S under the policy cap coincides with the conditional distribution under Laissez-faire: $H^P(r,L|S) = H^0(r,L|S)$. Under the assumption that $H^0(\cdot)$ is analytic over its domain—a standard assumption in the bunching literature—we then recover $H^0(\cdot)$ over the entire two-dimensional domain by extrapolating from the conditional distribution $H^P(r,L|S)$. Section 6.1 provides a detailed discussion of the statistical procedure that recovers $H^0(r,L)$ from $H^P(r,L|S)$.

Once we have the counterfactual and empirical distribution of contracts H^0 and H^P, we then take the difference between the two and define $D \equiv H^P - H^0$. Conceptually, we could do this market by market, and have a collection of D's across markets. We refer to D simply as the "distortion", as it represents the distortion in the distribution of contracts due to the interest rate cap. As an example, D for a particular market is visualized in Figure 7. In green are the regions in which there are "excess mass"—i.e. the observed joint distribution has more mass than the predicted counterfactual distribution. The excess mass is concentrated along the interest rate cap, where banks have "bunched" loans that would have otherwise existed *above* the cap. In pink are the regions in which there is "missing mass", where the observed distribution has less mass than the predicted counterfactual distribution. Since banks are not allowed to make loans above the cap, this missing mass is concentrated in the region above the cap. In principle, the two distributions should be identical strictly below the cap; any difference therein is due to imperfect fit in our estimation procedure.

We now discuss how to identify model parameters α and σ based on the collection of distortions D across markets. We form moment conditions guided by Proposition 4, which describes how banks would offer contracts in the presence of a size-dependent interest rate cap. The proposition describes which loans under a laissez faire regime, (L^*, r^*), would be distorted under a size

Figure 7: Difference between empirical distribution H^P and counterfactual distribution H^0



dependent interest rate cap, and importantly *where* they would be relocated, as functions of α and σ. In other words, it describes how the distortions we observe in D relate to the model parameters. It directly links the observables, L, r, and K, to the structural parameters. This approach enables us to be completely agnostic about borrower heterogeneity: we allow for an arbitrary distribution of borrower characteristics (z_i, p_i, δ_i), and the distribution could vary arbitrarily across markets.

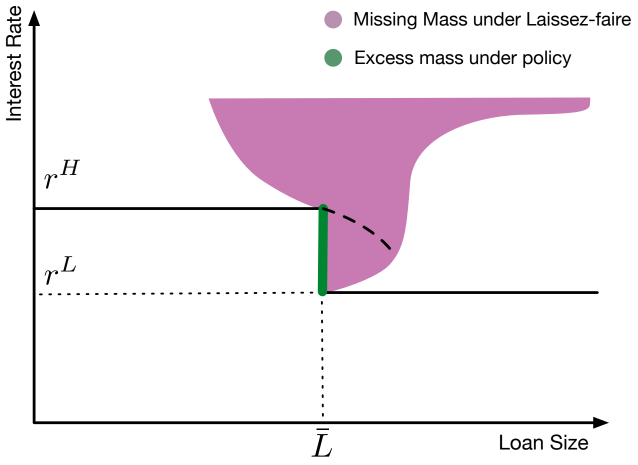
Specifically, we use two moment conditions per market for identification. These conditions equate the excess mass at the notch (where $L = \$50,000$) to specific regions of missing mass, as illustrated in figure 8. The notch is painted green to represent excess mass and the purple region represents missing mass. There are two purple regions separated by a dotted curve. Our first moment condition equates the excess mass at the point $\left(\bar{r}^H, \bar{L}\right)$ to the missing mass in the upper purple region, which contains the set of Laissez-faire contracts that would relocate to $\left(\bar{r}^H, \bar{L}\right)$ under the policy cap. Analogously, our second moment condition equates the excess mass along the notch but excluding the end points (i.e. excess mass over the set $B \equiv \left\{(r, L) \,\middle|\, r \in (\bar{r}^L, \bar{r}^H), L = \bar{L}\right\}$) to the missing mass in the lower purple region. The boundaries that define the two purple regions vary continuously with model parameters α and σ. Under the true structural parameters that generate the data, the excess mass in green should be exactly equal to the missing mass in purple.

The intuition for selecting these two moments is as follows. Recall that α and σ respectively capture 1) how lending surplus relates to loan size and 2) how each bank's market power varies with its market share. First consider a Laissez-faire contract with $r_i^* > \bar{r}^H$ and $L_i^* < \bar{L}$. If the loan is not rationed under the policy cap, Proposition 4 shows that the equilibrium contract will carry interest rate $r_i = \bar{r}^H$ and loan size $L_i = L_i^* \left(\frac{1+r_i^*}{1+\bar{r}^H}\right)^{\frac{1}{1-\alpha}} < \bar{L}$ if r_i^* and L_i^* are relatively small. For this contract, the rate cap $\bar{r}^H$ binds but the loan size remain locally unconstrained ($L_i < \bar{L}$), in which case the equilibrium loan size depends only on the structural parameter α and not σ. Intuitively, because interest rate cap binds, the bank's market power becomes irrelevant in choosing loan terms, and the parameter α governs the distortion in equilibrium loan size because it captures the elasticity of investment output to L. Consequently, the shape of the left boundary to the upper triangle—defined by the set $\left\{(r,L)\,|L\left(\frac{1+r}{1+\bar{r}^H}\right)^{\frac{1}{1-\alpha}} = \bar{L}\right\}$—is entirely pinned down by α and not σ.

Second, consider a Laissez-faire contract with $r_i^* \in \left(\bar{r}^L, \bar{r}^H\right)$ and $L_i^* > \bar{L}$. If the loan is not rationed, Proposition 4 shows that the corresponding equilibrium contract must either scale back loan size to $L_i = \bar{L}$ and charge a relatively higher interest rate $r_i \in [r_i^*, \bar{r}^H]$, or conform to the lower rate cap $r_i = \bar{r}^L$ and remain unconstrained on loan size $L_i > L_i^*$. For a fixed Laissez-faire contract, which of the two scenarios prevail in equilibrium depends on whether the contract falls to the left or to the right of the lower boundary of the lower purple region. The shape of that boundary in turn depends on banks' market power. Intuitively, when market power is low, the profit margin underlying the Laissez-faire contract is also low; hence, conforming to the lower rate cap $\bar{r}^L$ represents a disproportionately large decline in the profit margin and is relatively unattractive. In this case, banks are more likely to scale back loan size to maintain a larger profit margin. An extreme case is perfect competition and no market power—banks are perfect substitutes $\sigma \to \infty$ and there are multiple banks $K > 1$—the Laissez-faire rate r_i^* fully reflects marginal lending cost, and conforming to the lower rate cap $\bar{r}^L$ would generate losses to the banks; hence, banks have no choice but to scale back loan size and offer $r_i = r_i^*$, $L_i = \bar{L}$. Conversely, when the Laissez-faire contract has high profit margin, conforming to lower rate cap implies a relatively small decline in profit margin, and banks are more likely to find this option attractive relative to distorting loan size.

The discussion above illustrates that the lower boundary of the lower purple region depends on banks' market power and the choice elasticity. Because the choice elasticity is equal to $\sigma(1-1/K)$, we first use the moment condition to identify the choice elasticity within each market, and we then utilize the variation in choice elasticity across markets with varying number of banks and HHIs to recover the structural parameter σ.

