The Role of Regulation and Competition in Credit Allocation: Evidence from Small Business Lending*

Panagiotis Avramidis[†] Konstantinos Serfes[‡] Kejia Wu[§]

May 7, 2021[¶]

Abstract

Do policies that promote credit access have an impact on targeted borrowers? To address this question, we develop a theoretical model of information production, regulation and bank competition and test its predictions using the Community Reinvestment Act's small businesses lending program. The analysis reveals that, on average, regulation-induced surge of loans leads to an improvement in the credit score of small businesses. However, the effect vanishes in markets with high demand-adjusted bank competition, due to weaker bank incentives for information production. In order to improve credit allocation across all local markets, interventions should also enact policies that lower information costs.

Keywords: Competition; Regulation; Information production; Small business loans.

JEL classifications: G14; G21; L13; L50.

*We would like to thank Giovanni Compiani (discussant), Ishan Ghose, Robert M. Hunt and seminar participants at the 2021 IIOC and the Federal Reserve Bank of Philadelphia for useful comments and suggestions. Use of the National Establishment Time Series data was supported by the National Institute of Aging (grants 1R01AG049970, 3R01AG049970-04S1), Commonwealth Universal Research Enhancement (C.U.R.E) program funded by the Pennsylvania Department of Health - 2015 Formula award - SAP #4100072543, the Urban Health Collaborative at Drexel University, and the Built Environment and Health Research Group at Columbia University.

[†]Alba Graduate Business School, The American College of Greece, Athens, Greece. pavramidis@alba.acg.edu.

[‡]School of Economics, LeBow College of Business, Drexel University, Philadelphia PA, USA. ks346@drexel.edu.

[§]Federal Reserve Bank of Philadelphia, Philadelphia PA, USA and School of Economics, LeBow College of Business, Drexel University. kejia.wu@phil.frb.org.

[¶]The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia, the Federal Reserve Board of Governors, or the Federal Reserve System.

1 Introduction

Financial intermediaries play an important role in the economy by producing information necessary for the efficient allocation of credit (e.g., Diamond (1984) and Ramakrishnan and Thakor (1984)). Moreover, the effort banks need to exert to produce information varies across borrowers.[1] Banks seeking to minimize their costs will eschew borrowers with high implied screening costs giving rise to underfunded local markets.[2] The economic prospects of these local markets are consequently undermined.[3] To contain the problem of credit rationing and the formation of unbanked communities, policymakers have introduced programs that aim to propagate credit supply to targeted markets.[4] However, the impact on borrowers, targeted by these important initiatives, remains unexplored.

This study tries to shed light on the above market failure and the consequences of policy interventions on borrowers by developing a theoretical model and testing its predictions empirically. We show that the co-existence of heterogeneous entrepreneurs, with respect to the availability of information about their projects, means that some worthy projects remain unfunded. Imposing a regulation of a minimum level of credit supply improves efficiency under certain conditions. Specifically, banks identify and fund the worthy projects, improving the creditworthiness of targeted borrowers in local markets where the expected rents earned by banks exceed the screening costs. In local markets with insufficient expected rents, banks

[1] According to the 2017 Report to the Congress on the Availability of Credit to Small Businesses, prepared by the Board of Governors of the Federal Reserve System: *Lending to small businesses is further complicated by the "informational opacity" of many such firms. Obtaining reliable information on the creditworthiness of a small business is often difficult because little, if any, public information exists about the performance of most small businesses.*

[2] According to the same report as in footnote 1: *the relatively elevated costs of evaluating small business loan applications and the ongoing costs of monitoring firm performance have made loans to small businesses less attractive for some lenders, especially because, when expressed as a percentage of the (small) dollar amount of the proposed loan, these non interest costs are often quite high compared with loans to middle-market or large corporate borrowers.*

[3] Small businesses contribute to local economies by bringing growth and innovation to the community in which the business is established. According to the 2019 Small Business Profile Report by the Office of Advocacy, there are 30.7 million small businesses in the US accounting for 99.9% of total businesses and 47.3% of total US employment. Importantly, 37% of high-tech workers work for small businesses. Small businesses also provide opportunities for many people, including women and minorities, to achieve financial success and independence and they complement the economic activity of large organizations by providing them with components, services, and distribution of their products.

[4] Several long-standing government initiatives exist to support credit access for small businesses. Two such initiatives are of particular importance. The first is the Community Reinvestment Act (CRA) enacted in 1977 to encourage federally insured depository institutions to help meet the credit needs of their local communities, particularly low- and moderate-income neighborhoods. The second involves the loan programs sponsored by the U.S. Small Business Administration (SBA). The SBA provides financing to young and growing small firms through several channels such as the 7(a) Loan Program and SBA 504 Certified Development Companies (CDCs). More recently, the Federal Reserve established the Main Street Lending Program to support lending to small and medium-sized organizations before the onset of the COVID-19 pandemic.

extend credit indiscriminately to worthy and unworthy projects, offering no improvement in the average creditworthiness of targeted borrowers. We conclude that lowering screening technology costs is pivotal in enhancing the creditworthiness of targeted borrowers across all local markets.

The theoretical model consists of two main building blocks. The first is a model of bargaining between a bank and an entrepreneur about the division of expected surplus. The second is a search with frictions model linking outside options in the bargaining to the number of banks relative to entrepreneurs in the market, i.e., the demand-adjusted bank competition. Entrepreneurs are divided into four groups based on two criteria: the level of information availability and the project value. Specifically, we distinguish entrepreneurs between transparent and opaque and within each group we assume there are entrepreneurs with high and low value projects. Banks can identify the value of a transparent entrepreneur with no additional effort. For informationally opaque entrepreneurs, banks must invest in information production in order to identify the high value projects. Funding high value projects improves market efficiency, whereas funding low value projects is inefficient.

Expected equilibrium rents, i.e., the difference in expected equilibrium profits when a bank produces information and when it does not, is the main incentive for banks to acquire costly information (Grossman and Stiglitz (1980)). For a wide range of the screening technology cost parameter, there exists an equilibrium where banks do not invest in information production and only offer credit to the high value projects of the transparent entrepreneurs. This equilibrium is inefficient as banks forego funding the high value projects among the opaque entrepreneurs. A regulation forces banks to extend credit to the group of opaque entrepreneurs. The banks' incentives to acquire information are stronger when the regulation is in force, because if a bank does not invest in screening technology it funds all opaque entrepreneurs who on average have unprofitable projects.

Moreover, the expected equilibrium rents are a decreasing function of demand-adjusted bank competition, whether the regulation is in force or not. Therefore, in local markets with more entrepreneurs for each competing bank, the regulation induces banks to switch to an equilibrium where each bank invests in the screening technology. Banks extend credit to the opaque entrepreneurs with high value projects, improving the borrowers' creditworthiness. However, in local markets with fewer entrepreneurs for each competing bank, the expected rents fall short of the implied screening costs and the equilibrium continues to be the one with no information production. Banks, forced by the regulation and the lack of investment in screening capability, are equally likely to offer credit to entrepreneurs with low and high value projects, yielding no improvement in the average creditworthiness of borrowers.

The empirical evidence of the model's predictions is drawn from the Community Reinvestment Act (CRA), the US regulation aimed to encourage credit supply in underfunded local communities. First, we show that the regulatory eligibility of a community prompted banks to increase the number of loans to small businesses in the community. Second, our analysis reveals that, on average, credit was allocated efficiently leading to an improvement in the credit score of businesses. Since a higher credit score of a firm is linked to better loan terms, like lower interest rates, and to more favorable trade credit agreements, we conclude that the reported gains in efficiency may engender further economic improvement for the local market. However, and consistent with the predictions of the theoretical model, we observe no such improvement in those communities where demand-adjusted bank competition is high. Finally, we show that those communities with higher proportions of relatively more opaque borrowers are less likely to experience an increase in small businesses credit score, especially if the demand-adjusted bank competition is intense.

By performing a large-scale geocoding of National Establishment of Time Series (NETS) and Summary of Deposit (SOD) data, we are able to append to CRA loan data, small business and bank branch information aggregated at the census tract level. This unique dataset allows us to examine the impact of CRA lending on local small businesses in conjunction with bank competition which has been largely lacking from existing empirical studies. Specifically, CRA examiners divide tracts mainly into two groups, the upper-middle tracts with median family income (MFI) above 80% of the surrounding Metropolitan area's MFI and the low-moderate tracts with MFI ratio below 80%. Banks with branches in low-moderate tracts, the CRA eligible communities, are subject to examination in meeting the credit needs of these communities.

We exploit the fact that CRA eligibility is allocated on the basis of the MFI ratio to use the regression-discontinuity (RD) design. The underlying idea of our RD design is that the probability of CRA eligibility, conditional on the "running" variable of MFI ratio, jumps discontinuously at the cutoff, inducing heterogeneity in regulatory treatment assignment that is unrelated to potential confounders. We focus on the 2010 census which was incorporated to CRA examination in 2012 and we examine if firms located within the eligible tracts display higher credit scores compared to firms located within the ineligible tracts. Because of its local nature, we choose to estimate the RD effect of regulation eligibility using the local polynomial approximation. We confirm our findings using a variety of model parameters like kernel functions, bandwidth selection methods and the order of polynomial approximations.

Our research contributes to various streams of literature on banking which are discussed in more detail in the following section. Specifically, the stream of the extant literature that examines the effectiveness of government programs and policies that target underfunded communities, has limited its focus on whether these programs have succeeded in boosting the supply of credit to targeted borrowers in these markets. However, there has been little progress in examining whether the flow of credit to these communities benefits the local businesses. This is an important question for the evaluation of the economic efficacy of a government initiative that promotes credit supply to targeted borrowers. We try to fill this gap by introducing a theoretical model and providing empirical evidence that examines the efficiency of the CRA regulation.

Furthermore, our theoretical model's novelty compared to existing models of information acquisition and competition, e.g., Hauswald and Marquez (2006), is the introduction of heterogeneous borrowers with respect to their informational opacity. Specifically, the co-existence of opaque and transparent entrepreneurs with high value projects allows our model to explain why some markets remain inefficiently underfunded despite the efforts of policymakers to increase credit availability. Another novelty of the proposed model is the introduction of costly search and bargaining to banking literature, both realistic features in the context of small business lending. The search model also allows us to explicitly account for the number of banks and entrepreneurs in constructing the measure of demand-adjusted bank competition. Finally, in contrast to previous studies that resort to mixed strategy equilibria, our model yields an interest rate equilibrium in pure strategies. This differentiation is important because in practice lenders appear to prefer deterministic pricing rules, which implies that identical borrowers receive identical prices, to mixed strategies where banks randomize their pricing offers to identical borrowers.[5]

One implication of our findings is that bank competition does not lead inexorably to extension of credit to all entrepreneurs with high value projects. Banks will seek to minimize information production costs, by funding local markets comprised of borrowers of higher transparency due to information spillovers from past lending. In contrast, local markets that are deprived of these information spillovers can be drawn into a vicious cycle of low funding and high proportions of opaque entrepreneurs.

[5]Interest rate randomization would also imply that lenders would almost surely violate fair lending laws and regulations.

From a social-welfare perspective, this market imperfection entails serious economic consequences. A large number of communities remain underfunded despite the prolonged accommodative monetary policy and the recent growth in financial markets and bank assets. Moreover, small businesses see the supply of bank credit to be diverted to large corporations despite the introduction of several lending programs. For example, according to recent reports, the recent jump in corporate lending was almost entirely accounted for by big companies drawing down pre-existing credit lines, while smaller and midsized ones actually saw reduced use of credit lines.[6] The researchers point out that it is understandable that smaller businesses get poorer access to credit, given the paucity of timely, comprehensive and reliable financial information available to lenders.

These trends are unlikely to be reversed unless policymakers intervene to break the vicious cycle of low funding and shortfall of information externalities. However, our analysis shows explicitly that blanket policies promoting credit supply may produce mixed results. Banks will invest in information production to identify the high value projects of the opaque entrepreneurs only if the expected rents are sufficient. In local markets with high demand-adjusted bank competition, banks will extend credit to both high and low value projects offering no creditworthiness improvement to the business community.

The strategic role of information production that we identify has further implications for the role of screening technology in promoting efficient credit allocation. The empirical findings confirm the theoretical predictions under moderate information technology costs. According to our model, improvements in screening technology and information dissemination would render information production as the banks' equilibrium strategy independent of the local market characteristics.

These results suggest that, from a policy perspective, investing in information infrastructures may have a socially beneficial effect. Specifically, screening technologies help to curtail regulation-induced wasteful loans extended to entrepreneurs with low value projects. Furthermore, regulatory initiatives like CRA combined with screening technologies could engender a virtuous cycle of credit supply and positive information externalities that will eliminate the growing phenomenon of underserved communities.

The article is organized as follows. Section 2 offers a literature review. Section 3 presents a model of financial intermediation and information production. Section 4 derives the equilibrium for unregulated and

[6] Gabriel Chodorow-Reich, Harry Cooperman, Olivier Darmouni, Stephan Luck, and Matthew Plosser, "Weathering the Storm: Who Can Access Credit in a Pandemic?", Federal Reserve Bank of New York, Liberty Street Economics, Oct. 13, 2020. Article.

regulated markets. Empirical investigation is discussed in section 5, while section 6 presents the findings. Section 7 concludes. Some proofs are relegated to Appendix A and the remaining proofs, which are not essential for the understanding of the main derivations and results, are relegated to Appendix B, called Internet Appendix.

2 Literature review

Our paper joins a large body of literature on information production, bank competition and credit supply. Sharpe (1990) and Rajan (1992) study borrowing under adverse selection and how relationship building affects competition in the refinancing stage, while Broecker (1990) investigates the potentially negative effects of competition on loan markets under independent loan screening. In general, greater competition in credit markets leads to lower loan rates and enhanced credit access for marginalized borrowers (Allen and Gale (2004)). It may also encourage excess risk-taking of banks driven by the reduced profits (Hellmann, Murdock, and Stiglitz (2000), Martinez-Miera and Repullo (2010)) although moral hazard may actually lead to the opposite direction, causing banks to become more risky as their markets become more concentrated (Boyd and De Nicolo (2005)). Studies focusing on the interaction of information asymmetry with competition argue that greater competition may intensify the information asymmetry problem in local markets due to more dispersed borrower information (Marquez (2002), Dell'Ariccia and Marquez (2004)).

Ruckes (2004) and Hauswald and Marquez (2006) extended the literature of informational asymmetries and competition by endogenizing the generation of information. Specifically, in Ruckes (2004) lenders can improve the accuracy of the signal at a cost, but the level of competition plays no role, since the model only allows for two competing lenders. Hauswald and Marquez (2006) develop a theoretical model of strategic information acquisition by geographically differentiated financial intermediaries. In examining an unregulated market, they come to a similar conclusion to ours, namely, that investment in information falls with competition. Their main focus is on how the interaction of competition and lender-borrower relationship affects the incentives for information production. In contrast, our goal is to study the effect of the interaction between competition and regulation on the information production incentives. Our model is equipped to handle the introduction of regulation by avoiding all the technical complications that arise from mixed strategy equilibria and yields clean and simple testable predictions. A trade-off is that our model ignores any lender-borrower specific relationships, which are nevertheless beyond the scope of our research, since our focus is on underserved firms in low to medium income communities that are less likely to have formed strong relationships with lenders.

Our paper also adds to the broad literature on small business lending. Berger, Saunders, Scalise, and Udell (1998) report that changes in competition due to bank mergers lowers credit availability to small businesses in the short term. They also conclude that the merger-induced short-term distortions in the loan market are gradually eliminated by the entry of new banks and increased loan supply by existing banks. Cetorelli and Strahan (2006) provide evidence that local bank competition is associated with higher numbers of small businesses. They suggest that enhanced credit access in competitive bank markets lowers the barriers to entry for new firms. Moreover, by promoting bank competition, deregulation of branching has raised the rate of new business incorporation (Black and Strahan (2002)), has lowered the interest rates charged to small firms and helped banks to expand the supply of capital (Rice and Strahan (2010)). Although competition and deregulation have enhanced and broadened access to credit, there are specific groups of borrowers that remain unbanked or underbanked. The findings of our study contribute to small business lending and competition literature by offering evidence that government intervention may alleviate the obstacles to credit access and the improvement depends on the underlying bank competition.

The largest body of the literature on small business loans and competition focuses on lending relationship and how it ameliorates information production. Compared to large corporations, small businesses lack the public information or audited financial statements (hard information) that convey firm's quality to lenders, thus banks often depend on collection of private information (soft information) through their relationship with the borrower (Petersen and Rajan (1995)). Given that it is costly to acquire private information, small firms rely heavily on local banks in financing their operations (Petersen and Rajan (2002)). Lending distance, hence, is often obtained as a measure of soft information collection or strength of banking relationship (Agarwal and Hauswald (2010)). Notwithstanding the important role of lending relationship in reducing the obstacle of opacity, especially for small businesses, lending relationship premises some history of banking activity which does not exist for some groups of borrowers that remain unbanked or underbanked. Our findings show that government intervention may be necessary to break the vicious circle of unfunded communities and the paucity of information spillovers.

The information opacity of small businesses means that credit screening technologies are also important to banks. Hauswald and Marquez (2003) examine how progress in screening technologies affects competition and bank profitability. Levine, Lin, Peng, and Xie (2020) show how technology factors that enhance communication and reduce screening cost may lead to small-firm lending growth. Cole, Goldberg, and White (2004) find evidence that large banks make lending decisions mainly based on firms' financial statements, whereas small banks resort to a greater extent on information about the character of the borrower. The underlying assumption is that large banks have a comparative advantage on credit screening based on hard quantitative information, while small banks compete with large banks through acquisition of soft private information. Later studies challenge this view and argue that oversimplified classification of credit screening technologies (i.e. hard vs soft evaluation) may be misleading (Berger and Udell (2006), Berger and Black (2011)). Our study extends this literature by highlighting theoretically the important role of screening technology in the efficacy of government programs promoting access to credit.

Finally, this paper relates to a more niche literature examining the impact of the Community Reinvestment Act (CRA). Existing studies have focused on assessing the effects of the CRA regulation on the supply of residential mortgage loans (Dahl, Evanoff, and Spivey (2010), Bhutta (2011), Agarwal, Benmelech, Bergman, and Seru (2012), Avery and Brevoort (2015) and Ding and Nakamura (2020)) and the quality of financial services for mortgage-related products, Begley and Purnanandam (2021). Fewer studies focus on the effect of CRA on small business lending to low- and medium- income (LMI) communities. Bostic and Lee (2017) show that CRA-eligible census tracts attracted more small business lending activities than CRA-ineligible tracts but the banks' response was also driven by macroeconomic market conditions. Ding, Lee, and Bostic (2018) explore a policy shock that changes CRA-eligibility of census tracts and find that gaining (losing) CRA-eligibility lead to increase (decrease) in small business lending to LMI communities. The literature on the examination of the CRA regulation on other lending outcomes, such as efficiency or pricing, under the CRA regulation is scant. Avery, Bostic, and Canner (2005) evaluate survey responses from CRA-related lending and find that accomplishing CRA goals has exacted a price in terms of lending profit but the results are mixed at best. The empirical evidence on CRA lending beyond the volume of lending is really lacking, mostly due to the aggregated nature of CRA lending data. Our research goes a step further and examines whether the new loans, due to the CRA regulation, have positively impacted the targeted small businesses.

3 The theoretical model

We develop a dynamic equilibrium model, focusing on bank decisions about information production and the effect of bank competition and regulation on these decisions. Entrepreneurs are heterogeneous with respect to the value of their projects and how transparent this value is to the lender. The first building block of our theory is a search with frictions model where entrepreneurs match with banks. Search frictions, which is a natural assumption in this market, are meant to capture the cost of a deal delay if a bank and an entrepreneur do not reach an agreement and each one searches again for a new deal. In the second building block of our theory a bank and an entrepreneur, after they match, bargain over the financing terms. Bargaining is also a natural assumption in this context, where usually there are no posted 'prices' and how the surplus is shared depends on the parties' outside options. The outside options, in turn, depend on the level of demand-adjusted bank competition, i.e., the number of banks relative to the number of entrepreneurs. This affects the equilibrium terms of financing but more importantly it affects the banks' incentives for information production in order to identify the high value projects among the informationally opaque entrepreneurs. A regulation, that forces banks to extend loans to underfunded entrepreneurs, may affect these incentives differentially in low and high demand-adjusted bank competition markets.

3.1 Entrepreneurs, transparency and value of projects

Each entrepreneur has an investment project that generates a terminal cash flow X and requires one unit of capital the entrepreneur must borrow from a bank. This cash flow X can be an amount R with probability p, or 0 with probability $1-p$, where $p \in \{\underline{p}, \overline{p}\}$ with $\overline{p} > \underline{p}$ (see Sharpe (1990) and Hauswald and Marquez (2003)). Final cash flows are observable and contractible, but the project type is initially unknown to either borrower or lender. Let $H \equiv \overline{p}R$ be the high expected value project and $L \equiv \underline{p}R$ be the low expected value project. For simplicity we set $\underline{p} = 0$, so $L = 0$.

Each entrepreneur can be one of the following four types. There are two groups in terms of how transparent the expected value of the project is: *informationally Transparent* (T, henceforth) and *informationally Opaque* (O, henceforth) and within each group there are entrepreneurs with high expected value, $v = H > 0$, and low expected value, $v = L = 0$, projects. A bank can identify the type of a T entrepreneur's project without additional effort, while for an O entrepreneur it needs to invest in costly screening technology to screen out the low expected value projects.

When a high value project gets funded, the entrepreneur's creditworthiness increases. When a low expected value project receives funding, the creditworthiness of the entrepreneur decreases. This is because the entrepreneur will default.

3.2 Search and matching

The market consists of N_b banks (b) and N_e entrepreneurs (e). Each bank has one unit of capital, which means it can finance only one entrepreneur. All the results extend to the case where each bank can finance a finite number of entrepreneurs. Entrepreneurs and banks meet bilaterally and at random in continuous time with discount factor r.[7] Let $\theta \equiv \frac{N_b}{N_e}$ measure the demand-adjusted bank competition. We do not consider endogenous entry on either side of the market.[8] The measure of deals, or matches, per unit of time is given by the matching function $x(N_b, N_e)$, which we assume is increasing, concave and exhibits constant returns to scale. From the perspective of a bank the (Poisson) arrival rate of a deal is $\alpha_b(\theta) \equiv \frac{x(N_b,N_e)}{N_b}$; while the entrepreneur's deal arrival rate is $\alpha_e(\theta) \equiv \frac{x(N_b,N_e)}{N_e}$. As θ increases, i.e., more banks relative to entrepreneurs, $\alpha_b(\theta)$ decreases, while $\alpha_e(\theta)$ increases, with $\lim_{\theta\to 0} \alpha_b(\theta) = \lim_{\theta\to\infty} \alpha_e(\theta) = \infty$ and $\lim_{\theta\to\infty} \alpha_b(\theta) = \lim_{\theta\to 0} \alpha_e(\theta) = 0$.[9]

Whether an entrepreneur is transparent or opaque and the type of his project are match-specific. In particular, after a bank matches with an entrepreneur, there is a random and independent draw about the transparency of the entrepreneur and then there is another random and independent draw about the type of his project.[10] The probability of a T entrepreneur is q and the probability an O entrepreneur is $1 - q$. Furthermore, the probability of a high expected value project, conditional on the entrepreneur being in the T or O group, is given by $\Pr(v = H|T) = \nu$ and $\Pr(v = H|O) = \mu$, respectively. Banks will never fund a

[7]The search and matching part of our model follows closely Inderst and Müller (2004). In the search literature, our framework is commonly known as the Diamond-Mortensen-Pissarides model, Pissarides (2000).

[8]Given the short-run nature of our empirical exercise, this is a reasonable assumption.

[9]Consider, for example, a Cobb-Douglas matching technology $x(N_b, N_e) = \xi(N_b N_e)^{0.5}$, where $\xi > 0$ is a parameter. Then, the arrival rates are $\alpha_b(\theta) = \xi\theta^{-0.5}$ and $\alpha_e(\theta) = \xi\theta^{0.5}$.

[10]We can introduce positive correlation between any two successive draws from different banks for the same entrepreneur without affecting the results qualitatively.

low value project of the T group, since they can easily screen the low value projects of the T entrepreneurs out. Hence, we can set, without any loss of generality, $\nu = 1$.[11]

Match-specific heterogeneity can be justified on the grounds that banks may differ in terms of their (unobservable to the entrepreneurs) expertise (specialization) corresponding to an entrepreneur's particular project. Hence, an entrepreneur may be T for one bank but O for another. Likewise, an entrepreneur's project may be of high expected value for one bank but low for another, due to different assessments of the project's success probability.[12] Nevertheless, the corresponding probabilities dictate the relative size of each group of entrepreneurs. This allows us to sidestep unnecessary complications that would arise either from private information or from ex-ante heterogeneity, while the model retains the most relevant aspects of heterogeneity for the issues we intend to study.[13]

The cost of funds for the bank is $\bar{i} \geq 0$. We make the following assumption.

Assumption 1 $\mu H < 1 + \bar{i} < H$.

The above assumption implies that high expected value projects generate positive surplus, while pooling high and low expected value projects together, i.e., funding an O entrepreneur without screening, generates a negative expected surplus, since $L = 0$. In this case banks do not fund entrepreneurs, unless they are forced to by a regulator.

3.3 Bank information

The bank, after it matches with an entrepreneur and before they engage in bargaining, draws a signal s, which can be either high, h, or low, ℓ, if the entrepreneur belongs in the O group. Whether the entrepreneur is in the T or the O group and the signal realization become common knowledge between the bank and the entrepreneur. The signal's distribution function, conditional that $s = h, \ell$, is given by

$$\Pr(s = h|H) = \Pr(s = \ell|L) = \frac{1+\phi}{2} \text{ and } \Pr(s = h|L) = \Pr(s = \ell|H) = \frac{1-\phi}{2},$$

[11]If $\nu < 1$, in all the expressions that follow q should be replaced by $q\nu$, with no additional changes.

[12]For instance, lenders have to assess the idiosyncratic quality of "soft" capital in a business and this is a process that is both subject to error and cannot be subject to scale economies.

[13]Silveira and Wright (2016) make a similar assumption in a search model for the venture capital market. In particular, the fixed cost of getting the project off the ground and the return on the project are drawn from a distribution after a venture capitalist matches with an entrepreneur. See also Rogerson, Shimer, and Wright (2005), section 4.4, for a survey of search-theoretic models with match-specific heterogeneity.

where $\phi \in \{0, 1\}$ measures the accuracy of the signal. If $\phi = 0$, the signal is completely uninformative. On the other hand, if $\phi = 1$ the signal is perfect. Each bank can improve the precision of its own signal, from $\phi = 0$ to $\phi = 1$, if it incurs an exogenously given cost $c > 0$ per borrower it bargains with. We assume that the fixed cost of the technology is zero. We can think of this as an investment in a 'screening technology' the bank makes before it commences its search. It can also be interpreted as a costly effort the bank exerts to screen the high from the low value projects in the O group. Using Bayes' rule, we derive the following conditional probabilities as functions of the signal precision ϕ

$$\Pr(H|h) = \frac{(1+\phi)\mu}{(1+2\mu\phi-\phi)} \text{ and } \Pr(L|\ell) = \frac{(1+\phi)(1-\mu)}{(1-2\mu\phi+\phi)}. \tag{3.1}$$

3.4 Deal values

The interest rate a bank charges is i.[14] Let B^d and E^d denote the expected profit of the bank and the entrepreneur respectively when they strike a deal (d). For simplicity, we assume a transferrable utility (TU) environment where the interest rate i is the instrument that transfers surplus between a bank and an entrepreneur. When only high value (H) projects are funded, we denote the values of the deal as follows

$$B_H^d = i - \bar{i} \quad \text{and} \quad E_H^d = H - (1+i). \tag{3.2}$$

When banks fund entrepreneurs in the O group and cannot separate the high value from the low value projects the values of the deal are

$$B_O^d = \mu(1+i) - (1+\bar{i}) \quad \text{and} \quad E_O^d = \mu(H - (1+i)), \tag{3.3}$$

where, from Assumption 1, $B_O^d + E_O^d < 0$.

We assume that entrepreneurs' alternative to bank funding yields expected profit equal to zero. So, the lowest deal expected profit for an entrepreneur must be zero to satisfy the entrepreneur's participation constraint.

[14]There will be different interest rates depending on a bank's screening ability and on whether a bank chooses an equilibrium rate or deviates from the equilibrium. We do not make all these distinctions here, but they will become clear as we proceed with the analysis.

3.5 Bargaining

If the search is successful, the two sides bargain. The outside options in the bargaining, B^o and E^o, are the utilities from going back into the market and searching anew. Our bargaining concept is the Nash bargaining solution.[15] Bargaining takes place after the value of the project has become common knowledge, for a T entrepreneur, or after the signal has become common knowledge, for an O entrepreneur. The two parties bargain over the interest rate i to maximize

$$\left(B^d - B^o\right)^{\frac{1}{2}} \left(E^d - E^o\right)^{\frac{1}{2}}. \tag{3.4}$$

Taking the derivative of (3.4) with respect to E^d, noting that the derivative of B^d with respect to E^d is -1 due to the TU assumption, and re-arranging we obtain

$$E^d - E^o = B^d - B^o. \tag{3.5}$$

3.6 Outside options

An important feature of the model are the outside options, as they succinctly reflect the demand-adjusted bank competition. In equilibrium, there are two symmetric cases with respect to information: no bank has the ability to screen projects in the O group, $\phi = 0$, or all banks have this ability, $\phi = 1$.[16] In the latter case, only high value projects will be funded from both the T and the O groups. In the former case, under no regulation banks only fund T entrepreneurs (due to Assumption 1), while under regulation they must fund O entrepreneurs as well, and given their inability to screen they extend loans to entrepreneurs with high and low value projects. When a bank deviates in ϕ, then the outside option of the deviating bank is affected. We account for this when we solve for the equilibrium in Section 4.

[15]The axiomatic Nash bargaining solution can be derived as the limit of a non-cooperative bargaining game where the two parties bargain with an open time horizon under the risk of breakdown (Binmore, Rubinstein, and Wolinsky (1986)). Our qualitative results do not depend on the specifics of the Nash bargaining solution. Furthermore, for simplicity we assume equal bargaining powers, but we could have allowed for general bargaining powers without any effect on our results.

[16]There can also be an asymmetric outcome where some banks acquire information and some do not.

Given that the market is stationary, the expected profit from going back into the market equals the expected profit from entering the market in the first place. Hence, B^o and E^o represent both the outside options in the bargaining as well as the overall profits from searching.

We begin with the bank's outside option B^o. Assume $\phi = 0$ and that banks must fund all projects. Consider the value of not having a deal with an entrepreneur for a short period Δt. The flow benefit is zero. The Poisson arrival rate of a deal for the bank is $\alpha_b(\theta)$. At the end of $t + \Delta t$ the bank either continues with no deal with probability $e^{-\alpha_b(\theta)\Delta t}$, or strikes a deal with an entrepreneur with probability $1 - e^{-\alpha_b(\theta)\Delta t}$. With probability q the entrepreneur is in the T group and the value of the deal is B_H^d and with probability $(1-q)$ the entrepreneur is in the O group and the value of the deal is B_O^d. Thus, the value of the outside option is expressed as follows

$$B^o = e^{-r\Delta t}\left(e^{-\alpha_b(\theta)\Delta t}B^o + \left(1 - e^{-(\alpha_b(\theta)q)\Delta t}\right)B_H^d + \left(1 - e^{-(\alpha_b(\theta)(1-q))\Delta t}\right)B_O^d\right).$$

Solving for B^o and letting $\Delta t \to 0$, using l'Hopital's rule, we obtain

$$B^o = \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)}\left(qB_H^d + (1-q)B_O^d\right). \tag{3.6}$$

Next, we turn to the entrepreneurs' outside option E^o. Following similar steps as in the derivation of B^o we obtain

$$E^o = \frac{\alpha_e(\theta)}{r + \alpha_e(\theta)}\left(qE_H^d + (1-q)E_O^d\right). \tag{3.7}$$

The outside options, B^o and E^o, are a decreasing function of the discount rate r, which captures the cost of a deal delay, and an increasing function of the speed of matching $\alpha_b(\theta)$ and $\alpha_e(\theta)$.