Figure 8: Excess and Missing Mass Regions Used for Identification


● Missing Mass under Laissez-faire
● Excess mass under policy
Interest Rate
r^H
r^L
$\bar{L}$
Loan Size

Formally, let

$$S_K^1 \equiv \left\{ (r_i, L_i) \left| \left(\frac{(\bar{L}/L_i)^{\alpha} - \alpha \bar{L}/L_i}{1 + (1-1/K)\,\sigma} \right) > \left(\frac{1 + (1-1/K)\,\sigma\alpha}{1 + (1-1/K)\,\sigma} \left(\frac{1 + r_i - \delta_i}{1 + \bar{r}^L - \delta_i} \right)^{\frac{\alpha}{1-\alpha}} - \alpha \left(\frac{1 + r_i - \delta_i}{1 + \bar{r}^L - \delta_i} \right)^{\frac{1}{1-\alpha}} \right), L_i > \bar{L} \right. \right\}$$

$$S_K^2 \equiv \left\{ (r_i, L_i) \left| (1 + r_i) \frac{(\bar{L}/L_i)^{\alpha-1} + (1-1/K)\,\sigma\alpha}{1 + (1-1/K)\,\sigma\alpha} < 1 + \bar{r}^H, L_i > \bar{L} \right. \right\}.$$

The intersection $S_K \equiv S_K^1 \cap S_K^2$ corresponds to the lower purple region in the figure above, and Laissez-faire contracts $(r^*, L^*) \in S_K$ will bunch into region $B \equiv \left\{ (r, L) \,\middle|\, r \in (\bar{r}^L, \bar{r}^H), L = \bar{L} \right\}$ under the size-dependent interest rate cap. Intuitively, S_K^1 picks out the Laissez-faire contracts that scale back to $\bar{L}$ (instead of over-lending), and S_K^2 picks out Laissez-faire contracts that charge strictly less than $\bar{r}^H$ under the policy intervention. That the excess mass in the region defined by B is equal to the missing mass in the region defined by S_K is our first moment condition for all markets with K banks.

To formalize the second moment condition, let

$$R_K^1 \equiv \left\{ (r, L) \,|L > \bar{L}, \right\}$$

$$R_K^1 \equiv \left\{ (r, L) \left| L \geq \bar{L}, \frac{(1+r)\,(L/\bar{L})^{1-\alpha} + (1+r)\,(1-1/K)\,\sigma\alpha}{1 + (1-1/K)\,\sigma\alpha} \geq 1 + \bar{r}^H \right. \right\}$$

$$R_K^2 \equiv \left\{ (r,L) \middle| r \geq \bar{r}^H, L\left(\frac{1+r}{1+\bar{r}^H}\right)^{\frac{1}{1-\alpha}} \geq \bar{L} \right\}$$

$$R_K^3 \equiv \left\{ (r,L) \middle| \left(1+\bar{r}^H - \frac{(1+r)}{\frac{1+(1-1/K)\sigma\alpha}{\alpha+(1-1/K)\sigma\alpha}}\right)\bar{L} \geq \left(1+\bar{r}^L - \frac{(1+r)}{\frac{1+(1-1/K)\sigma\alpha}{\alpha+(1-1/K)\sigma\alpha}}\right) L\left(\frac{1+r}{1+\bar{r}^L}\right)^{\frac{1}{1-\alpha}} \right\}.$$

The intersection $R_K \equiv R_K^1 \cap R_K^2 \cap R_K^3$ corresponds to the upper purple region, the Laissez-faire contracts in which will bunch back to the point $\left(\bar{r}^H, \bar{L}\right)$.

For each market structure K, we generate the following moment condition:

$$\begin{cases} \iint_{(r,L)\in B} dD(r,L) + \iint_{(r,L)\in S_K} dD(r,L) = 0 \\ D\left(\bar{r}^H, \bar{L}\right) + \iint_{(r,L)\in R_K} dD(r,L) = 0. \end{cases}$$

Intuitively, the two moment conditions for a single market are sufficient to recover α and the choice elasticity of that market. We exploit cross-market variation in order to recover σ, which governs how the choice elasticity varies with market structure. The model is therefore over-identified.

For computational simplicity, we sort markets by HHI and group them into 8 bins. We assume all banks are symmetric, and we set $1/K$ to be the average HHI of markets within each group. We estimate (σ, α) by exploiting the two moment conditions across groups of markets with varying average HHI.

6 Estimation and Results

Here we describe the empirical procedure for estimating the counterfactual distribution and the estimation of the model parameters using the set of indifference conditions that equate the excess and missing mass in our two-dimensional setting.

6.1 Estimating the counterfactual joint distribution

A precise measure of the excess mass requires that we compare the observed distribution of contracts to the counterfactual distribution that would exist in the absence of a notch. Therefore, in this section our goal is to get a reasonable estimate of the joint distribution of loan amount and interest rates in a hypothetical world in which the Small Business Administration did not impose a size-dependent interest rate cap. This is a nontrivial problem as we only observe loans created in an environment subject to this rate cap.

Following the identification argument in section 4, we restrict our sample to the subset of contracts that have interest rates below the interest rate cap. Within this subsample, we estimate the joint distribution of loan size and interest rate, allowing for a flexible correlation structure between the two variables.

While the distribution of both interest rates and loan size below the lower cap appears lognormally distributed, the presence of pronounced "round number bunching" at familiar basis point or dollar multiples generates some empirical challenges. For example, in fitting a smooth lognormal counterfactual joint distribution to both L and r, we would fail to reflect the "spiky" nature of $H^P(r,L|r<\bar{r}^L)$. Therefore we take a more non-parametric approach: we first fit a flexible polynomial with round number dummies to the marginal distribution of r, accounting for the fact that we observe only the truncated distribution of $(r|r<\bar{r})$. We then divide the loan size distribution into \$2,500 bins, and estimate the distribution of loan size *conditional* on interest rate. Using the estimated parameters, we then predict the distribution of contracts, $\left(\hat{r},\hat{L}\right)$, for $r>\bar{r}$.

Below is a more detailed description of the estimation procedure. It is composed of 3 central steps:

- Step 1: Derive an estimate for the marginal probability function $H_r^0 = P(R=r)$
- Step 2: Estimate the conditional probability function $H_{L|r}^0 = P(L=l|R=r)$
- Step 3: Combine the output from steps 1 and 2 to compute the counterfactual distribution $H^0 = P(L=l,R=r)$. Rescale the counterfactual mass so that it matches the observed mass below point of truncation.

1. Estimation of marginal density function, $P(R=r)$

We focus initially on estimating the marginal distribution of interest rates using the observed set of contracts with interest rates strictly below the rate cap. Our key assumption in using this set for estimation is that these contracts were not altered by the interest rate cap, and thus an identical set of loan contracts would have existed in the counterfactual world.

The distribution of observed contracts displays distinct round number bunching at predictable intervals, and is also truncated at the notch. Figure 17in the appendix contains the histogram and CDF of observed loans with rates below 6.5% and 4.5% respectively. In both the histograms and CDFs, we see significant spikes occurring at integer interest rates, and at multiples of 50 basis points and 25 basis points. These are marked by red, blue, and green lines respectively.