Following similar steps as in the derivations above, if all banks fund T entrepreneurs only, the outside options become

$$B_T^o = \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)}qB_H^d, \qquad E_T^o = \frac{\alpha_e(\theta)}{r + \alpha_e(\theta)}qE_H^d. \tag{3.8}$$

Finally, if all banks fund high value projects only, both from the T and the O groups, the outside options become

$$B_H^o = \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)}(q + (1-q)\mu)B_H^d, \qquad E_H^o = \frac{\alpha_e(\theta)}{r + \alpha_e(\theta)}(q + (1-q)\mu)E_H^d. \tag{3.9}$$

It should be highlighted that when banks fund both T and the O entrepreneurs, like going from (3.8) to (3.9), the outside options increase.

3.7 Inflows and outflows

To close the model, we need to specify what the inflows and outflows are. Stationarity requires that the inflow of banks and entrepreneurs matches their respective outflow. Let n_b and n_e denote the measure of banks and entrepreneurs arriving in the market over one unit of time. The inflows n_b and n_e are fully, and uniquely, determined by the stationarity conditions $n_b = \alpha_b(\theta)N_b$ and $n_e = \alpha_e(\theta)N_e$, respectively.

4 Analysis and theoretical results

We begin with an unregulated market. There are two possible symmetric equilibria in pure strategies. Either no bank produces information, $\phi = 0$, or all banks produce information, $\phi = 1$. There may also exist an asymmetric equilibrium in pure strategies where a fraction of banks chooses $\phi = 0$ and the remaining banks choose $\phi = 1$. We then introduce a regulation that forces banks to extend loans to underfunded entrepreneurs, but does not mandate screening.[17] Again, there are two possible symmetric equilibria and an asymmetric equilibrium (in pure strategies). We are interested in the effect of regulation on banks' (steady state) equilibrium incentives to invest in ϕ.[18]

4.1 Unregulated market equilibrium

We determine the range of the information cost parameter c that supports either $\phi = 0$, or $\phi = 1$, or both as equilibria. We deal with asymmetric equilibria in Appendix B.5.

4.1.1 Equilibrium where no bank invests in the screening technology, $\phi = 0$

We denote the choices of all other banks by $\hat{\phi}$, and NR stands for no regulation. The unilateral incentive to invest in the screening technology, Δ^{NR}, is the difference between the expected profits when a bank has

[17] This is consistent with the CRA regulation, which is the focus of our empirical study.

[18] In what follows, when we say equilibrium it should be understood as a steady state equilibrium.

chosen $\phi = 1$, assuming all other banks have not invested in the screening technology and the profits when all banks have chosen $\phi = 0$. This is expressed as follows (all the derivation details are in Appendix A.1.1)

$$\Delta^{NR}(\theta|\hat{\phi} = 0) = \frac{\alpha_b(\theta)\left(H - (1+\bar{i})\right)}{r + \alpha_b(\theta)} \left[\frac{q(A_H^{dev}(\theta) - A_T(\theta))}{(1 + A_H^{dev}(\theta))(1 + A_T(\theta))} + \frac{(1-q)\mu A_H^{dev}(\theta)}{(1 + A_H^{dev}(\theta))}\right], \tag{4.1}$$

where the $A_T(\theta)$ and $A_H^{dev}(\theta)$ terms are given by (A.2) and (A.5) and capture the bank's market power. The market power is a decreasing function of demand-adjusted competition θ. $A_T(\theta)$ arises when all banks fund T entrepreneurs and $A_H^{dev}(\theta)$ arises when the bank funds all high value projects, both from the T and the O group, as a deviation, i.e., all other banks fund T entrepreneurs only.

As $\theta \to 0$, $\Delta^{NR}(\theta|\hat{\phi} = 0) \to [H - (1+\bar{i})]\frac{2\mu(1-q)}{(2-q)(2-q-\mu(1-q))}$. This situation represents the highest market power for banks. Furthermore, as $\theta \to \infty$, $\Delta^{NR}(\theta|\hat{\phi} = 0) \to 0$. This situation represents the lowest market power for banks. Unsurprisingly, banks in this case have zero incentives to invest in the screening technology.[19]

In Appendix A.1.1 we show that $A_H^{dev}(\theta) > A_T(\theta)$. A bank has a higher market power when it can screen low value projects out than when it does not have this ability, and in either case all other banks cannot separate high from low value projects in the O group. This is because in the former case the bank has a higher probability of striking a deal than in the latter, which implies a higher outside option. In turn, this implies a higher interest rate and higher expected profits. Assumption 1 then guarantees $\Delta^{NR}(\theta|\hat{\phi} = 0) > 0$.

We can conclude that $\phi = 0$ is a symmetric equilibrium if and only if $\Delta^{NR}(\theta|\hat{\phi} = 0) < c$, i.e., the expected profit gain from a unilateral deviation is lower than the cost of the screening technology.

4.1.2 Equilibrium where all banks invest in the screening technology, $\phi = 1$

The reduction in expected profits when a bank unilateral does not invest in the screening technology, assuming all other banks have chosen $\hat{\phi} = 1$, is given by (all the derivation details are in Appendix A.1.2)

$$\Delta^{NR}(\theta|\hat{\phi} = 1) = \frac{\alpha_b(\theta)\left(H - (1+\bar{i})\right)}{r + \alpha_b(\theta)} \left[\frac{q(A_H(\theta) - A_T^{dev}(\theta))}{(1 + A_H(\theta))(1 + A_T^{dev}(\theta))} + \frac{(1-q)\mu A_H(\theta)}{(1 + A_H(\theta))}\right], \tag{4.2}$$

[19]For the derivation of these limits we use the results in footnotes 32 and 33.

where $A_H(\theta)$ is given by (A.10) and captures the bank's market power when all banks fund high value projects both from T and O entrepreneurs. $A_T^{dev}(\theta)$ is given by (A.12) and reflects the bank's market power when a bank funds T entrepreneurs only as a deviation, i.e., all other banks fund high value projects both from T and O entrepreneurs.

As in Section 4.1.1, when $\theta \to 0$, $\Delta^{NR}(\theta|\hat{\phi}=1) \to [H-(1+\bar{i})]\frac{2\mu(1-q)}{(2-q)(2-q-\mu(1-q))}$ and when $\theta \to \infty$, $\Delta^{NR}(\theta|\hat{\phi}=1) \to 0$.[20]

We can conclude that $\phi = 1$ is a symmetric equilibrium if and only if $\Delta^{NR}(\theta|\hat{\phi}=1) > c$, i.e., the profit reduction from a unilateral deviation to no investment is higher than the cost savings.

4.1.3 Equilibrium conditions

Banks and entrepreneurs have rational expectations, so they anticipate the equilibrium. If a bank unilaterally deviates in ϕ, then it will become common knowledge to the bank and the entrepreneur that matches with the bank (and with entrepreneurs that possibly match with the deviating bank in the future). Other banks and entrepreneurs do not observe this deviation. The bank's outside option is affected, but the entrepreneur, who matches with the deviating bank, does not expect his outside option to change. This is because: i) the quality of the project is match-specific, so if a deal is not reached the project's expected value draw is independent from past draws and ii) bank and entrepreneur expect all other banks to have chosen the equilibrium ϕ. An equilibrium under no regulation is characterized by the following conditions:

i) the equilibrium interest rate, (A.3) or (A.9), maximizes the Nash product (3.4);

ii) the outside options (B^o, E^o) satisfy the asset value equations, (3.6) and (3.7), or (3.8), or (3.9);

iii) the flows and stocks of entrepreneurs and banks, (n_e, n_b) and (N_e, N_b) respectively, satisfy the stationarity conditions $n_b = \alpha_b(\theta)N_b$ and $n_e = \alpha_e(\theta)N_e$; and

iv) the equilibrium choice of the screening technology satisfies

$$\phi^\star = \arg\max_{\phi\in\{0,1\}} E\pi^{NR}(\phi|\hat{\phi}=\phi^\star) - \mathbb{1}_\phi c,$$

[20] For the derivation of these limits we use the results in footnotes 35 and 37.

given agents' beliefs and all other banks' choices $\hat{\phi}$, where the expected profits $E\pi^{NR}$ are given by (A.4) and (A.7), when $\hat{\phi} = 0$, (A.11) and (A.13) when $\hat{\phi} = 1$ and $\mathbb{1}_\phi$ is an indicator function that takes the value of 1 if $\phi = 1$ and 0 if $\phi = 0$.

4.1.4 Results

The Lemma below summarizes the equilibria under no regulation.[21]

Lemma 1 *The conditions under which the different equilibria in investments ϕ in screening technology arise under no regulation are described below.*

- *Suppose $\Delta^{NR}(\theta|\hat{\phi} = 1) > \Delta^{NR}(\theta|\hat{\phi} = 0)$.*
  - *If $c > \Delta^{NR}(\theta|\hat{\phi} = 1)$ the unique symmetric equilibrium is $\phi^\star = 0$.*
  - *If $c < \Delta^{NR}(\theta|\hat{\phi} = 0)$ the unique symmetric equilibrium is $\phi^\star = 1$.*
  - *If $c \in (\Delta^{NR}(\theta|\hat{\phi} = 0), \Delta^{NR}(\theta|\hat{\phi} = 1))$ there are two symmetric equilibria, $\phi^\star = 0$ and $\phi^\star = 1$ and an asymmetric equilibrium, where a fraction $\hat{x}_0$ of banks chooses $\phi^\star = 0$ and the rest choose $\phi^\star = 1$.*
- *Suppose $\Delta^{NR}(\theta|\hat{\phi} = 0) > \Delta^{NR}(\theta|\hat{\phi} = 1)$.*
  - *If $c > \Delta^{NR}(\theta|\hat{\phi} = 0)$ the unique symmetric equilibrium is $\phi^\star = 0$.*
  - *If $c < \Delta^{NR}(\theta|\hat{\phi} = 1)$ the unique symmetric equilibrium is $\phi^\star = 1$.*
  - *If $c \in (\Delta^{NR}(\theta|\hat{\phi} = 1), \Delta^{NR}(\theta|\hat{\phi} = 0))$ a symmetric equilibrium does not exist. Nevertheless, an asymmetric equilibrium exists, where a fraction $\hat{x}_0$ of banks chooses $\phi^\star = 0$ and the rest choose $\phi^\star = 1$.*

The numerical Example 1 we present later illustrates that $\Delta^{NR}(\theta|\hat{\phi} = 1)$ can be higher or lower than $\Delta^{NR}(\theta|\hat{\phi} = 0)$ depending on parameter values. That is why Lemma 1 considers both possibilities.

In the next Lemma we state the effect of competition on the banks' unilateral incentives to invest in the screening technology.

[21] The proof of the asymmetric equilibrium is presented in Appendix B.5.

Lemma 2 *Bank unilateral incentives to invest in the screening technology,* $\Delta^{NR}(\theta|\hat{\phi} = 0)$*, see* (4.1)*, and* $\Delta^{NR}(\theta|\hat{\phi} = 1)$*, see* (4.2)*, are decreasing in demand-adjusted bank competition,* θ*, when the fraction of informationally opaque entrepreneurs is high, i.e., high* $1 - q$*, and when the fraction of informationally opaque entrepreneurs is low, i.e., low* $1 - q$*, they can be inverted U-shaped.*

It is easy to verify that both the (candidate) equilibrium and the deviation profits are decreasing in θ. The interesting question, the answer to which determines the incentives to invest in ϕ, is which one is decreasing faster? By inspecting (4.1), or (4.2), we can see that the first and the third terms are decreasing in θ but the second term's dependence on θ is in general ambiguous. It turns out that if q is sufficiently high, the second term is initially increasing in θ, causing the incentives to invest in ϕ to be initially increasing in θ. Moreover, for high values of θ the incentives are decreasing in θ. This gives rise to the inverted-U relationship. But for low q, the incentive to invest is monotonically decreasing in competition.

We make the following assumption.

Assumption 2 *The information cost,* c*, satisfies*

$$c > \max\left\{\Delta^{NR}(\theta|\hat{\phi} = 0), \Delta^{NR}(\theta|\hat{\phi} = 1)\right\}, \textit{ for all } \theta.$$

The above assumption rules out very low values of the information cost c that would make the problem uninteresting and perhaps also unrealistic. For very low values of c, $\phi^\star = 1$ is the unique equilibrium (at least for some θ), banks fund the high value projects among O entrepreneurs and therefore there is no need for regulation. We summarize in the Lemma below.

Lemma 3 *Suppose Assumptions 1-2 hold and the regulation is not in force. Then,* $\phi^\star = 0$*, i.e., banks do not produce information. As a result, in equilibrium, banks cannot separate high from low value projects of the informationally opaque (O) group and only fund high value projects of the informationally transparent (T) entrepreneurs. The equilibrium interest rate for the T group of entrepreneurs is given by* (A.3).

4.2 Efficiency

A social planner would invest in the screening technology if and only if the incremental benefit was higher than the information cost. The incremental benefit is the additional expected profit from funding the high value projects in the O group. Hence, the social benefit (incentive), per borrower, is: $(1-q)\mu[H-(1+\bar{i})]$.

Social and private incentives for information production are not aligned. The reason for this is the fact that banks do not appropriate all the surplus in their bargaining with the entrepreneurs, even when $\theta \to 0$.[22] This, however, does not imply that private incentives are always weaker than social incentives, they can be stronger.

When banks have the highest market power, as is the case when $\theta \to 0$, then, as the analysis in Sections 4.1.1 and 4.1.2 revealed, the private incentives are $(1-q)\mu[H-(1+\bar{i})]\frac{2}{(2-q)(2-q-\mu(1-q))}$. It can be verified that when $\theta \to 0$, the private incentives are weaker than the social incentives if and only if q and μ are low.[23] If private incentives are lower than the social incentives, when θ is close to zero, then they will be lower for any θ if q is low, since in this case the private incentives are decreasing in θ, Lemma 2.

Therefore, the market equilibrium can entail 'too little' or 'too much' investment in screening technology, relative to the efficient outcome. We focus on the possibility of too little investment, which arises when the fraction of informationally opaque entrepreneurs is high, consistent with the theme of our study. In this case, regulation is meaningful and has the potential to restore efficiency. We summarize in the Lemma below.

Lemma 4 *Suppose the fraction of informationally opaque entrepreneurs, $1-q$, is high. In addition, assume that the probability of high value projects among the informationally opaque entrepreneurs, μ, is low. Then, the market incentives for information production fall short of the social incentives.*

The implication of the above Lemma is that for a range of the information cost c, that is below the social incentives and above the private incentives, the market does not produce socially efficient information.

[22]If instead of the standard Nash bargaining problem, (3.4), we had considered a generalized Nash bargaining problem with a weight $\eta \in (0,1)$ for the bank, then the bank would appropriate all the surplus as $\eta \to 1$. In this case social and private incentives for information production would be perfectly aligned. This would be analogous to the case of a monopolist who practices perfect price discrimination and as a result produces the socially optimal level of quantity.

[23]More specifically, when $q=0$ the private incentives are given by $\mu[H-(1+\bar{i})]\frac{2}{2(2-\mu)}$ and are lower than the social incentives for any $\mu \in [0,1]$. In addition, for any $q \in (0, 2-\sqrt{2})$, there exists a threshold given by $\frac{q^2-4q+2}{q^2-3q+2}$ such that social incentives are stronger than private if and only μ is below that threshold. When $q > 2-\sqrt{2}$ private incentives are stronger than social for any μ.

4.3 Regulated equilibrium

The analysis so far and in particular Lemmas 3 and 4 have shown that in the unique unregulated equilibrium $\phi^{\star} = 0$, while the social planner, provided that the cost of information c is not very high, prefers $\phi^{fb} = 1$, where fb stands for first-best.

When the regulation is in force (R), banks must also fund entrepreneurs in the O group. However, financing low value projects is inefficient. Hence, the regulator runs the risk of creating a new inefficiency, if banks do not enhance their screening ability. In that scenario banks, in order to comply with the regulation, extend loans indiscriminately to the O group of entrepreneurs, so low value projects receive funding too. But if the regulation induces banks to switch to the $\phi = 1$ equilibrium, then efficiency is enhanced, since banks can identify the high value projects among the O entrepreneurs. The regulator's goal is to eliminate one inefficiency (i.e., the lack of screening ability) with the 'threat' of another one (i.e., financing unprofitable projects).

4.3.1 Equilibrium where no bank invests in the screening technology, $\phi = 0$

The incentive to unilaterally invest in the screening technology, and all other banks have chosen $\hat{\phi} = 0$, is given by (all the derivation details are in Appendix B.2.1)

$$\Delta^R(\theta|\hat{\phi} = 0) =$$
$$\frac{\alpha_b(\theta)(H - (1 + \bar{i}))}{r + \alpha_b(\theta)} \left[\frac{q(A_H^{dev}(\theta) - A(\theta))}{(1 + A_H^{dev}(\theta))(1 + A(\theta))} + (1 - q) \left(\frac{\mu A_H^{dev}(\theta)}{1 + A_H^{dev}(\theta)} - \frac{\mu H - (1 + \bar{i})}{H - (1 + \bar{i})} \right) \right],$$

where $A(\theta)$ is given by (B.3) and captures the bank's market power when all banks fund all projects in the O group.

When banks invest in screening technology, the regulation is not binding. In this case banks fund entrepreneurs in the O group with or without regulation. The regulation becomes binding when banks do not invest in information production. It can be readily verified that $A_H^{dev}(\theta) > A(\theta)$. Combined with Assumption 1 we have that $\Delta^R(\theta|\hat{\phi} = 0) > 0$. Hence, $\phi = 0$ is a symmetric equilibrium under regulation if and only if $\Delta^R(\theta|\hat{\phi} = 0) < c$.

4.3.2 Equilibrium where all banks invest in the screening technology, $\phi = 1$

The unilateral incentive not to invest in the screening technology when the regulation is in force, and all other banks have chosen $\hat{\phi} = 1$, is given by (all the derivation details are in Appendix B.2.2)

$$\Delta^R(\theta|\hat{\phi} = 1) = \\ \frac{\alpha_b(\theta)(H - (1 + \bar{i}))}{r + \alpha_b(\theta)} \left[\frac{q(A_H(\theta) - A^{dev}(\theta))}{(1 + A_H(\theta))(1 + A^{dev}(\theta))} + (1 - q) \left(\frac{\mu A_H(\theta)}{1 + A_H(\theta)} - \frac{\mu H - (1 + \bar{i})}{H - (1 + \bar{i})} \right) \right],$$

where $A^{dev}(\theta)$ is given by (B.7) and measures the bank's market power when it funds all projects in the O group as a deviation, i.e., all other banks fund high value projects only.

It can be readily verified that $A_H(\theta) > A^{dev}(\theta)$. Combined with Assumption 1 we have that $\Delta^R(\theta|\hat{\phi} = 1) > 0$. The deviating bank experiences lower profits, but it saves the cost c. We can conclude that $\phi = 1$ is a symmetric equilibrium under regulation if and only if $\Delta^R(\theta|\hat{\phi} = 1) > c$.[24]

4.4 Results

The Lemma below, which very much resembles Lemma 1, summarizes the equilibria under regulation.[25]

Lemma 5 *The conditions under which the different equilibria in investments ϕ in screening technology arise under regulation are described below.*

- *Suppose* $\Delta^R(\theta|\hat{\phi} = 1) > \Delta^R(\theta|\hat{\phi} = 0)$.
  - *If* $c > \Delta^R(\theta|\hat{\phi} = 1)$ *the unique symmetric equilibrium is* $\phi^\star = 0$.
  - *If* $c < \Delta^R(\theta|\hat{\phi} = 0)$ *the unique symmetric equilibrium is* $\phi^\star = 1$.
  - *If* $c \in (\Delta^R(\theta|\hat{\phi} = 0), \Delta^R(\theta|\hat{\phi} = 1))$ *there are two symmetric equilibria,* $\phi^\star = 0$ *and* $\phi^\star = 1$ *and an asymmetric equilibrium, where a fraction* $\hat{x}_0$ *of banks chooses* $\phi^\star = 0$ *and the rest choose* $\phi^\star = 1$.
- *Suppose* $\Delta^R(\theta|\hat{\phi} = 0) > \Delta^R(\theta|\hat{\phi} = 1)$.

[24] The equilibrium conditions under regulation are very similar with the ones presented in Section 4.1.3 and so we omit them.
[25] The analysis for asymmetric equilibria follows closely the one under no regulation, see Appendix B.5, and is omitted.

- *If* $c > \Delta^R(\theta|\hat{\phi} = 0)$ *the unique symmetric equilibrium is* $\phi^\star = 0$*.*
- *If* $c < \Delta^R(\theta|\hat{\phi} = 1)$ *the unique symmetric equilibrium is* $\phi^\star = 1$*.*
- *If* $c \in (\Delta^R(\theta|\hat{\phi} = 1), \Delta^R(\theta|\hat{\phi} = 0))$ *a symmetric equilibrium does not exist. Nevertheless, an asymmetric equilibrium exists, where a fraction* $\hat{x}_0$ *of banks chooses* $\phi^\star = 0$ *and the remaining banks choose* $\phi^\star = 1$*.*

The Lemma below compares the banks' incentives to produce information between no regulation and regulation.

Lemma 6 *The unilateral incentives to invest are higher under regulation, that is*

$$\Delta^R(\theta|\hat{\phi} = 1) > \Delta^{NR}(\theta|\hat{\phi} = 1) \text{ and } \Delta^R(\theta|\hat{\phi} = 0) > \Delta^{NR}(\theta|\hat{\phi} = 0)$$

and for μ *low enough the following holds*

$$\max\left\{\Delta^{NR}(\theta|\hat{\phi} = 0), \Delta^{NR}(\theta|\hat{\phi} = 0)\right\} < \min\left\{\Delta^{R}(\theta|\hat{\phi} = 0), \Delta^{R}(\theta|\hat{\phi} = 0)\right\}.$$

Regulation implies low bank profits when $\phi^\star = 0$, due to the fact that banks are forced to fund the O entrepreneurs and cannot screen out the low value projects. Hence, the incremental profit when a bank chooses $\phi = 1$ is higher under regulation.

How does demand-adjusted bank competition, θ, affect the incentives to invest? The answer is the same as in the case of no regulation, see Lemma 2, so a formal proof is omitted.

Example 1 *Suppose the matching technology is given by the example in footnote 9. Also assume,* $\bar{\imath} = 0.1$*,* $\mu = 0.2$*,* $H = 5$*,* $r = 0.2$ *and* $\xi = 2$*.*

In Figure 1, we present the (unilateral) incentives to invest in ϕ both under no regulation and regulation, for a low q (left panel) and a high q (right panel), using the numerical values from Example 1. Presented in Figure 1 are two key results of the equilibrium search model of bank competition. First, (unilateral) incentives to invest in ϕ are stronger under regulation and second, higher θ either monotonically, or eventually, diminishes these incentives.

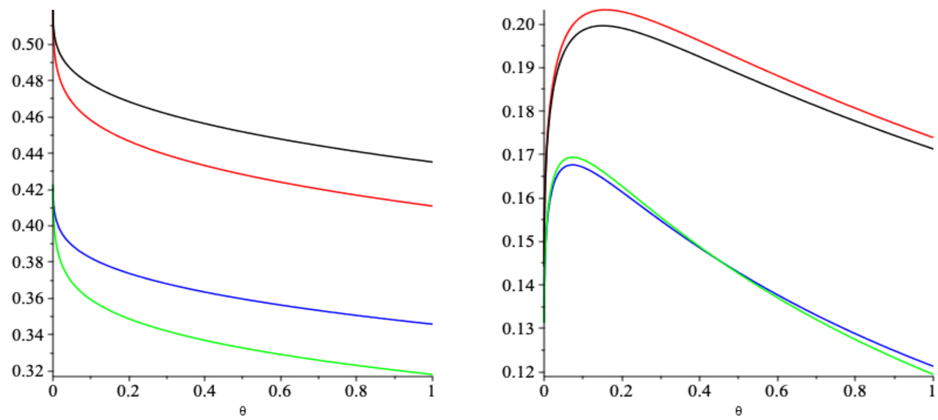

Figure 1: Unilateral incentives to invest in screening technology as function of competition θ. $\Delta^{NR}(\theta|\hat{\phi} = 0)$–(blue curve); $\Delta^{NR}(\theta|\hat{\phi} = 1)$–(green curve); $\Delta^{R}(\theta|\hat{\phi} = 0)$–(black curve) and $\Delta^{R}(\theta|\hat{\phi} = 1)$–(red curve). In the left panel $q = 0.2$ in the right panel $q = 0.9$.

4.4.1 The effects of competition on the incentives to invest in the screening technology

Recall from Assumption 2 that the cost of information c is above a very low threshold. This implies that under no regulation banks, in equilibrium, make no investments in the screening technology. We also assume, following Section 4.2, that this equilibrium is socially inefficient, since a social planner would want banks to invest.[26] Thus, we assume that c is higher than the threshold given in Assumption 2 and below $(1-q)\mu[H-(1+\bar{i})]$. This is the interesting range, as a c outside this range would imply that either information production is not even socially efficient (if c is very high), or information is always produced (if c is very low). We also assume that $1-q$, the fraction of O entrepreneurs, is high so that the incentives to invest are decreasing in θ, see Lemma 2.[27]

How does regulation affect banks' incentives to invest in ϕ? Recall that $\phi^{\star} = 1$ denotes an equilibrium in which all banks invest in the screening technology and $\phi^{\star} = 0$ an equilibrium in which no bank invests. We will divide the remaining range of the information cost into three regions: high, intermediate and low. The first two Propositions state the main results for high or low information cost values. The Proofs are straightforward and are omitted.

[26] Provided that $c < (1-q)\mu[H-(1+\bar{i})]$.

[27] This is a reasonable assumption, since in our data we have small businesses in low to medium income census tracts, where the fraction of informationally transparent businesses is arguably low.

Proposition 1 *Suppose Assumptions 1-2 hold and the information cost is high, i.e.,*

$$c > \max\left\{\Delta^R(\theta|\hat{\phi}=1), \Delta^R(\theta|\hat{\phi}=0)\right\}, \text{ for all } \theta. \tag{4.3}$$

Then, $\phi^\star = 0$, both under no regulation and regulation, for any θ.

Proposition 2 *Suppose Assumptions 1-2 hold and the information cost is low, i.e.,*

$$c < \min\left\{\Delta^R(\theta|\hat{\phi}=1), \Delta^R(\theta|\hat{\phi}=0)\right\}, \text{ for all } \theta. \tag{4.4}$$

Then, the following hold:

- *When the regulation is not in force $\phi^\star = 0$, for any θ.*
- *When the regulation is in force $\phi^\star = 1$, for any θ.*

A more plausible case is arguably when the information cost is intermediate. The Proposition below states the result. The proof is in Appendix B.4.

Proposition 3 *Suppose Assumptions 1-2 hold and there are two separate markets, one with high demand-adjusted bank competition, $\hat{\theta}$, and one with low demand-adjusted bank competition, $\tilde{\theta}$, where $\tilde{\theta} < \hat{\theta}$. Moreover, assume that the cost of information is intermediate, i.e.,*

$$\max\left\{\Delta^R(\hat{\theta}|\hat{\phi}=1), \Delta^R(\hat{\theta}|\hat{\phi}=0)\right\} < c < \min\left\{\Delta^R(\tilde{\theta}|\hat{\phi}=1), \Delta^R(\tilde{\theta}|\hat{\phi}=0)\right\}. \tag{4.5}$$

Then, the following hold:

- *When the regulation is not in force, $\phi^\star = 0$, for any θ.*
- *When the regulation is in force*
  - *in the low demand-adjusted bank competition market, $\tilde{\theta}$, $\phi^\star = 1$ and*
  - *in the high demand-adjusted bank competition market, $\hat{\theta}$, $\phi^\star = 0$.*

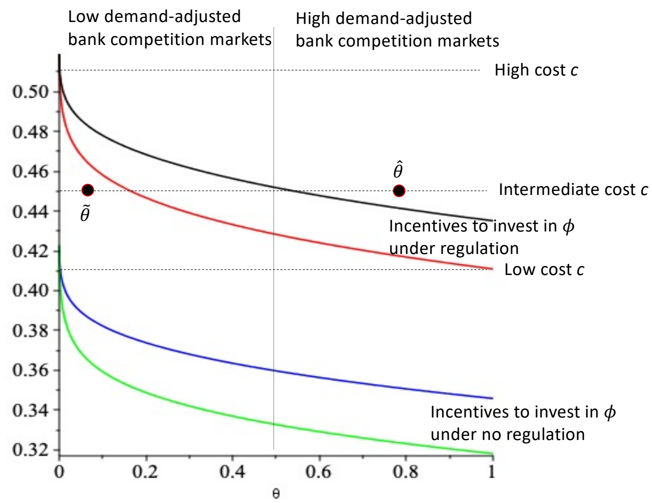

Figure 2: The effect of demand-adjusted bank competition, θ, and the cost of information, c, on the equilibrium information production, ϕ. The curves are the same with the left panel of Figure 1.

The proofs of all three Propositions combined can be best understood with the aid of Figure 2. The high cost line, that satisfies condition (4.3), is above all the curves that represent the incentives to deviate, so in all equilibria $\phi^{\star} = 0$. The low cost line, that satisfies condition (4.4) and Assumption 2, is above the curves under no regulation, so $\phi^{\star} = 0$ and below the curves under regulation, so $\phi^{\star} = 1$. Finally, the intermediate cost line, that satisfies condition (4.5), is above the curves under no regulation, so $\phi^{\star} = 0$, below the curves under regulation for low demand-adjusted bank competition, $\theta = \tilde{\theta}$, so $\phi^{\star} = 1$, but above the curves under regulation for high demand-adjusted bank competition, $\theta = \hat{\theta}$, so $\phi^{\star} = 0$.[28]

The main take-away is that, when the cost of information c is high, regulation has no effect on banks' ability to screen high value from low value projects. When the regulation does not induce banks to switch to an equilibrium where they invest in the screening technology, the average effect on the credit worthiness in the market can be negative or zero. This is because entrepreneurs with low value projects also obtain loans. If c is low, regulation incentivizes banks to invest regardless of how many entrepreneurs relative to bank branches are in the market. When banks can target the new loans they have to give due to the regulation to entrepreneurs with high value projects, the average creditworthiness in the market increases. Lastly, when the cost of information is intermediate, regulation incentivizes banks to enhance their screening ability

[28]We pick two markets, $\tilde{\theta}$ and $\hat{\theta}$ in Figure 2, that are far apart from each other so that we obtain a sharp contrast in terms of the equilibrium information production strategies banks adopt in each market.

only in markets with many entrepreneurs relative to competing bank branches. Hence, in those markets we expect the average creditworthiness to rise. The effect in markets with few entrepreneurs relative to competing bank branches should be negative or zero.

4.5 Testable hypotheses

The analysis in Section 4.4.1 leads to the following competing testable hypotheses. Recall that the theoretical model assumes that regulation will increase, in expectation, the number of loans.