Using this observed data, we fit the following model using nonlinear least squares:

$$P(R\leq r) = \frac{e^{\eta}}{1+e^{\eta}}$$

where the linear predictor, η, is given by:

$$\eta = \alpha + \sum_{p=1}^{P} \beta_p r^p + \delta_1 \sum_{i=1}^{I} \mathbf{1}\{r \geq .01i\} + \delta_2 \sum_{j=1}^{J} \mathbf{1}\{r \geq .005 + .01(j-1)\}$$
$$+ \delta_3 \sum_{k=1}^{K} \mathbf{1}\{r \geq .0025 + .005(k-1)\}$$

Here, δ_1 measures the discontinuous jump in the linear predictor when r reaches a round integer interest rate, δ_2 measures the discontinuous jump when r reaches a multiple of 50 basis points, and δ_3 measures the discontinuous jump when r reaches a multiple of 25 basis points. α is an intercept of the linear predictor and β is the slope measuring how the linear predictor moves continuously with changes in r. The linear predictor is then inserted as an argument to the Sigmoid function as shown above giving the desired estimate. We vary the degree of the polynomial, P, over various specifications.

Using these coefficients, we can then estimate the $P(R \leq r)$ for $r \geq \bar{r}$ by imposing the assumption that the jumps to the linear predictors (measured by the δ coefficients) are the same for interest rates that would be located above the cap. Importantly for interpretation, this assumption does not impose that the same jump occurs in the CDF at, for example, 3% and 7% since the slope of the Sigmoid function is much greater at $\eta(.03)$ than at $\eta(.07)$. The graph in Figure 9 overlays the estimated CDF from the model (in red) with the unconditional CDF from the data (in black).

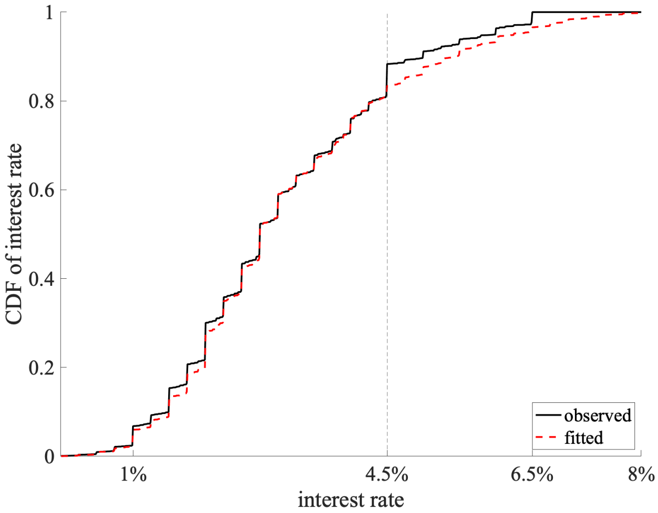
2. Estimation of conditional density function $P(L = l|R = r)$

In this next step, we estimate the density of loan size conditional on interest rates, $P(L = l|R = r)$. If L were independent of r, we could repeat step one, estimating the marginal distribution of loan size. However, the correlation between L and r in our sample is negative and significant. We therefore need to include the dependence on r in our model for the distribution of L.

To model this conditional probability distribution, we begin by discretizing L into bins of width \$2,500 and fitting the linear predictor such that it includes interaction terms between $log(L)$ and r:

$$\chi = \alpha + \sum_{p=1}^{P} \beta_p r * log(L)^p + \delta_1 \sum_{i=1}^{I} \mathbf{1}\{L \geq 5,000i\} + \delta_2 \sum_{j=1}^{J} \mathbf{1}\{L \geq 10,000j\}$$
$$+ \delta_3 \sum_{k=1}^{K} \mathbf{1}\{L \geq 25,000k\} + \delta_4 \sum_{m=1}^{M} \mathbf{1}\{L \geq 30,000j\} + \sigma_1 \sum_{i=1}^{I} r * \mathbf{1}\{L \geq 5,000i\}$$
$$+ \sigma_2 \sum_{j=1}^{J} r * \mathbf{1}\{L \geq 10,000j\} + \sigma_3 \sum_{k=1}^{K} r * \mathbf{1}\{L \geq 25,000k\} + \sigma_4 \sum_{m=1}^{M} r * \mathbf{1}\{L \geq 30,000j\}$$

Figure 9: Observed vs. Estimated Marginal Distribution of Interest Rates



This figure plots the estimated (in red) and observed (in black) marginal CDF of interest rates. The estimated CDF is created by fitting the model $P(R \leq r) = \frac{e^{\eta}}{1+e^{\eta}}$ using nonlinear least squares where the linear predictor, η, is given by $\eta = \alpha + \sum_{p=1}^{P} \beta_p r^p + \delta_1 \sum_{i=1}^{I} \mathbb{1}\{r \geq .01i\} + \delta_2 \sum_{j=1}^{J} \mathbb{1}\{r \geq .005 + .01(j-1)\} + \delta_3 \sum_{k=1}^{K} \mathbb{1}\{r \geq .0025 + .005(k-1)\}$. The use of various dummies variables accounts for the visible "spikes" occurring in the distribution at integer interest rates, and at multiples of 50 basis points and 25 basis points.

We then follow the estimation procedure described in step one, using NLLS to estimate the parameters of the linear predictor once inserted into a sigmoid function.

3. Creation of $H^0(L,r)$ and rescaling

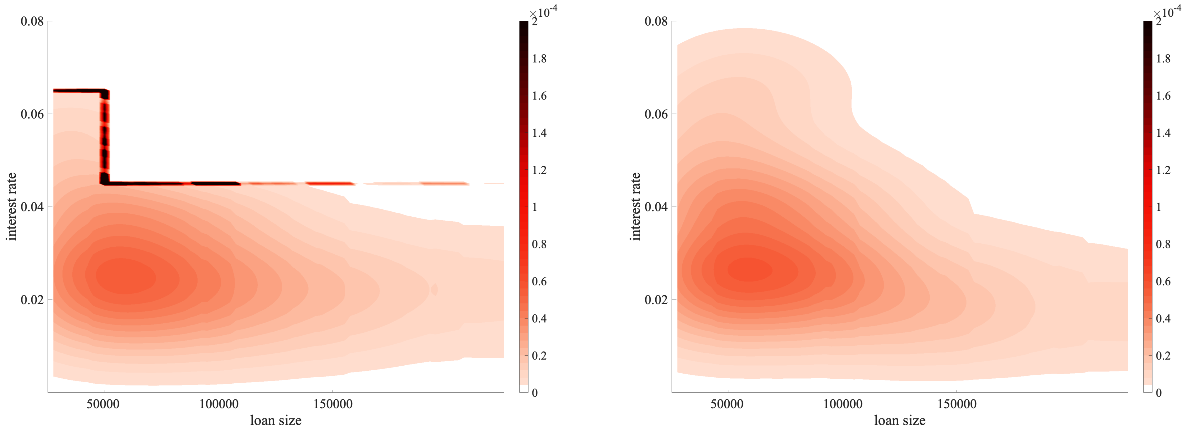
We are able to create the joint predicted distribution $H^0(L,r) = P(L = l, R = r) = P(L = l|R = r) * P(R = r)$ by multiplying the marginal and conditional distributions estimated in steps 1 and 2. We again rescale the counterfactual distribution such that:

- For each discretized value of r, the marginal Ldistribution is equal to 1 at the maximum loan size and equal to the observed distribution at $L = \$47,500$ for loans with interest rates from 4.5 to 6.5%.

- For each discretized value of L, the marginal r distribution is equal to 1 at the maximum interest rate, which we set to be 8%. Additionally, we need the observed and counterfactual marginal distributions of r to obtain the same value at 6.49% or 4.49% in the case of loans $\geq$ \$50,000.

Figure 10 plots the observed loan distribution and the predicted counterfactual distribution, pooling over all markets and loans. In the observed distribution, the excess mass at the threshold, \$50,000, and along the interest rate cap is pronounced. The predicted counterfactual distribution spreads this excess mass throughout the region where loan contracts would have been located in the absence of the discontinuity, both above and to the right of the threshold.