Hypothesis 1 *Regulation decreases (or leaves unchanged) the average creditworthiness in any local market, regardless of the demand-adjusted bank competition intensity.*

Hypothesis 2 *Regulation increases the average creditworthiness in any market, regardless of the demand-adjusted bank competition intensity.*

Hypothesis 3 *Regulation increases the average creditworthiness in local markets with low demand-adjusted bank competition intensity and decreases (or leaves unchanged) the average creditworthiness in local markets with high demand-adjusted bank competition intensity.*

The above three Hypotheses follow directly from Propositions 1, 2 and 3. We turn to the empirical analysis next, where we test the Hypotheses.

5 Empirical investigation

In the theoretical model we assume that a market consists of two types of entrepreneurs, the informationally opaque and the transparent. According to Lemma 3, without a regulation in force, banks only fund transparent entrepreneurs because they are more profitable for the banks due to lower screening costs. This means that worthy projects in markets with a high concentration of opaque entrepreneurs will remain unfunded. To deal with this market imperfection, a regulation is required to compel banks to produce information.

In the empirical investigation, we focus on the Community Reinvestment Act (CRA) regulation which is aimed to address the credit market imperfection responsible for the low credit availability in low-and moderate-income communities. These communities contain a proportionally larger number of opaque entrepreneurs. Specifically, banks economize on information costs, by using information generated from past lending, both their own lending and that of other lenders. A loan therefore creates a positive externality to future lenders and borrowers in the form of valuable information. Since low-and moderate-income communities are systematically underfunded, they are deprived of the positive information externalities and therefore they include a larger number of opaque entrepreneurs. Thus, as predicted by our model, banks will refrain from funding low-and moderate-income communities without any regulation. The practical implication of the information spillover and the engendering path dependence is that businesses do not have a credit report if they do not have credit and this can be a source of persistent inequality.

5.1 Institutional background

The Community Reinvestment Act (CRA) deals with the lack of banks' incentives to screen and give loans to opaque borrowers, by requiring banks to lend to small businesses located in underfunded neighborhoods in which the bank has presence. As of 1995, banks are regularly examined for their compliance in the areas of lending, investment and services by three federal agencies: the Federal Reserve, the Office of the Comptroller of the Currency (OCC), and the Federal Deposit Insurance Corporation (FDIC). The lending area has been given the highest weight in determining overall ratings. The CRA ratings are publicly released and due to elevated public attention, institutions need to show satisfactory responsiveness to the credit needs of assessment areas based on where they operate and particularly to the financing needs of small businesses in low-to-moderate income communities.

The empirical evidence discussed below supports the notion that CRA regulation is associated with an increase in small business lending. The effectiveness of CRA is linked to the two channels that incentivize compliance by depository institutions (Ding, Lee, and Bostic (2018)). First, federal regulators must consider an institution's CRA performance when evaluating an application by that institution for a merger or acquisition, the formation of branch, or other business activity. Second, and more indirectly, community activists and public interest groups monitor banking institutions and provide an independent source of bank discipline.

5.2 Identification

Census tracts are identified as low and moderate-income areas based on demographic information provided by the US census. According to the Federal Financial Institutions Examination Council (FFIEC), until the end of 2011 a tract's eligibility status assignment and banks' CRA performance assessment was conducted using the 2000 census data and from January 2012 forward it was conducted using 2010 census data. As a result, approximately 15 percent of the tracts that were marginally non-eligible in the 2000 census became CRA eligible tracts.

We focus on newly eligible tracts for two main reasons. First, census tract eligibility does not necessarily imply an automatic increase in small business lending. Specifically, bank compliance to CRA is assessed at higher geographical level (i.e. assessment areas are at county or district level) which means that census tract eligibility increases the probability of small business loans growth. Therefore, we need to identify the eligible tracts that actually experience an unexpected increase in the loan supply. This supply-driven loan growth is more easily observed among the newly eligible tracts compared to census tracts that have been eligible since the 2000 census data and thus have a more predictable supply of loans ever since.

Second, focusing on newly eligible tracts where banks are, presumably, less likely to have built long-term relationships with the local businesses helps us to avoid any confounding effect from relationship lending on the credit score of businesses. In contrast, if we had focused on census tracts that have been eligible for a long period, it would have been difficult to identify whether the impact of regulation eligibility is due to banks' screening effort to reduce the adverse selection, or a result of the relationship lending benefits enjoyed by targeted borrowers.

We exploit the fact that the community's CRA eligibility is allocated on the basis of the observed ratio of census tract's median family income to the median family income of the Metropolitan area to use the regression-discontinuity (RD) design. In RD design, units receive treatment on the basis of whether the value of an observed covariate, known as the running variable, is above or below a known cutoff. The underlying idea is that the probability of receiving treatment conditional on this covariate jumps discontinuously at the cutoff, inducing heterogeneity in treatment assignment that is unrelated to potential confounders. In sharp RD design assignment to treatment coincides with the actual treatment and thus the jump in the probability of receiving treatment at the cutoff is from zero to one. When treatment compliance is imperfect, as in the case of CRA, the method is the fuzzy RD design and the jump in the probability of receiving treatment at the cutoff is less than one.

Specifically, if a community i, defined at census tract level, has a median family income ratio X_i that is less than the cutoff value $c = 80\%$, the loans given to the community count as CRA-eligible while those census tracts with ratio above the cutoff are not. This generates a discontinuity in eligibility as a function of the tract's median family income ratio. The discontinuity in eligibility implies that affected tracts are more likely to experience an unexpected increase in the loan supply disproportional to the growth in small business loans among unaffected census tracts. We define as "treatment" the CRA-eligibility induced increase in small business lending and we denote it with the dummy variable T_i. Then, fuzzy RD design assumes that $\lim_{\epsilon\to 0} \Pr(T_i = 1|X_i = c - \epsilon) - \lim_{\epsilon\to 0} \Pr(T_i = 1|X_i = c + \epsilon) > 0$.

At the same time, there appears to be no reason, other than the eligibility of the tract, for the credit score Y_i of small businesses located in the tract, to be a discontinuous function of the tract's median family income ratio. In this setting, the treatment effect of the increase in the probability of small business loan growth on the outcome variable Y_i can be written as

$$\tau = \frac{\lim_{\epsilon\to 0} \mathrm{E}(Y_i|X_i = c - \epsilon) - \lim_{\epsilon\to 0} \mathrm{E}(Y_i|X_i = c + \epsilon)}{\lim_{\epsilon\to 0} \mathrm{E}(T_i|X_i = c - \epsilon) - \lim_{\epsilon\to 0} \mathrm{E}(T_i|X_i = c + \epsilon)} = \frac{\mu_{Y^-} - \mu_{Y^+}}{\mu_{T^-} - \mu_{T^+}}. \tag{5.1}$$

5.3 Estimation methodology

Estimation of the eligibility effect τ within the RD design is performed using local polynomial estimators. This estimation strategy involves approximating the regression functions above and below the cutoff by means of weighted polynomial regressions, typically of order 1 or 2, with weights computed by applying a kernel function on the distance of each observation to the cutoff point. In particular, we employ the robust local linear polynomial estimator of Calonico, Cattaneo, and Titiunik (2014) which corrects for the bias of bandwidth. Formally, the conventional local-linear estimator, for a positive bandwidth h, kernel function $K_h = K((X_i - c)/h)$ and a set of covariates $\mathbb{Z}_i$ is found by optimizing the equations:

$$\hat{\mu}_{U^-} = \arg\min_{\{b_0,b_1\}} \sum_n (\mathbb{1}(X_i < c)(U_i - b_0 - b_1(X_i - c) - \gamma\mathbb{Z}_i)^2 K_h \tag{5.2}$$

$$\hat{\mu}_{U^+} = \arg\min_{\{b_0,b_1\}} \sum_n (\mathbb{1}(X_i \geq c)(U_i - b_0 - b_1(X_i - c) - \gamma\mathbb{Z}_i)^2 K_h \tag{5.3}$$

where the random variable U_i is equal to either Y_i or T_i. The bias-corrected local linear sharp RD estimator of the probability of treatment, i.e. small business loans increase, employing a local-quadratic estimate of the leading bias is

$$\hat{\tau}_T = \hat{\mu}_{T^-} - \hat{\mu}_{T^+} - h_T^2 \mathbb{B}_T(h_T, b_T) \tag{5.4}$$

and the bias-corrected local linear fuzzy RD estimator on the outcome variable (i.e. the average credit score) employing a local-quadratic estimate of the leading bias is

$$\hat{\tau}_Y = \frac{\hat{\mu}_{Y^-} - \hat{\mu}_{Y^+}}{\hat{\mu}_{T^-} - \hat{\mu}_{T^+}} - h_Y^2 \mathbb{B}_Y(h_Y, b_Y) \tag{5.5}$$

where b_T, b_Y are the bias-correction bandwidths and $\mathbb{B}_T(\cdot,\cdot)$, $\mathbb{B}_Y(\cdot,\cdot)$ the bias correction factors see, Calonico, Cattaneo, and Titiunik (2014) for more details.

Increasing the order of the polynomial generally improves the accuracy of the approximation but also increases the variability of the estimator. In particular, it can be shown that the asymptotic variances of the local constant ($p = 0$) and local linear ($p = 1$) polynomial estimators are equal, but the linear approximation has a smaller asymptotic bias (Cattaneo, Idrobo, and Titiunik (2019)). Thus, our main model setup uses the recommended choice of local linear polynomial estimator while a higher order (quadratic) polynomial estimator ($p = 2$) is employed for additional robustness analysis. Moreover, in the analysis that follows, we employ the Epanechnikov kernel function, often called the optimal kernel because it yields the lowest possible asymptotic Mean Square Error (MSE). In additional robustness tests we also use the Uniform kernel function.

A more consequential issue is the choice of the bandwidth because, given the choices of order approximation and kernel function, the accuracy of the approximation is essentially determined by the bandwidth. For all point estimators of the RD effects we use the recommended, Mean Square Error (MSE) optimal plug-in bandwidth h_{MSE} which minimizes the asymptotic variance and bias of the estimator (Cattaneo, Idrobo, and Titiunik (2019)). When we run the quadratic polynomial estimator, we use the common MSE-optimal bandwidth selector for the sum of regression estimates bandwidth $h_{MSE,SUM}$ which is useful for regularization purposes to minimize over-fitting. Furthermore, we use the CER optimal bandwidth h_{CER}, that minimizes the asymptotic coverage error rate of the bias corrected confidence interval, for inference in the falsification tests. Finally, we use the cluster at census tract level robust plug-in residuals variance estimator to compute the variance-covariance matrix.

5.4 Data

Our empirical analysis assembles data from three sources. First, CRA small business loan data is obtained from FFIEC. The CRA regulations require financial institutions, which include all state member banks, state nonmember banks, national banks and savings associations that meet or exceed the asset size thresholds, to report commercial loans extended to small businesses. The filing asset size threshold is published every year by FFIEC and it is typically over $1 billion for each of prior two calendar years. Small business loans are considered the loans with origination amounts of $1 million or less. The reported annual loan data refers to new loans and it includes the total number of loans and the total loan amount at origination aggregated at U.S. census tract level. CRA small business lending data covers roughly 86 percent of all loans of $1 million or less (Greenstone, Mas, and Nguyen (2020)).

Along with the loan data, FFIEC also publishes census data at tract level used in preparing CRA data. Variables that are pertaining to our analysis include Median Family Income (MFI) for census tracts and their surrounding Metropolitan Statistical Area (MSA) or Metropolitan Division (MD). Qualified CRA activities hinge on the income level of census tracts and geographic distribution of loans within assessment areas. The income group classification is determined by tract MFI as a percentage of MSA/MD level MFI (or MFI percentage). For tracts located outside a MSA/MD, area MFI is the statewide non-MSA/MD MFI. Updates of tract MFI or MSA/MD MFI impact the eligibility of tracts and the CRA activities subsequently. According to FFIEC website, 2012 MFI income data for both census tracts and MSA (MD) are collected from 2010 Census while 2011 MFI income data are from 2000 Census. Consistent with the income data, census tracts are classified on the basis of 2000 U.S. Census data for loans originated in 2011. Loans originated in 2012 were classified on the basis of 2010 Census data. Additional demographic variables used in our analysis include the census tract total population and the minority population as a percentage of the total population.

Second, business data is obtained from National Establishment Time Series (NETS) compiled by Walls & Associates. Because small businesses are mainly single establishments, we use the terms firm and establishment interchangeably. Among all NETS establishments, we exclude those from NAICS-industries of Agriculture, Forestry, Fishing and Hunting whose CRA loans are reported separately, and from Education Services and Public Administration which do not typically depend on small business bank loans for their operations. We further dropped establishments with missing NAICS codes, data year and employment numbers, and inactive establishments such as those with bankruptcy filing indicator.

We use the Small Business Administration's definition to characterize a NETS establishment as a small business if it has less than 500 employees and it is in the manufacturing industry or three year average annual sales less than $7.5 million for all non-manufacturing industries. Moreover, we use the Dun & Bradstreet paydex score, the credit score for an establishment similar to FICO for individuals, that is derived from the firm's trade references with vendors and suppliers. Specifically, the paydex score is a value weighted, according to the size of obligations, score that ranges from 0 to 100 with higher values implying higher ability to pay back the obligations. Credit scores like paydex are therefore important in assessing a firm's creditworthiness and banks use them to determine important loan terms such as the loan's tenure, the interest rate and the loan's collateral requirements. Another important feature of paydex score is that it relates with hard data of debt payments including trade credit. This means that any increase in the score is due to an improvement in the firm's cash flow rather than due to the loan-certification effect (James (1987)).

To assign establishments to census tracts, we perform geocoding from an establishment's latitude and longitude provided by NETS and determine whether an establishment falls within the boundary of a census tract. Census tracts boundary files are downloaded from the U.S. Census website. To be consistent with CRA lending data, NETS data before 2012 are mapped to census tracts based on 2000 census map, while 2012 NETS data are mapped based on 2010 census map. To exclude confounding factors that contributed to a business move and CRA lending, we do not include establishments that moved during sample periods. We also dropped establishments with missing coordinates, as well as establishments which are reported to have only one employee (i.e. self-employed). For each year, we calculated the total number of small establishments and the average credit score per census tract.

Finally, bank branch data come from the Summary of Deposits (SOD), the annual survey of branch office deposits for all FDIC-insured institutions, including insured U.S. branches of foreign banks. Following standard treatment, we keep only brick-and-mortar branches and remove any duplicated, in respect to address, branches. Additionally, we dropped branches missing physical addresses. SOD data include latitudes and longitudes of physical branches, however for branches missing latitudes and longitudes we impute coordinates from geocoding the branch addresses. To determine the number of bank branches serving a census tract, we draw a circle with a search radius around the population center of each census tract.[29] We apply a variable distance for urban, mixed and rural tracts. Specifically, we assume an entrepreneur searches within 2 miles around the tract center for banking services in an urban area, 5 miles in a mixed area and 10 miles in a rural area.[30] Any bank branch with its coordinates falling within the corresponding search radius to the tract center is considered serving entrepreneurs in that census tracts. For each tract year, we are able to identify all bank branches serving a particular census tract. Same bank branches within a tract, although a rare event, are treated as distinct units because they also compete for small business loans and because screening using soft information is branch specific. Finally, we also calculated separately the total number of branches of large banks, defined as banks or savings associations with assets exceeding the annually-determined size threshold by CRA regulation and the total number of branches of small banks, defined as banks or savings associations with assets below the annually-determined size threshold by CRA regulation. The latter separation is important due to the different strategic focus between large and small banks with literature suggesting that smaller sized banks are more likely to fund local small businesses (Berger, Miller, Petersen, Rajan, and Stein (2005)).

5.5 Summary statistics

Census tract, the primary unit of analysis, is a relatively permanent statistical subdivision of a county. Tracts are defined to contain 4,000 inhabitants and therefore vary in size across urban and rural areas. Our sample focuses on the period 2011-2012 and includes two types of census tracts, the newly eligible tracts after the incorporation of the latest Census data and the tracts that remained ineligible after the Census data update. Overall, we have 73,366 tract-year observations with 69,315 observations belonging to tracts with median family income above the 80% cut off and 4,051 observations belonging to tracts with median family income below the 80% cut off.