Figure 10: Counterfactual, Observed, and Difference in Density for average Loan Distribution



This figure plots the difference in density between the observed loan distribution and the counterfactual distribution, pooling over all markets and loans. Here the excess mass at the threshold, \$50,000, is pronounced and equal to 2 percentage points. One can also see that there is some missing mass to the right of the threshold, where loan contracts would have been located in the absence of the discontinuity.

6.2 Estimation of Parameters

For each market k we calculate the observed empirical joint probability density, $\hat{H}_k^P$ over a 2-dimensional grid, with grid points defined by the intervals $L = [25,000 : 2,500 : 250,000]$ and $r = [0 : .0001 : .8]$. The visible bunching in the loan size distribution at round number multiples requires that we use this discrete, rather than a continuous, approach. Using the method described above, we predict the counterfactual density, $\hat{H}_k^0$, over this same domain and calculate the difference between the two as $\hat{D}_k = \hat{H}_k^P - \hat{H}_k^0$.

Using $\hat{D}_k$, we calculate the empirical analogues to our theoretical moment conditions. Our estimation routine then chooses $(\hat{\alpha},\hat{\sigma}) = \arg\min R(\alpha,\sigma)$, where

$$R(\alpha,\sigma) = \sum_k \left[\left(\hat{E}_{k,1} + \hat{M}_{k,1} \right)^2 + \left(\hat{E}_{k,2} + \hat{M}_{k,2} \right)^2 \right],$$

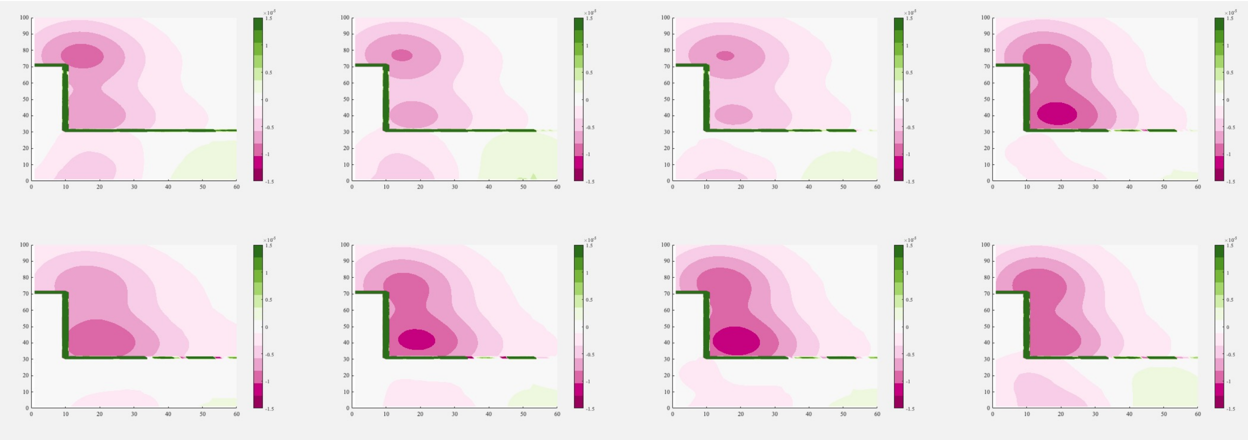
where $\hat{E}_{k,i}$ is the excess mass for the i-th moment condition in market k and $\hat{M}_{k,i}$ is the corresponding missing mass. The routine uses a grid-search approach across values of $\alpha = (0,1)$ and $\sigma = (0,25)$ to find the minimum of the objective function.

6.3 Implementation and Results

Our main specification splits the data into 8 quantiles of inverse HHI. The most concentrated market group has an average K of 1.4, while the less concentrated group has an average K of 9.45. In other specifications we divide the data into more than two quantiles of market concentration.

We construct $\hat{H}_k^0$, and $\hat{D}_k$ for $k = 1$ through 8 using the procedure outlined in section 6.1. Figure 11 plots the distribution of $\hat{D}_k$ across loan size and interest rate for the various groups; pronounced excess mass (in green) occurs along the border of the interest rate cap, where laissez-faire contracts with higher interest rates have been forced to "bunch". Missing mass (in pink) is concentrated above the cap and to the right of the notch. Visually, the missing mass shifts down as K increases, implying that markups are lower in more competitive markets.

Figure 11: Difference in counterfactual and observed (L, r) distribution across 8 quintiles of market concentration



This figure plots the difference in density between the observed loan distribution and the counterfactual distribution after dividing the data into 8 quintiles of inverse HHI, K. Excess mass is shown in green and is concentrated along the interest rate threshold. Missing mass is plotted in pink, and indicates where loan contracts would have been located in the absence of the cap. As predicted by our model, the missing mass is concentrated primarily above the cap and to the right of the notch. It is lower and more diffuse in the more competitive markets.

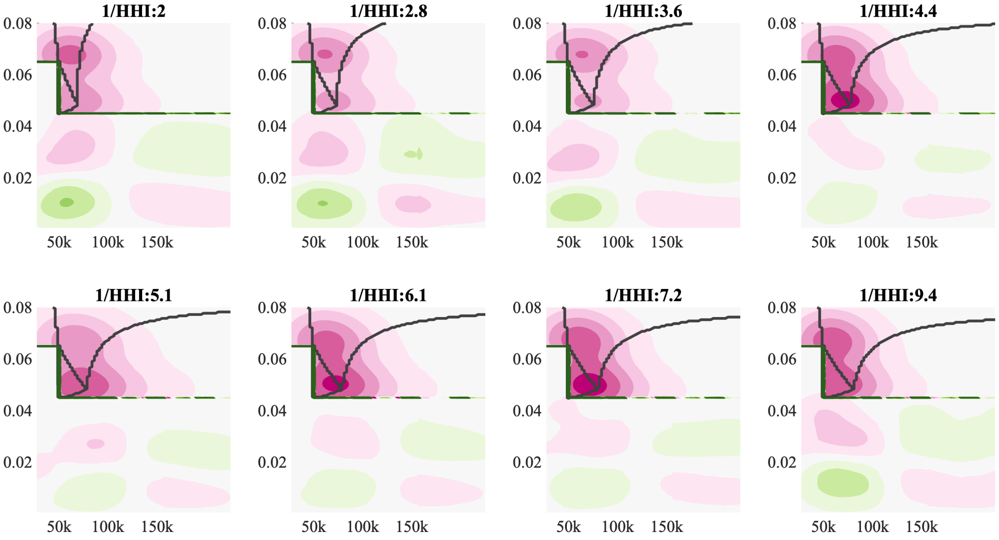
Using the various $\hat{D}_k$, we then choose the set of parameters that minimizes the difference between observed excess and missing mass in all markets. Figure **??** overlays the boundary of the missing mass region estimated by these parameters over the observed difference in distributions. Our estimate of the parameters implies that in a symmetric duopolistic market, lenders capture 34% of surplus (see Figure 13).

Table 2: Parameter Estimates

	α	σ
Estimate	.81	5.876
N		240,188

This table reports the estimated parameter values using two markets, above and below median K, for estimation. We use a grid search over possible values of $\sigma$and $\alpha$to find the set of parameters that minimizes the difference between observed excess and missing mass in both markets. These parameters imply that in a symmetric duopolistic market, lenders capture 34% of surplus.

Figure 12: Difference between Counterfactual and Observed Loan Distributions ($\hat{D}_k$) for large and small K

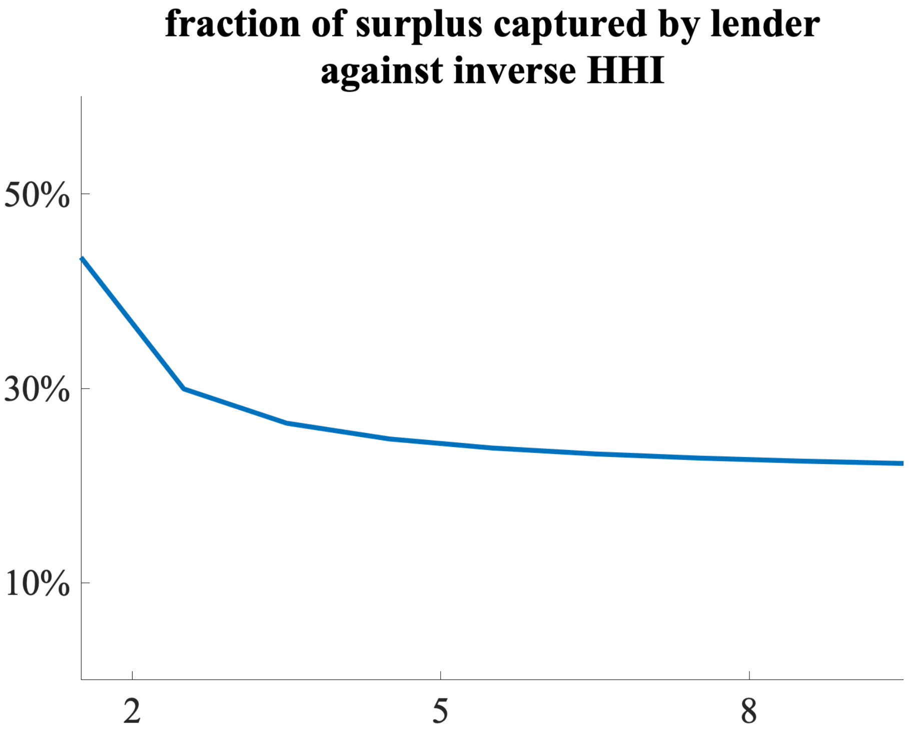


This figure plots the difference in density between the observed loan distribution and the counterfactual distribution for 8 quantiles of inverse HHI. This is our main estimation sample. We overlay the boundaries of the missing mass region in gray, which is determined by the estimated parameters as well as K.

7 Counterfactual Analysis

Our counterfactual analysis uses the predicted Laissez Faire contract distribution and estimated parameters $(\hat{\alpha}, \hat{\sigma})$ to compute the impact of several policies commonly implemented to address market power. We measure how a uniform interest cap, credit subsidy, and increase in bank competition change the distribution of contracts, and consequently the size and division of surplus between borrowers and lenders. We also analyze the welfare impact of the existing policy, the "notched" interest rate cap. In appendix D we provide the theoretical formulas used to calculate changes in welfare and surplus.

Figure 13: Fraction of Surplus captured by Lenders across *K*



7.1 Empirical Results

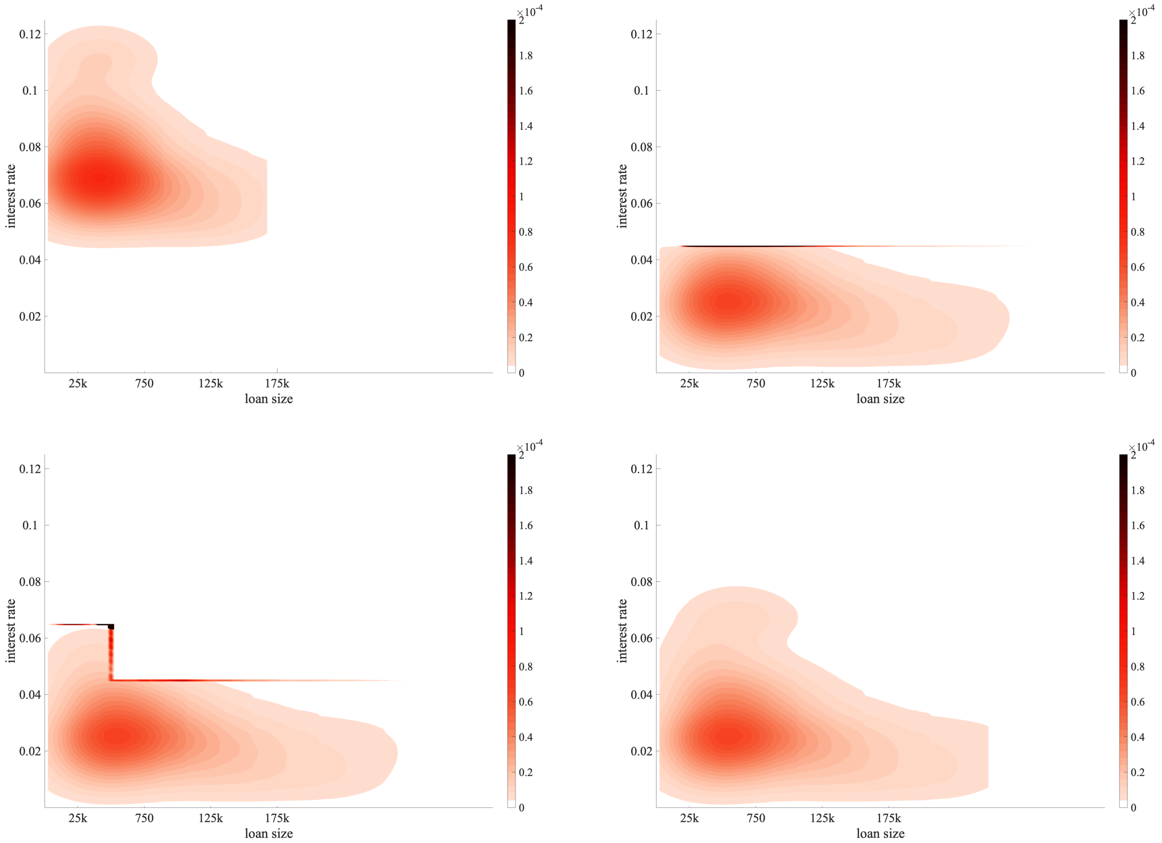
We use the counterfactual distribution constructed during the estimation procedure in Section 6 and the theory in appendix D as the basis of our counterfactual analysis. For every laissez-faire contract (r^{*}, L^{*}), we compute the counterfactual value of *r*and *L*under the following policy interventions: 1) a uniform interest rate cap of 5%, 2) the "notched" interest rate cap we find in our setting, 3) a guarantee-based subsidy to the lender that reimburses losses at a 50% rate, and 4) an increase in market competition of 20%.[4] Figure 14shows the resulting loan-interest rate distributions under each scenario.

Table 3 reports the impact of these policies both on the average values of *r*and *L*in the distribution, as well as on total surplus, lender surplus, and borrower surplus relative to the laissez-faire baseline. The table also reports the percentage of laissez-faire loans that are potentially rationed, or lost, under the counterfactual scenarios.[5] We repeat the exercise for both a concentrated market (K=1.6) and competitive market to show the non-linear policy response across markets of different sizes.

[4] For the baseline laissez-faire scenario, in which there is no government intervention, we first "remove" the impact of the 50% guarantee that exists in the data.

[5] Note that rationing would only occur under scenarios 1 and 2, which both involve a interest rate cap.