[29]Tract population center files are downloaded from census website (https://www.census.gov/geographies/reference-files/time-series/geo/centers-population.2010.html)

[30]Using the same distance across sparsely and densely populated areas assumes that it takes the same amount of time to cover a distance of 10 miles in rural and urban areas, which is counterfactual. Consequently, a fixed distance would inflate the number of branches in more densely populated areas while underestimating the figure for rural areas. The same approach was adopted by the Bank Policy Institute's report "Some Facts about Bank Branches and LMI Customers".

Table 1: Descriptive statistics for upper-middle income tracts & lower-moderate income tracts.

	Panel A: Bank branch and small business loan data							
	80%≤ MFI (N=69,315)				MFI< 80% (N=4,051)			
VARIABLES	mean	sd	min	max	mean	sd	min	max
Branches of small banks	3.152	4.233	0	78	2.681	3.559	0	52
Branches of large banks	12.47	24.13	0	754	10.24	18.09	0	485
Number of small business loans	78.30	79.51	0	2,770	56.17	76.05	1	2,652
Total amount ($Th) of small business loans	2,932	4,266	0	139,875	2,479	4,696	2	123,345
Demand-adjusted bank competition	12.64	31.11	0	2,933	15.12	123.4	0	7,671
	Panel B: Demographic and business data							
	80%≤ MFI (N=69,315)				MFI< 80% (N=4,051)			
VARIABLES	mean	sd	min	max	mean	sd	min	max
Population	4,314	1,747	23	33,041	4,138	1,747	37	16,326
Minority population (%)	22.36	22.17	0	99.93	42.68	28.57	1.030	99.87
Tract MFI	68,841	28,827	20,833	250,001	45,564	10,789	2,499	85,912
Number of small businesses	182.8	162.4	1	5,774	158.4	171.2	5	5,515
Credit score	71.07	3.355	57.68	77.46	69.61	3.897	57.68	77.46
Years of establishment	16.11	4.319	4	51	17.23	4.830	5.914	50.56
Proprietorship/partnership legal form	0.742	0.0880	0.250	1	0.726	0.0969	0.250	1

Table 1 Panel A summarizes the bank data and particularly the number of branches split by bank size and the number and amount of small business loans. It also includes the demand-adjusted bank competition measured as the number of branches per 100 small firms. Table 1 Panel B summarizes the demographic data including the tract population, the percentage of minority population and the median family income. Moreover it includes the business data and particularly the number of small businesses, the average years of firm's establishment, the proportion of businesses with proprietorship/partnership legal form and the average credit (paydex) score.

Direct comparison between the two groups of tracts reveals that upper-middle income tracts on average have more bank branches, higher number and amount of small business loans, lower local bank competition adjusted for the number of borrowers, and higher firm credit score compared to the lower-moderate income tracts. Moreover, they are slightly more populated and have a lower proportion of minority population. In addition, they display greater economic activity, measured by the number of small businesses which are also slightly younger in age. Finally, both tract groups have approximately the same percentage of proprietorship/partnership entities.

If we concentrate on the tracts that are near to the cut off, we see that the differences between the upper-middle and lower-moderate income tracts are eliminated. Specifically, Table 2 Panels A and B summarize the characteristics of tracts with median family income between 80% and 90% and tracts with median

Table 2: Descriptive statistics for upper-middle income tracts & lower-moderate income tracts with MFI close to the cut off point

Panel A: Bank branch and small business loan data								
	80%≤ MFI<90% (N=10,863)				70%<MFI< 80% (N=2,608)			
VARIABLES	mean	sd	min	max	mean	sd	min	max
Branches of small banks	2.788	3.660	0	72	2.661	3.313	0	28
Branches of large banks	9.752	17.23	0	730	9.803	13.83	0	328
Number of small business loans	55.24	54.68	0	1,517	57.52	81.19	1	2,652
Total amount ($Th) of small business loans	2,234	3,318	0	60,523	2,452	4,644	2	123,345
Demand-adjusted bank competition	12.58	37.23	0	2,933	12.21	21.73	0	350
Panel B: Demographic and business data								
	80%≤ MFI<90% (N=10,863)				70%<MFI< 80% (N=2,608)			
VARIABLES	mean	sd	min	max	mean	sd	min	max
Population	4,153	1,612	37	16,532	4,247	1,722	494	16,326
Minority population (%)	29.47	26.57	0	99.76	39.09	28.13	1.040	99.67
Tract MFI	47,805	11,267	20,833	95,777	48,726	10,303	24,274	85,912
Number of small businesses	154.0	127.9	1	4,024	161.4	183.4	8	5,515
Credit score	70.39	3.659	57.68	77.46	69.81	3.760	57.68	77.46
Years of establishment	17.34	4.708	5	43.50	17.16	4.599	5.914	35.82
Proprietorship/partnership legal form	0.733	0.0945	0.261	1	0.730	0.0947	0.380	1

family income between 70% and 80%. We conclude that tracts from the two sides of the cut off point have similar demographics as well as similar business characteristics and banking activities.

In addition to the sample containing the total number of census tracts, we create a separate sample of matched tracts. At first glance, the use of a matched sample in RD-design seems unnecessary. Heckman, LaLonde, and Smith (1999) argue that regression discontinuity estimators constitute a special case of selection on observables, and thus the RD estimator is a limit form of matching. Indeed, as we observed above, the heterogeneity of observable tracts characteristics is eliminated if we focus on tracts with median family income ratio close to the cut off point. Nonetheless, tracts from different states may be subject to different regulatory and legal frameworks especially concerning state's branch banking policies. Such disparities, if they exist, could interfere with the supply of bank credit. We therefore use a matched-sample for further robustness analysis to address any remaining concerns about unobservable state induced heterogeneity.

Specifically, for every tract that has become eligible under CRA regulation, we find a matched tract located at the same Metropolitan Statistical Area that remained ineligible with similar characteristics such as population, percentage of minority populations, median family income, number of small businesses, number of large bank branches and number of small bank branches. Table 3 Panels A and B summarize the characteristics of tracts from the same Metropolitan Statistical Area with median family income above and below the cut off point of 80%. We conclude that upper-middle income and lower-moderate income tracts from the same Metropolitan Statistical Area have similar demographics as well as similar business and banking activities.

Table 3: Descriptive statistics for upper-middle income tracts & lower-moderate income tracts from the same Metropolitan Statistical Area

	Panel A: Bank branch and small business loan data							
	80%≤ MFI (N=10,190)				MFI< 80% (N=3,974)			
VARIABLES	mean	sd	min	max	mean	sd	min	max
Branches of small banks	2.919	3.854	0	65	2.671	3.555	0	52
Branches of large banks	11.29	18.51	0	365	10.25	18.10	0	485
Number of small business loans	57.73	76.01	0	2,770	56.03	76.44	1	2,652
Total amount ($Th) of small business loans	2,397	4,307	0	139,875	2,483	4,722	2	123,345
Demand-adjusted bank competition	13.93	27.43	0	630.8	15.22	124.6	0	7,671
	Panel B: Demographic and business data							
	80%≤ MFI (N=10,190)				MFI< 80% (N=3,974)			
VARIABLES	mean	sd	min	max	mean	sd	min	max
Population	4,101	1,539	91	16,532	4,138	1,751	37	16,326
Minority population (%)	33.26	26.78	0.450	99.77	42.94	28.58	1.030	99.87
Tract MFI	50,867	13,487	20,833	200,001	45,322	10,668	2,499	85,912
Number of small businesses	158.0	167.7	5	5,774	157.9	171.8	5	5,515
Credit score	70.12	3.747	57.68	77.46	69.60	3.909	57.68	77.46
Years of establishment	16.91	4.666	5.231	49.80	17.23	4.851	5.914	50.56
Proprietorship/partnership legal form	0.728	0.0956	0.261	1	0.726	0.0969	0.250	1

6 Empirical results

6.1 CRA regulation and small business loan growth

The fundamental assumption of the fuzzy RD design in this study is that CRA eligibility increases the probability of small business loan growth. We begin the empirical investigation by providing explicit evidence in support of this assumption. In particular, we estimate the discontinuity jump in the probability of small business loan increase i.e. $\hat{\tau}_T = \mu_{T^-} - \mu_{T^+}$. Table 4 columns (1)-(2) present the results for the two samples, the unmatched and the matched samples respectively. The estimate of the discontinuity jump is statistically significant suggesting that the probability of a small business loan increase in eligible tracts is significantly higher compared to the probability of a small business loan increase in non-eligible tracts. Moreover, the results in Table 4 columns (3)-(6) from the median-split sample based on the demand-adjusted bank competition, reveal that the regulation is equally likely to enhance small business loan availability in census tracts with low and high demand-adjusted bank competition.

Table 4: Effect of CRA eligibility on the small business lending growth.

	All tracts		Low demand-adjusted bank competition tracts		High demand-adjusted bank competition tracts	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Conventional	0.262***	0.346***	0.268***	0.348***	0.255***	0.335***
	(0.0193)	(0.0199)	(0.0274)	(0.0265)	(0.0274)	(0.0284)
Bias-corrected	0.268***	0.350***	0.276***	0.347***	0.258***	0.341***
	(0.0193)	(0.0199)	(0.0274)	(0.0265)	(0.0274)	(0.0284)
Robust	0.268***	0.350***	0.276***	0.347***	0.258***	0.341***
	(0.0222)	(0.0229)	(0.0317)	(0.0306)	(0.0323)	(0.0327)
Observations	73366	14164	36683	7087	36683	7077
Robust 95% CI	[0.224 ; 0.312]	[0.305 ; 0.395]	[0.214 ; 0.338]	[0.287 ; 0.407]	[0.195 ; 0.322]	[0.277 ; 0.405]
Kernel Type	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov
BW Type	mserd	mserd	mserd	mserd	mserd	mserd
VCE method	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Total	Matched	Total	Matched	Total	Matched
Order Loc.Poly.(p)	1.000	1.000	1.000	1.000	1.000	1.000
Order Bias(q)	2.000	2.000	2.000	2.000	2.000	2.000
BW Loc.Poly.(h)	11.632	11.850	10.728	12.786	12.358	12.574
BW Bias (b)	21.348	21.852	19.146	23.191	21.042	23.588

Note: The outcome variable is the probability of a small business loan increase. Covariates include the tract's population, the proportion of minority populations, the median family income and the number of small and large bank branches. The demand-adjusted bank competition is measured by the number of branches located within the tract per 100 establishments. Conventional estimator of RD is based on the first order (linear) polynomial estimators in (5.2) and (5.3). (Robust) Bias-corrected is the first order polynomial RD estimator in (5.4). Reported bandwidths are percentage points of the MFI ratio. Standard errors in parentheses. Significance levels are ***1%, **5% and *10%.

The conclusion that demand-adjusted bank competition does not impede the effectiveness of the regulation in increasing the loan availability, is attributed to the fact that CRA compliance is monitored by multiple local groups. Specifically, if the branch is located in a census tract with relatively low number of branches, then compliance scrutiny is likely to be higher both by the regulator and the community, increasing the compliance incentives discussed earlier. If the branch is located in a census tract with a relatively high number of branches, then compliance scrutiny at bank level may be less intense but the coexistence of several banks increases the probability that some banks will opt to fund new loans in the eligible tract. Taken together, tract eligibility on average raises the likelihood of a small business loan increase.

Overall, the above findings are in line with the extant CRA literature which supports the notion that banks complied with the law by engaging more with local community organizations and shifting their lending activities to CRA-eligible neighborhoods. Specifically, Ding, Lee, and Bostic (2018) show that CRA promotes small business lending, especially the number of loan originations, while Agarwal, Benmelech, Bergman, and Seru (2012) find that in the six quarters surrounding a CRA exam, lending by banks is elevated on average by 5 percent.

6.2 Regulation and small businesses credit score

The expected increase in loans documented in the previous section does not imply automatically that small businesses' creditworthiness is enhanced. A necessary condition for the latter is that banks invest in information production so the new loans are given to entrepreneurs who, on average, have profitable projects. The theoretical model predicts that banks' incentives for information production are stronger under regulation (see Lemma 6). Specifically, regulation increases the likelihood that banks invest in screening technology and thus, they end up funding, on average, more entrepreneurs with profitable projects strengthening the small businesses' creditworthiness. In this section, we provide empirical evidence in support of the hypothesis that regulation, on average, increases the credit score of small businesses. The evidence rejects Hypothesis 1.

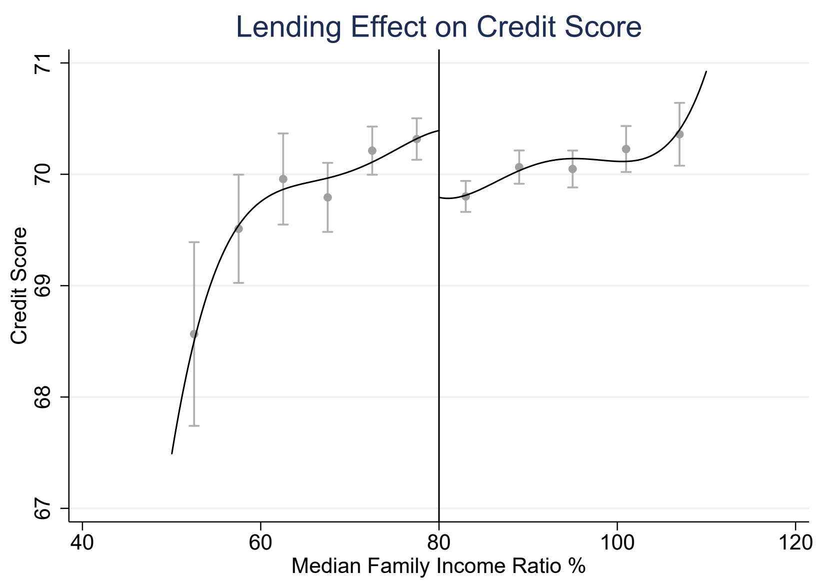

Figure 3: The credit score discontinuity at the cut off point of 80% in the median family income ratio

We begin with plotting the credit score distribution one year after the eligibility status over the support of the ratio of the census tract's median family income to the area median family income, using the evenly spaced method to select the number of bins. The RD plot in Figure 3 reveals that there is a discontinuity jump at the cut off point. In particular, the small businesses located in census tracts with income ratio lower than 80% have an average credit score, one year after the tracts eligibility, significantly higher than the average credit score of small businesses located in census tracts with income ratio higher than 80%.

Moreover, the lower value of the 95% confidence interval for the sample mean of the credit score within the bin close to but below the cut off is higher than the upper value of the 95% confidence interval for the sample mean of the credit score within the bin close to but above the cut off, suggesting that the jump is statistically significant at 5% level. Beside the discontinuity point at the cut off, the plot has an upward trend indicating that small businesses located in census tracts with higher income family ratio display on average higher credit score.

Table 5: Effect of CRA regulation on firm's credit score.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Conventional	0.824**	1.659***	1.362**	1.690***	1.893***	1.874***
	(0.415)	(0.498)	(0.630)	(0.414)	(0.479)	(0.630)
Bias-corrected	0.933**	1.610***	1.372**	1.777***	1.920***	1.922***
	(0.415)	(0.498)	(0.630)	(0.414)	(0.479)	(0.630)
Robust	0.933*	1.610***	1.372**	1.777***	1.920***	1.922***
	(0.483)	(0.568)	(0.692)	(0.484)	(0.554)	(0.682)
Observations	73366	73366	73366	14164	14164	14164
Robust 95% CI	[-0.014 ; 1.879]	[0.498 ; 2.723]	[0.016 ; 2.728]	[0.829 ; 2.724]	[0.834 ; 3.005]	[0.585 ; 3.258]
Kernel Type	Epanechnikov	Epanechnikov	Uniform	Epanechnikov	Epanechnikov	Uniform
BW Type	mserd	mserd	msecomb1	mserd	mserd	msecomb1
VCE method	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster
Covariates	No	Yes	Yes	No	Yes	Yes
Sample	Total	Total	Total	Matched	Matched	Matched
Order Loc.Poly.(p)	1.000	1.000	2.000	1.000	1.000	2.000
Order Bias(q)	2.000	2.000	3.000	2.000	2.000	3.000
BW Loc.Poly.(h)	14.347	14.117	15.346	12.408	10.636	11.016
BW Bias (b)	25.131	27.128	25.450	21.348	19.124	19.089

Note: The outcome variable is the one year forward tract-level average credit score of small businesses located in the census tract. Covariates include the tract's population, the proportion of minority populations, the median family income and the number of small and large bank branches. Conventional estimator of RD is based on the first order (linear) polynomial estimators in (5.2) and (5.3). (Robust) Bias-corrected is the first order polynomial fuzzy RD estimator in (5.5). Specifications (3) and (6) are based on the bias-corrected second order (quadratic) polynomial estimator. Reported bandwidths are percentage points of the MFI ratio. Standard errors in parentheses. Significance levels are ***1%, **5% and *10%.

Findings in Table 5 confirm that in eligible census tracts, the average credit score of the small businesses located within the community increases one year after the change in regulation eligibility status. The actual effect measured by the bias corrected local linear estimate of $\hat{\tau}_Y$ in equation (5.5) varies from 0.933 to 1.922 points depending on the sample, the inclusion of covariates and the choice of kernel and polynomial approximation, but it remains statistically singificant under all model specifications. The effect is equal to half of the credit scores' standard deviation and it is economically significant. According to paydex score definition, one point increase in the score implies approximately one to two days shorter delinquency. Thus, we conclude that regulation increases, on average, the incentives of banks to invest in information production and as a result, the increasing number of small business loans is channeled to entrepreneurs with profitable projects.

6.3 Regulation, bank competition and small businesses credit score

The effect of regulation on the small business credit score documented above applies to the average tract in our sample. In this section, we examine this effect conditional on the demand-adjusted bank competition within the tract. In particular, we use the median split of the demand-adjusted bank competition measure to create two subsamples: the tracts with a low bank competition (below the median) and the tracts with a high bank competition (above the median).

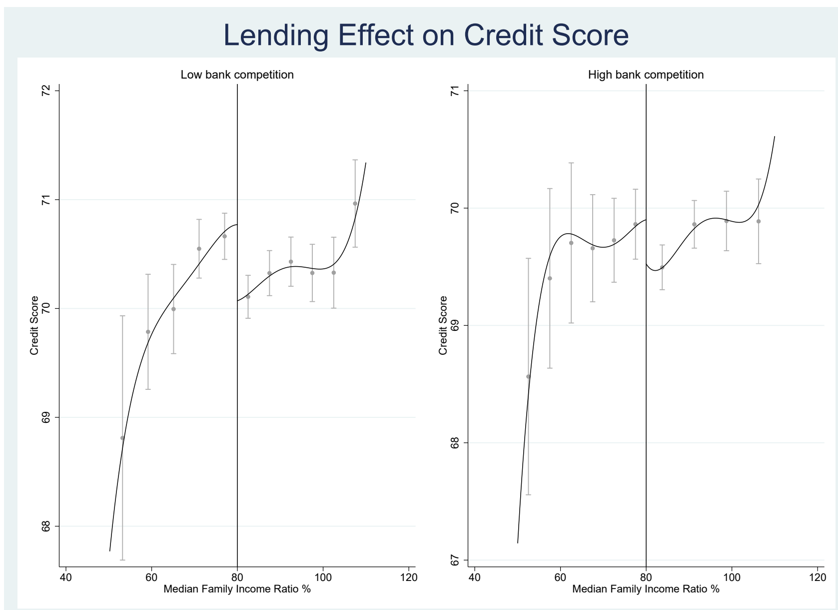

Figure 4: The credit score discontinuity and demand-adjusted bank competition

We begin with plotting the credit score distribution one year after the eligibility status over the ratio of the census tract's median family income to the area's median family income using the median split samples. The left RD plot in Figure 4 reveals that there is a discontinuity jump at the cut off point among tracts with low demand-adjusted bank competition. In particular, the lower value of the 95% confidence interval for the sample mean of the outcome variable within the bin close to but below the cut off is higher than the upper value of the 95% confidence interval for the sample mean of the outcome variable within the bin close to but above the cut off, suggesting that the jump is statistically significant at 5% level. In contrast, the right RD plot in Figure 4 reveals that the discontinuity jump at the cut off point among tracts with high bank competition is not statistically significant since the lower value of the 95% confidence interval for the sample mean of the outcome variable within the bin close to but below the cut off overlaps with the upper value of the 95% confidence interval for the sample mean of the outcome variable within the bin close to but above the cut off.

Table 6: Bank competition and the effect of CRA regulation on firm's credit score.

	Low demand-adjusted bank competition tracts		High demand adjusted bank competition tracts	
VARIABLES	(1)	(2)	(3)	(4)
Conventional	2.056***	2.244***	0.916	1.300*
	(0.611)	(0.535)	(0.856)	(0.760)
Bias-corrected	2.007***	2.191***	0.793	1.390*
	(0.611)	(0.535)	(0.856)	(0.760)
Robust	2.007***	2.191***	0.793	1.390
	(0.706)	(0.611)	(1.012)	(0.886)
Observations	36683	7087	36683	7077
Robust 95% CI	[0.623 ; 3.392]	[0.993 ; 3.389]	[-1.191 ; 2.777]	[-0.346 ; 3.127]
Kernel Type	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov
BW Type	mserd	mserd	mserd	mserd
VCE method	Cluster	Cluster	Cluster	Cluster
Covariates	Yes	Yes	Yes	Yes
Sample	Total	Matched	Total	Matched
Order Loc.Poly.(p)	1.000	1.000	1.000	1.000
Order Bias(q)	2.000	2.000	2.000	2.000
BW Loc.Poly.(h)	15.553	14.466	12.089	11.781
BW Bias (b)	27.987	27.784	20.271	20.843

Note: The outcome variable is the one year forward tract-level average credit score of small businesses located in the census tract. Covariates include the tract's population, the proportion of minority populations, the median family income and the number of small and large bank branches. The demand-adjusted bank competition is measured by the number of branches located within the tract per 100 establishments. Conventional estimator of RD is based on the first order (linear) polynomial estimators in (5.2) and (5.3). (Robust) Bias-corrected is the first order polynomial fuzzy RD estimator in (5.5). Reported bandwidths are percentage points of the MFI ratio. Standard errors in parentheses. Significance levels are ***1%, **5% and *10%.

Findings in Table 6 columns (1)-(2) confirm that in markets with low demand-adjusted bank competition, banks investing in information production, are able to fund the high value projects of the opaque entrepreneurs which reflects positively on the credit score of those entrepreneurs one year after the eligibility status. The actual effect on the credit score of these businesses exceeds 2 scale points, an economically significant increase of more than half of the credit scores' standard deviation. Moreover, Table 6 columns (3)-(4) show that the effect of regulation on the average credit score of the small businesses located in tracts with high demand-adjusted bank competition is not statistically significant. Equivalently, in high competition markets, there is no evidence that banks perform substantial screening. Instead, they are equally likely to fund high and low value projects of the opaque entrepreneurs. We conclude that, high competition erodes the banks' incentives to invest in screening due to the lower number of borrowers reachable by the competing banks. The findings reject Hypothesis 2. Hence, and given that we have already rejected Hypothesis 1, we conclude that the empirical evidence lends support to Hypothesis 3. Furthermore, the evidence is consistent with the theoretical model's prediction under the condition of intermediate cost of screening, see Proposition 3.

6.4 The role of firm's opacity

The theoretical underpinning of the effect of regulation on information production and the role of bank competition presumes the existence of two groups of entrepreneurs, the transparent and the opaque entrepreneurs. In reality, we observe the co-existence of entrepreneurs with varying degrees of informational opacity. Compared to the large, publicly listed corporations, small busines is a group of informationally opaque firms for many reasons (Berger and Udell (1998)). First, small businesses do not issue traded securities that are continuously priced in public markets. Moreover, many of the smallest firms do not have audited financial statements that can be shared with the bank. Finally, small firms do not enter into contracts that are publicly visible or widely reported in the press and contracts with their labor force, their suppliers, and their customers are generally kept private. As a result, small firms often cannot credibly convey their quality. But even within the class of small businesses there can be multiple levels of opacity. For example, Hyytinen and Pajarinen (2008) find that opacity, measured by disagreement between credit information companies, is inversely related to the age of firms. Younger firms do not have the track record and sufficient reputation to signal their intrinsic value.[31] Moreover, according to Bianco, Bontempi, Golinelli, and Parigi (2013) family firms are typically more opaque as a result of lower level of trust among agents and the desire to keep the firm's control within the family.

Based on the above observations, we create a construct of relative informational opacity among small businesses. In particular, we use age and legal form variables from the NETS database and we define as highly opaque the small businesses with age less than ten years and legal form of sole proprietorship or

[31] According to the U.S. Census Bureau's Annual Survey of Entrepreneurs young small businesses are more likely to tap into informal sources of credit such as funding from owners or family and friends, while older firms are more likely to receive funding from more traditional sources. The FED's 2017 Report to the Congress on the Availability of Credit to Small Businesses explains that this difference is likely tied to the greater informational opacity of new firms. This opacity might make evaluating creditworthiness more difficult for arms-length lenders, which could reduce the supply of more formal credit available to young firms.

Table 7: Effect of CRA eligibility on the small business lending growth for different levels of bank competition and firm informational opacity.

Panel A: Low demand-adjusted bank competition tracts						
	Low firm opacity tracts		Medium firm opacity tracts		High firm opacity tracts	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Conventional	0.288***	0.388***	0.335***	0.394***	0.207***	0.290***
	(0.0424)	(0.0473)	(0.0538)	(0.0539)	(0.0640)	(0.0607)
Bias-corrected	0.300***	0.399***	0.345***	0.395***	0.216***	0.296***
	(0.0424)	(0.0473)	(0.0538)	(0.0539)	(0.0640)	(0.0607)
Robust	0.300***	0.399***	0.345***	0.395***	0.216***	0.296***
	(0.0493)	(0.0547)	(0.0636)	(0.0629)	(0.0765)	(0.0717)
Observations	12227	2494	12175	2315	12281	2278
Robust 95% CI	[0.203 ; 0.396]	[0.292 ; 0.506]	[0.22 ; 0.469]	[0.272 ; 0.519]	[0.066 ; 0.366]	[0.155 ; 0.436]
Kernel Type	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov
BW Type	mserd	mserd	mserd	mserd	mserd	mserd
VCE method	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Total	Matched	Total	Matched	Total	Matched
Order Loc.Poly.(p)	1.000	1.000	1.000	1.000	1.000	1.000
Order Bias(q)	2.000	2.000	2.000	2.000	2.000	2.000
BW Loc.Poly.(h)	10.462	10.883	8.328	8.163	7.079	7.239
BW Bias (b)	18.264	19.550	13.632	14.126	10.938	11.737
Panel B: High demand-adjusted bank competition tracts						
	Low firm opacity tracts		Medium firm opacity tracts		High firm opacity tracts	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Conventional	0.318***	0.378***	0.248***	0.403***	0.148***	0.289***
	(0.0436)	(0.0465)	(0.0532)	(0.0595)	(0.0567)	(0.0631)
Bias-corrected	0.324***	0.381***	0.241***	0.425***	0.143**	0.278***
	(0.0436)	(0.0465)	(0.0532)	(0.0595)	(0.0567)	(0.0631)
Robust	0.324***	0.381***	0.241***	0.425***	0.143**	0.278***
	(0.0514)	(0.0536)	(0.0634)	(0.0686)	(0.0678)	(0.0743)
Observations	12030	2216	13270	2513	11383	2348
Robust 95% CI	[0.223 ; 0.424]	[0.276 ; 0.486]	[0.117 ; 0.365]	[0.291 ; 0.560]	[0.011 ; 0.276]	[0.132 ; 0.424]
Kernel Type	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov
BW Type	mserd	mserd	mserd	mserd	mserd	mserd
VCE method	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Total	Matched	Total	Matched	Total	Matched
Order Loc.Poly.(p)	1.000	1.000	1.000	1.000	1.000	1.000
Order Bias(q)	2.000	2.000	2.000	2.000	2.000	2.000
BW Loc.Poly.(h)	11.985	15.239	9.474	6.967	10.857	7.735
BW Bias (b)	20.549	28.564	15.074	12.059	17.321	13.287

Note: The outcome variable is the probability of a small business loan increase. Covariates include the tract's population, the proportion of minority populations, the median family income and the number of small and large bank branches. The demand-adjusted bank competition is measured by the number of branches located within the tract per 100 establishments. Firm opacity is measured by years of establishment (less than 10) and legal form (sole proprietorship/partnership). Conventional estimator of RD is based on the first order (linear) polynomial estimators in (5.2) and (5.3). (Robust) Bias-corrected is the first order polynomial RD estimator in (5.4). Reported bandwidths are percentage points of the MFI ratio. Standard errors in parentheses. Significance levels are ***1%, **5% and *10%.

partnership, which is more common among the family businesses. We then examine how the extent of highly opaque entrepreneurs observed within a tract interplays with bank's competition and how it affects the credit allocation emanated by the regulation. We do so by applying the median-split based on bank competition and then splitting further the subsamples based on three equal-centile intervals of the proportion of highly opaque small businesses located within each tract.

Initially, we examine whether the interaction of firm opacity with competition has an effect on the regulation-induced small business loans growth. If banks consider the extent of firm opacity to strategically choose which LMI tracts they should invest in, then failing to observed any gains in firm's credit score could be attributed to the variability of small business loans supply. The results in Table 7 reveal that the effect of regulation on the likelihood of a small business loan increase is significant, independent of the tract's demand-adjusted bank competition and proportion of highly opaque firms. Equivalently, banks do not seem to use any these two local market characteristics to circumvent the CRA regulation's burden, probably because they are not, ex-ante, aware of the actual extent of highly opaque firms within each tract before performing any screening.

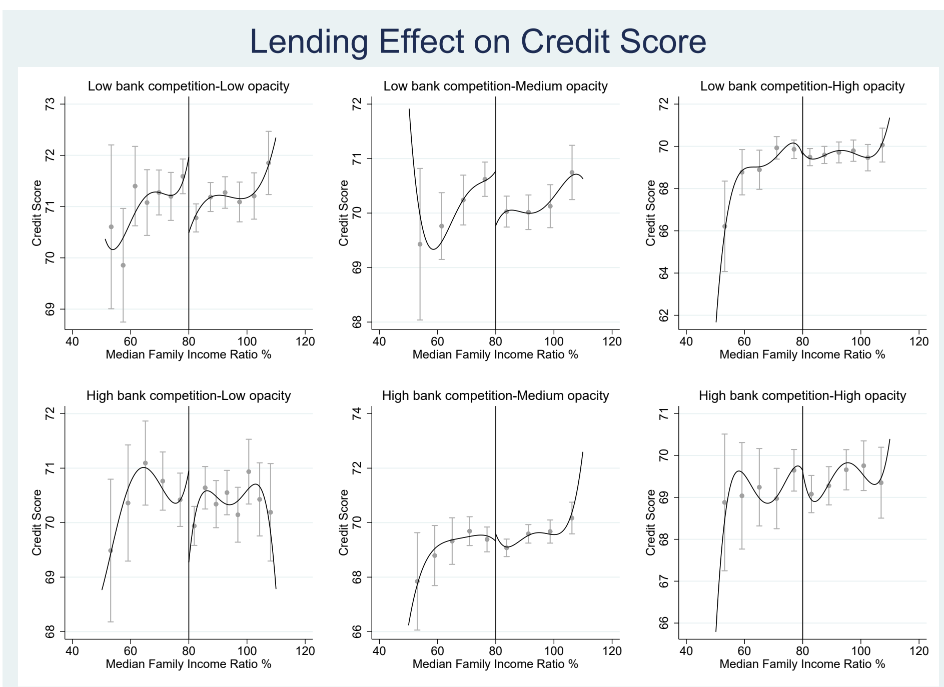

Figure 5: The credit score discontinuity, demand-adjusted bank competition and the firm opacity level

We continue by plotting the small businesses credit score across all different groups of tracts. Specifically, the top row of the RD plots in Figure 5 contains the low demand-adjusted bank competition tracts split into three different groups of highly opaque firms proportion. For the tracts with low and medium proportion of highly opaque firms, the lower value of the 95% confidence interval for the sample mean of the outcome variable within the bin close to but below the cut off is higher than the upper value of the 95% confidence interval for the sample mean of the outcome variable within the bin close to but above the cut off, suggesting that the jump in market efficiency is statistically significant at 5% level. In contrast, the RD plot of the tracts with high level of opaque firms reveals no discontinuity jump at the cut off point.