Figure 14: Distribution of (L,r) contracts under various counterfactual scenarios



These four plots show the predicted distribution of L and $r$under the laissez faire baseline (top left), an interest rate cap of 5% (top right), the existing interest rate cap “notch” stucture (bottom left), and a 50% guarantee (bottom right).

Table 3: Counterfactual Scenarios Calculated for Small and Large K Markets

K = 1.6

	LF	*Cap*	*Notch*	*Guar*	*IncK*
AvgR	8.43	4.97	6.12	4.00	7.75
AvgL	91255.00	103116.00	63625.00	138750.00	91259.00
TS/TS*	1.00	0.92	0.95	0.90	1.00
LS/LS*	1.00	0.66	0.63	1.44	0.86
BS/BS*	1.00	1.09	1.17	1.17	1.10
Rationed	0.00	0.08	0.02	0.00	0.00

K = 6.29

	LF	*Cap*	*Notch*	*Guar*	*IncK*
AvgR	8.31	4.97	6.12	4.00	8.25
AvgL	92160.00	103116.00	63625.00	138750.00	92168.00
TS/TS*	1.00	0.75	0.84	0.90	1.00
LS/LS*	1.00	0.41	0.49	1.44	0.98
BS/BS*	1.00	0.85	0.94	1.23	1.01
Rationed	0.00	0.30	0.11	0.00	0.00

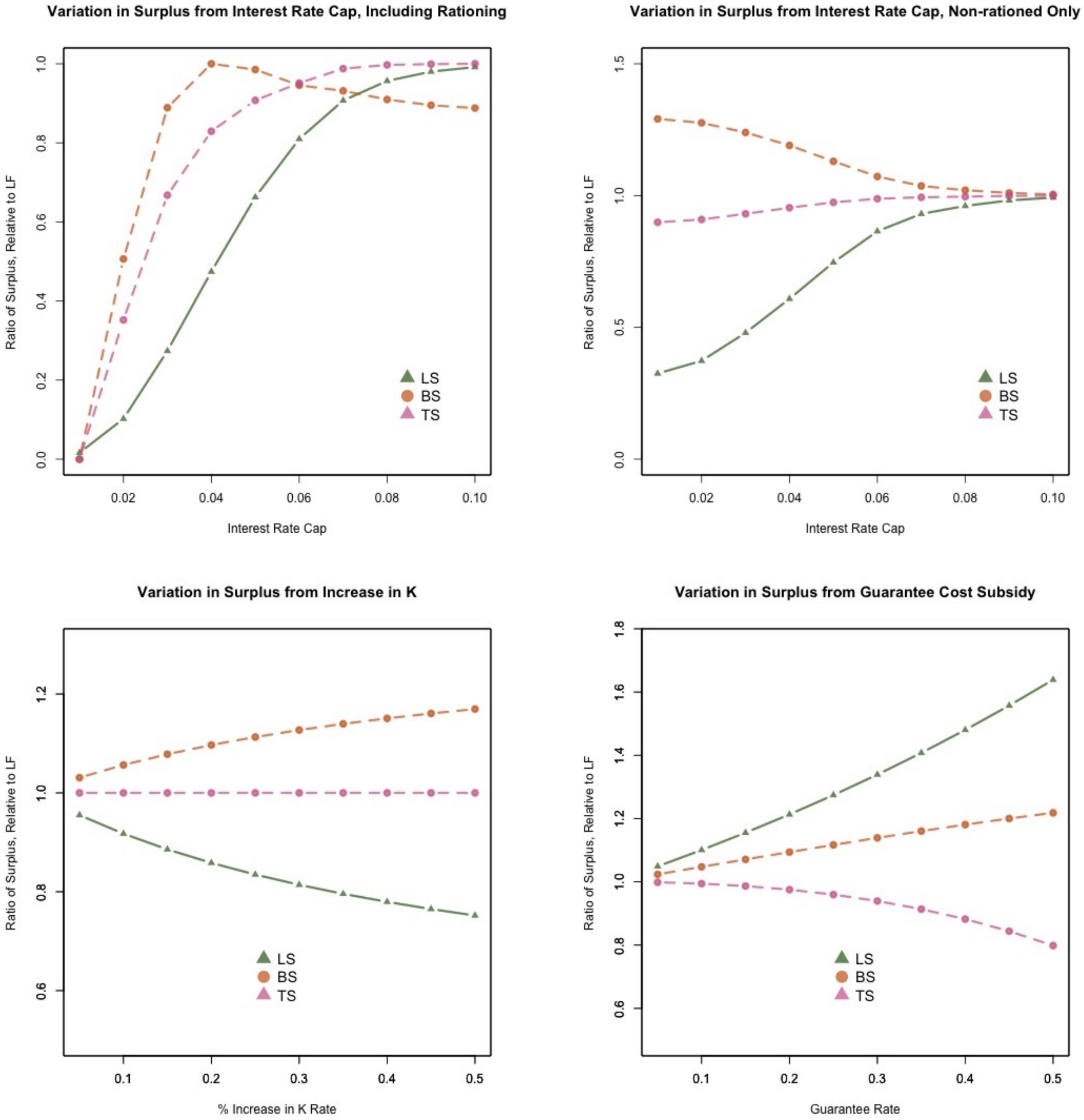
This table reports the impact of counterfactual policies both on the average values of *r*and *L*in the distribution, as well as on total surplus, lender surplus, and borrower surplus relative to the laissez-faire baseline. We analyze 1) a uniform interest rate cap of 5%, 2) the "notched" interest rate cap we find in our setting, 3) a guarantee-based subsidy to the lender that reimburses losses at a 50% rate, and 4) an increase in market competition of 20%. The table also reports the percentage of laissez-faire loans that are potentially rationed, or lost, under the counterfactual scenarios. We repeat the exercise for both a concentrated market (K=1.6) and competitive market to show the non-linear policy response across markets of different sizes.

There is some reduction in total surplus due to the distortions induced in scenarios 2-4 – loan size deviates from its efficient size under both the rate caps and the guarantee, generating inefficiencies. Total surplus remains constant when K increases (scenario 5), since increasing competition will only impact the division, but not the size, of surplus. While both interest rate cap policies increase borrower surplus for affected but *non-rationed* borrowers, they negatively impact borrowers that are rationed. In the setting where K=6.29 this rationing is so extensive that the net effect is a decrease in borrower surplus. The guarantee policy lowers costs for lenders, which in turn both increases loan size and decreases interest rates. However, the cost of the guarantee subsidy must be born by the government and therefore lowers total surplus. This subsidy is more beneficial for the lender than the borrower, since the lender does not entirely "pass-through" the reduction in costs.

The graphs in Figure 15plot proportional changes in surplus as a continuous function of the policy variables. They show the ratios of lender surplus (green), borrower surplus (orange), and total surplus (pink), relative to the Laissez Faire baseline in each counterfactual scenario for a market where originally K=1.6. Along the horizontal axis we vary the intensity of the counterfactual policy. The top two graphs show the impact of a more/less stringest interest rate cap for the entire population (left), and for only non-rationed loans (right). In the rationing case, borrower surplus initially increases as the cap is lowered due to lower prices. However, after a certain point ($\bar{r} = 4\%$)

so many loans are rationed that borrower surplus falls. When we focus only on the non-rationed population, we can see that borrower surplus increases monotonically with decreases in the cap. Total surplus falls slightly because of the distortions generated in loan size. The bottom left graph shows the impact of an increase in market competition – total surplus remains constant since loan size does not vary. However, the markup decreases in K, which causes an increase in BS and decrease in LS. The final plot, bottom right, shows the impact of a guarantee that reimburses 10-50% of lost principal to the lender. This decreases the lenders expected costs. The guarantee decreases total surplus, since it is costly to the government to provide the subsidy and loans become inefficiently large relative to the laissez-faire baseline. Both lender and borrower surplus increases – for the lenders due to a decrease in costs, and for the borrowers due to a decrease in price. However, due to market power, the increase in surplus is captured primarily by the lender. When moving from a 0 to a 50% guarantee, a $1 subsidy increases lender surplus by $0.44, and borrower surplus by only $0.30.

Figure 15: Counterfactual Scenarios Calculated for Continuum of Policy Levels, at K = 1.6



These four figures plot the ratios of lender surplus (green), borrower surplus (orange), and total surplus (pink), relative to the Laissez Faire baseline in each counterfactual scenario. Along the horizontal axis we vary the intensity of the counterfactual policy. The top two graphs show the impact of a more/less stringest interest rate cap for the entire population (left), and for only non-rationed loans (right). The bottom left graph shows the impact of an increase in market competition. The final plot, bottom right, shows the impact of a guarantee that reimburses 10-50% of lost principal.

References

Best, M. and H. Kleven (2018). Housing market responses to transaction taxes: Evidence from notches and stimulus in the uk. *Review of Economic Studies* (85), 157–193.

Dreschler, I., A. Savov, and P. Schnabl (2017). Do investment-cash flow sensitivities provide useful measures of financing constraints? *Quarterly Journal of Economics 132*(4), 1819–1876.

Kaplan, S. and L. Zingales (1997). Do investment-cash flow sensitivities provide useful measures of financing constraints? *Quarterly Journal of Economics 112*(1), 169–215.

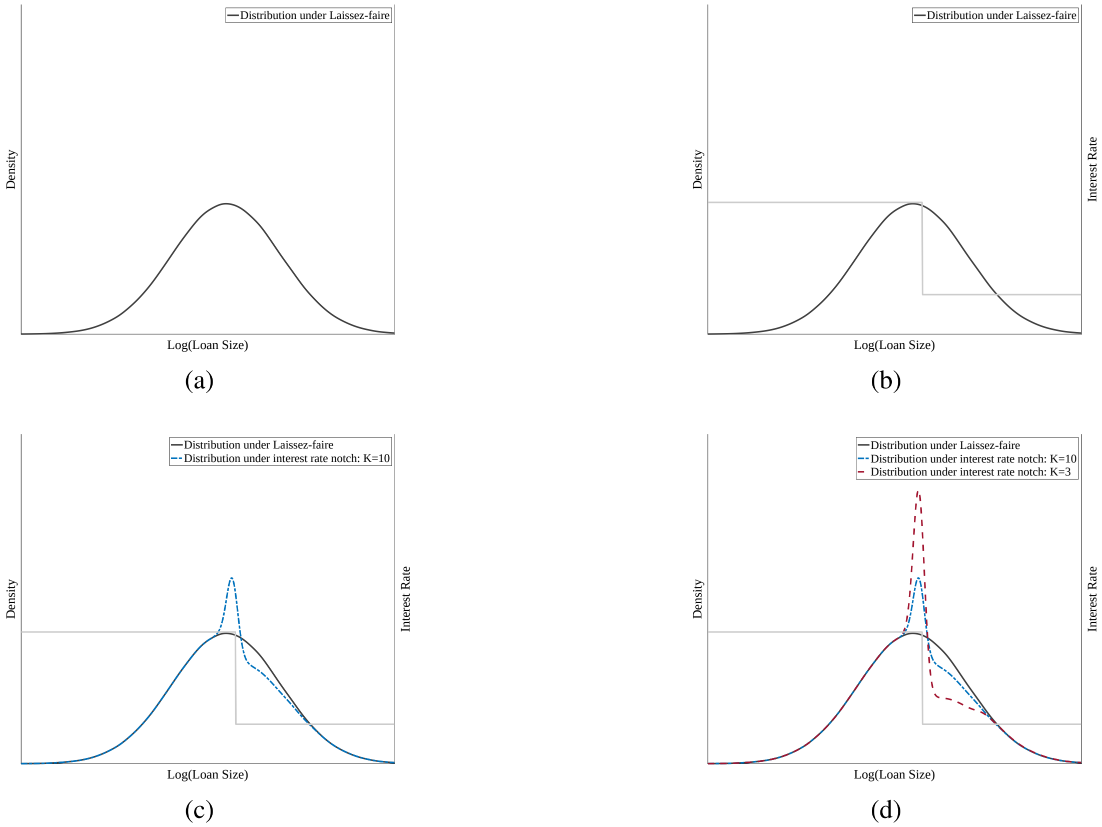
Kleven, H. (2016). Bunching. *Annual Review of Economics* (8), 435–464.

Kleven, H. J. and M. Waseem (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *Quarterly Journal of Economics 128*(2), 669–723.

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

Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy 2*(3), 180–212.

A Model Simulations



(a)

(b)

(c)

(d)

Figure 16: Model Simulations of Distributional Response to an Interest Rate Ceiling, Across Markets of Varying Concentration

B SBA Express Program

The SBA Express program was established in 1995 (under the original name FA$TTRAK) and provides a 50% loan guarantee on loans up to $350k. It is the second most popular SBA lending program, besides the 7(a) guarantee program.

The primary differences between the Express program and the SBA's flagship 7(a) Loan Program is in the maximum loan amounts, which are lower in the Express Program, the prime interest rates, which are higher in the Express program, and the SBA review time, which is typically shorter for Express loans. The documentation necessary for the SBA Express loan is less taxing compared to the standard SBA 7(a) loans, at the cost of higher interest rates.

There are two types of SBA Express loans. The first type of loans is for businesses that export goods, and the second type is for all other business. Lenders can approve a loan or line of credit up to $350,000 with an SBA Express loan. Loans can go to $500,000 if it is an Export Express Loan. The SBA Export Express loan program can help businesses that export goods get up to $500,000. The SBA will respond within 36 hours following the submission of a loan application for an Express Loan, while the eligibility review will take up to 24 hours for an Export Express Loan.

The type of loan and the type of collateral determine the amount of repayment time. The (expected) life of collateral is used to determine the repayment time: for example, using real estate for collateral is expected to lead to a longer repayment period, compared to securing a loan against equipment collateral. In particular, the maximum SBA Express loan terms are up to 25 years for real estate term loans, up to 10 years for leasehold improvement term loans, ranging between 10 and 25 years for equipment, fixtures or furniture term loans, up to 10 years for inventory or working capital term loans, and up to 7 years for revolving lines of credit.

C Additional Figures

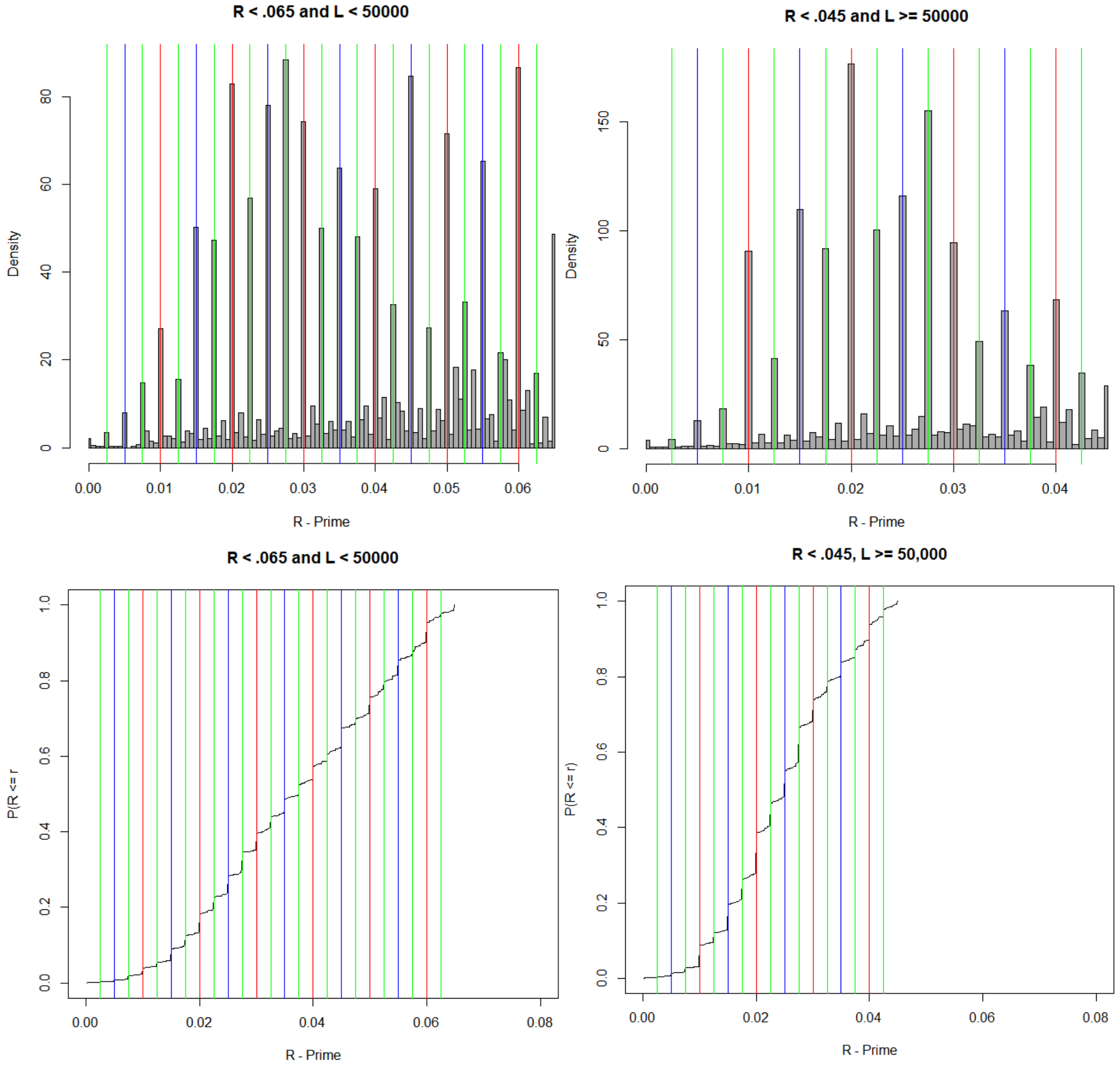


Figure 17: Observed Marginal Distribution of Interest Rates

D Counterfactual Welfare formulas

Let LS denote lender surplus, BS borrower surplus, and TS total surplus. Total surplus is solely pinned down by loan size, whereas the interest rate serves as transfer between borrower and lender. We can use the following formulas to compute proportional changes in borrower surplus and lender surplus as functions of changes in loan size (η), the interest rate (λ), as well as market structure (μ).

D.1 Uniform Interest Rate Cap and Increase in K

Let r and L denote the equilibrium interest rate and loan size under rate cap, $\eta \equiv \frac{L}{L^*}$ be the proportional distortion in equilibrium loan size, $\lambda \equiv \frac{1+r}{1+r^*}$ be the distortion in the gross interest rate, and $\mu \equiv \frac{1+(1-HHI)\alpha\sigma}{\alpha+(1-HHI)\alpha\sigma}$ denote the market structure, which depends on HHI as well as structural parameters σ and α.

We can show:

$$\frac{TS}{TS^*} = \frac{\eta^\alpha - \alpha\eta}{1-\alpha}.$$

Lender surplus is

$$\begin{aligned}
\frac{LS}{LS^*} &= \frac{(1+r-c^*)L}{(1+r^*-c^*)L^*} \\
&= \frac{(1+r-c^*)}{(1+r^*)\frac{\mu-1}{\mu}} \times \eta \\
&= \frac{\mu(1+r)-(1+r^*)}{(1+r^*)(\mu-1)} \times \eta \\
&= \eta\frac{\mu\lambda-1}{\mu-1}.
\end{aligned}$$

Borrower surplus is

$$\begin{aligned}
\frac{BS}{BS^*} &= \frac{TS-LS}{TS^*-LS^*} \\
&= \frac{TS-LS^*\left(\frac{LS}{LS^*}\right)}{TS^*\left(1-\frac{1}{1+\sigma(1-HHI)}\right)} \\
&= \frac{TS/TS^*}{\left(1-\frac{1}{1+\sigma(1-HHI)}\right)} - \frac{LS^*}{TS^*\left(1-\frac{1}{1+\sigma(1-HHI)}\right)}\left(\frac{LS}{LS^*}\right) \\
&= \frac{\eta^\alpha-\alpha\eta}{1-\alpha}\frac{1+\sigma(1-HHI)}{\sigma(1-HHI)} - \frac{\eta}{\sigma(1-HHI)}\frac{\mu\lambda-1}{\mu-1}.
\end{aligned}$$

D.2 Credit Subsidies

Under a credit subsidy, the planner subsidizes the cost of lending.

We know

$$L = \left(\frac{\alpha pz}{c-\delta}\right)^{\frac{1}{1-\alpha}}$$

Let $c' < c$ denote the new cost of lending. Doing so costs the government $(c'-c)L'$, where $'$ variables denote the new equilibrium under the subsidy. We know $L^* = \left(\frac{\alpha z}{c^*}\right)^{\frac{1}{1-\alpha}}$ and $L' = \left(\frac{\alpha z}{c'}\right)^{\frac{1}{1-\alpha}}$

hence

$$\frac{L'}{L^*} = \left(\frac{c}{c'}\right)^{\frac{1}{1-\alpha}}$$

The constant markup implies that:

$$\frac{R^{'}}{R^*} = \frac{c'}{c}$$

Hence

$$\frac{LS'}{LS^*} = \frac{(R'-c')\,L'}{(R^*-c)\,L^*} = \left(\frac{c}{c'}\right)^{\frac{\alpha}{1-\alpha}}$$

We know

$$\begin{aligned} LS^* + BS^* &= z\,(L^*)^{\alpha} - cL^* \\ &= z\left(\frac{\alpha z}{c}\right)^{\frac{\alpha}{1-\alpha}} - c\left(\frac{\alpha z}{c}\right)^{\frac{1}{1-\alpha}} \\ &= (1-\alpha)\,z^{\frac{1}{1-\alpha}}\,c^{-\frac{\alpha}{1-\alpha}} \end{aligned}$$

Likewise $LS' + BS' = (1-\alpha)\,z^{\frac{1}{1-\alpha}}\,(c')^{-\frac{\alpha}{1-\alpha}}$. Hence

$$\frac{LS' + BS'}{LS^* + BS^*} = \left(\frac{c}{c'}\right)^{\frac{\alpha}{1-\alpha}}$$

Borrower surplus also changes accordingly:

$$\frac{BS'}{BS^*} = \left(\frac{c}{c'}\right)^{\frac{\alpha}{1-\alpha}}.$$