Similarly, the bottom row of the RD plots in Figure 5 contains the high demand-adjusted bank competition tracts split into the three different groups of highly opaque firms proportion. With the exception of the tracts with low proportion of highly opaque firms, which displays marginal evidence of a discontinuity jump in credit score at the cut off point, the other two plots with the medium and high proportions of highly opaque firms reveal no discontinuity in credit score. Taken together, the RD plots give us an indication of the interplay between bank competition and the proportion of highly opaque entrepreneurs which is directly linked to the screening effort required by the banks.

The findings presented in Table 8 confirm the role of borrowers' opacity and the implied screening costs on credit allocation. If banks have to choose between different groups of opaque firms, they will invest in the relatively less opaque group to minimize their costs. Thus, as the proportion of highly opaque small businesses and the implied screening costs increase, the effect on firms credit score declines and it does so faster for the markets with fewer borrowers relative to the number of competing banks. Specifically, Table 8 Panel A columns (1)-(4) confirm that in the low demand-adjusted competitive markets with low and moderate proportion of highly opaque businesses, banks are able to distinguish the high value projects leading to an improvement in the average firm credit score. However, Table 8 Panel A columns (5)-(6) show that even within low competitive markets, banks fail to allocate credit efficiently if the market is fraught with highly opaque entrepreneurs due to the increasing implied screening costs.

Moreover, Table 8 Panel B columns (1)-(2) show some evidence that in the high competitive markets with low proportion of highly opaque businesses, banks may be able to distinguish the high value projects. In contrast, Table 8 Panel B columns (3)-(6) show that as the proportion of highly opaque entrepreneurs increases, banks fail to distinguish the high value projects. The proportion of highly opaque firms that

Table 8: Firm opacity, bank competition and the effect of small business loans supply increase on firm's credit score.

	Panel A: Low demand-adjusted bank competition tracts					
	Low firm opacity tracts		Medium firm opacity tracts		High firm opacity tracts	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Conventional	1.827**	2.329***	2.005*	2.691***	0.723	1.450
	(0.751)	(0.706)	(1.031)	(0.927)	(2.191)	(1.543)
Bias-corrected	1.903**	2.299***	1.976*	2.438***	-0.0341	1.069
	(0.751)	(0.706)	(1.031)	(0.927)	(2.191)	(1.543)
Robust	1.903**	2.299***	1.976*	2.438**	-0.0341	1.069
	(0.871)	(0.816)	(1.194)	(1.075)	(2.574)	(1.793)
Observations	12227	2494	12175	2315	12281	2278
Robust 95% CI	[0.197 ; 3.61]	[0.701 ; 3.898]	[-0.364 ; 4.315]	[0.332 ; 4.545]	[-5.08 ; 5.012]	[-2.445 ; 4.582]
Kernel Type	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov
BW Type	mserd	mserd	mserd	mserd	mserd	mserd
VCE method	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Total	Matched	Total	Matched	Total	Matched
Order Loc.Poly.(p)	1.000	1.000	1.000	1.000	1.000	1.000
Order Bias(q)	2.000	2.000	2.000	2.000	2.000	2.000
BW Loc.Poly.(h)	14.712	13.562	10.535	11.060	7.872	8.504
BW Bias (b)	26.649	25.105	18.074	19.433	12.816	14.398
	Panel B: High demand-adjusted bank competition tracts					
	Low firm opacity tracts		Medium firm opacity tracts		High firm opacity tracts	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Conventional	1.187	1.834*	-0.543	-0.407	1.925	0.822
	(0.829)	(1.012)	(1.298)	(1.266)	(4.035)	(2.168)
Bias-corrected	1.085	2.295**	-1.262	-0.856	0.881	0.206
	(0.829)	(1.012)	(1.298)	(1.266)	(4.035)	(2.168)
Robust	1.085	2.295*	-1.262	-0.856	0.881	0.206
	(0.940)	(1.175)	(1.525)	(1.469)	(4.725)	(2.601)
Observations	12030	2216	13270	2513	11383	2348
Robust 95% CI	[-0.756 ; 2.927]	[-0.007 ; 4.597]	[-4.252 ; 1.728]	[-3.734 ; 2.023]	[-8.380 ; 10.142]	[-4.892 ; 5.305]
Kernel Type	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov
BW Type	mserd	mserd	mserd	mserd	mserd	mserd
VCE method	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Total	Matched	Total	Matched	Total	Matched
Order Loc.Poly.(p)	1.000	1.000	1.000	1.000	1.000	1.000
Order Bias(q)	2.000	2.000	2.000	2.000	2.000	2.000
BW Loc.Poly.(h)	16.134	11.842	14.188	7.436	8.852	7.369
BW Bias (b)	31.796	20.756	24.776	12.586	15.064	12.030

Note: The outcome variable is the one year forward tract-level average credit score of small businesses located in the census tract. Covariates include the tract's population, the proportion of minority populations, the median family income and the number of small and large bank branches. Demand-adjusted bank competition is measured by the number of branches located within the tract per 100 establishments. Firm opacity is measured by years of establishment (less than 10) and legal form (sole proprietorship/partnership). Conventional estimator of RD is based on the first order (linear) polynomial estimators in (5.2) and (5.3). (Robust) Bias-corrected is the first order polynomial fuzzy RD estimator in (5.5). Reported bandwidths are percentage points of the MFI ratio. Standard errors in parentheses. Significance levels are ***1%, **5% and *10%.

credit allocation efficiency disappears is less for high demand-adjusted competitive markets compared to low demand-adjusted competitive markets because the bank incentives for information production remain strong for the latter but not the former group of markets.

6.5 Falsification analysis

A main advantage of the RD design is that the mechanism by which regulatory eligibility is assigned is based on the ratio of tract's MFI over the area's MFI and none of these quantities can be manipulated by any of the involved parties. In particular, neither the local entrepreneurs nor banks have any influence on this objectively determined ratio. If involved parties are unable to precisely manipulate the assignment variable, a consequence is that the variation in regulatory eligibility near the threshold is randomized as though from a randomized experiment (Lee and Lemieux (2010)). Given the objective determination of the assignment variable we focus on the continuity assumption of the regression function near the cutoff, that ensures the validity of the RD design. Although the continuity assumption cannot be directly tested, the RD design offers some empirical methods that provide indirect evidence about the validity of the assumption. Following Cattaneo, Idrobo, and Titiunik (2019), we consider three empirical validation tests based on (i) the insignificant eligibility effect on pretreatment covariates (placebo outcomes), (ii) the eligibility effect observed at artificial cutoff values (placebo cut offs) and (iii) the exclusion of observations too close to the cutoff (the donut-hole analysis).

Specifically, the eligibility to regulation should not be linked to any predetermined characteristic of the census tracts. Since the eligibility effect on predetermined covariates is zero by construction, consistent evidence of non-zero effects on covariates that are likely to be confounders would raise questions about the validity of the RD assumption. For example, if tracts display significant differences at the cut off with respect to their population or to the number of entrepreneurs, that would indicate a discontinuity in the demand for bank loans, which could provide an alternative explanation of the findings. Similarly, if tracts display significant differences at the cut off with respect to the number of bank branches, that would indicate a discontinuity in access to bank credit. Therefore, we examine the significance of potential discontinuity jumps at the cut off for the model covariates, namely, the tract's population, number of small businesses, number of small bank branches and number of large bank branches. The p-values of the regulatory eligibility effect on the covariates in Table 9 columns (1)-(4) provide no evidence at 5% level of significance, that tracts display any discontinuity jumps near the cut off with respect to their size, the number of entrepreneurs and the number of small and large bank branches.

Table 9: Continuity-Based Analysis for Covariates

VARIABLES	(1) Population	(2) No of small businesses	(3) No of small banks	(4) No of large banks
Conventional p-value	0.0484	0.174	0.486	0.103
Bias-corrected p-value	0.0598	0.179	0.476	0.0851
Robust p-value	0.0708	0.204	0.500	0.0992
Observations	73366	73366	73366	73366
Kernel Type	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov
BW Type	cerrd	cerrd	cerrd	cerrd
VCE method	Cluster	Cluster	Cluster	Cluster
Order Loc.Poly.(p)	1.000	1.000	1.000	1.000
Order Bias(q)	2.000	2.000	2.000	2.000
BW Loc.Poly.(h)	6.384	6.964	6.215	5.001
BW Bias (b)	20.408	16.948	16.974	14.625

Note: Conventional estimator p-value is based on the first order (linear) polynomial estimators in (5.2) and (5.3) and (Robust) Bias-corrected p-value is based on the first order polynomial RD estimator in (5.4) with outcome variables the covariates of the tract's population, the proportion of minority populations and the number of small and large bank branches. Reported bandwidths are percentage points of the MFI ratio.

We continue with the falsification analysis by testing whether there are significant effects at placebo cutoff values. Evidence of discontinuity away from the cut off would imply that CRA regulatory eligibility is not the only "treatment" effect, casting doubt on our RD design. In the placebo cut off test, we replace the true cutoff value of 80% by another value at which the regulatory eligibility status does not really change, and perform estimation and inference using this placebo cutoff point. We choose as placebo cut off points the ratios of 70% and 90% i.e. one lower and one higher than the actual cut off point. In Table 10 columns (1) and (3), the estimated effects of the placebo cut off points in the probability of small business loans increase is not statistically significant. We can therefore conclude that the discontinuity in the probability of small business loans increase occurs only at the CRA eligibility threshold. Similarly, the estimated effects of the placebo cut off points in the small business credit score in Table 10 columns (2) and (4) are not statistically significant. We can therefore conclude that the discontinuity in the small business credit score is due to the increase in the small business loans induced by CRA eligibility.

We conclude the falsification analysis by investigating how sensitive the findings are to the existence of a few tracts who are located very close to the cut off point. This strategy is useful to assess the sensitivity of the results to the unavoidable extrapolation involved in local polynomial estimation, as the observations

Table 10: Placebo cut off

	Cut off 70%		Cut off 90%	
	(1)	(2)	(3)	(4)
VARIABLES	Credit supply	Credit score	Credit supply	Credit score
Conventional	-0.0501	10.52	-0.0144	-8.211
	(0.0353)	(9.625)	(0.0143)	(12.24)
Bias-corrected	-0.0457	13.42	-0.0142	-15.93
	(0.0353)	(9.625)	(0.0143)	(12.24)
Robust	-0.0457	13.42	-0.0142	-15.93
	(0.0420)	(11.44)	(0.0172)	(12.99)
Observations	73366	73366	73366	73366
Kernel Type	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov
BW Type	mserd	mserd	mserd	mserd
VCE method	Cluster	Cluster	Cluster	Cluster
Covariates	Yes	Yes	Yes	Yes
Order Loc.Poly.(p)	1.000	1.000	1.000	1.000
Order Bias(q)	2.000	2.000	2.000	2.000
BW Loc.Poly.(h)	10.833	10.753	5.533	5.529
BW Bias (b)	17.666	17.523	8.754	15.340

Note: The outcome variables are the probability of a small business loan increase and the one year forward tract-level average credit score of small businesses located in the census tract. Covariates include the tract's population, the proportion of minority populations, the median family income and the number of small and large bank branches. Conventional estimator of RD is based on the first order (linear) polynomial estimators in (5.2) and (5.3). (Robust) Bias-corrected is the first order polynomial RD estimators in (5.4) (Columns (1) and (3)) and the (5.5) (Columns (2) and (4)). Alternative cut off points of the median family income ratio are 70% and 90%. Standard errors in parentheses. Reported bandwidths are percentage points of the MFI ratio. Significance levels are ***1%, **5% and *10%.

Table 11: Continuity-Based Analysis for the Donut-Hole Approach

	<79% & >81%		<78% & >82%	
	(1)	(2)	(3)	(4)
VARIABLES	Credit supply	Credit score	Credit supply	Credit score
Conventional	0.256***	1.869***	0.246***	2.855**
	(0.0232)	(0.706)	(0.0296)	(1.145)
Bias-corrected	0.261***	1.872***	0.251***	2.943**
	(0.0232)	(0.706)	(0.0296)	(1.145)
Robust	0.261***	1.872**	0.251***	2.943**
	(0.0270)	(0.830)	(0.0351)	(1.373)
Observations	72051	72051	70769	70769
Kernel Type	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov
BW Type	mserd	mserd	mserd	mserd
VCE method	Cluster	Cluster	Cluster	Cluster
Covariates	Yes	Yes	Yes	Yes
Order Loc.Poly.(p)	1.000	1.000	1.000	1.000
Order Bias(q)	2.000	2.000	2.000	2.000
BW Loc.Poly.(h)	13.114	11.628	12.721	9.355
BW Bias (b)	29.455	24.655	30.736	21.145

Note: The outcome variables are the probability of a small business loan increase and the one year forward tract-level average credit score of small businesses located in the census tract. Covariates include the tract's population, the proportion of minority populations, the median family income and the number of small and large bank branches. Conventional estimator of RD is based on the first order (linear) polynomial estimators in (5.2) and (5.3). (Robust) Bias-corrected is the first order polynomial RD estimators in (5.4) (Columns (1) and (3)) and the (5.5) (Columns (2) and (4)). Tracts with the median family income ratio within 1% and 2% radius around the cut off point of 80% are excluded from the sample. Reported bandwidths are percentage points of the MFI ratio. Standard errors in parentheses. Significance levels are ***1%, **5% and *10%.

nearest to the cut off are more influential data points since they receive the higher kernel weights. If the exclusion of those tracts changes the significance of the results then the external validity of the findings is at doubt. Excluding such tracts and then repeating the estimation and inference analysis using the remaining sample is sometimes referred to as a "donut hole" approach. Initially we remove the tracts that are within a radius of 1% from the cut off point and then we repeat the test by removing tracts that are within a radius of 2% from the cut off point. The results in Table 11 columns (1) - (4) show that the estimated effects of the regulation on the probability of small business loans increase and on the small business credit score remain statistically significant and economically equivalent to the main results, even after removing approximately 1.5% and 3.4% of the nearest to the cut off point, and thus more influential, tracts from our sample.

7 Conclusion

Do government initiatives that support credit access to marginalized groups of borrowers promote credit allocation efficiency? We answer this question by providing evidence of the positive impact of CRA's small business lending program on the credit score of businesses located in low and moderate income communities. We conclude that without regulatory policies, borrowers with higher implied screening costs will remain underfunded. Implementing a new policy however is not a panacea. The program's success depends on the demand-adjusted bank competition of the local market. When there are fewer borrowers for each competing branch, the incentives of banks to invest in information production are weaker, inducing banks to lend to entrepreneurs with both high and low value projects. Effectively, for bank branches located in markets with many rival banks relative to the number of borrowers, policies like CRA represent a regulatory tax that they are willing to accept. This is because the alternative for a bank is to produce information that allows it to identify the opaque borrowers with high value projects, a strategy that does not yield enough additional profits to cover the screening cost.

The findings of the study lead us to some policy recommendations. Specifically, lowering screening technology cost is pivotal in increasing the efficacy of government intervention programs. This could be achieved through incentivizing banks to invest, similar to SBA loan programs that presume a certain level of screening by the lender if it is to receive the benefits of the loan guarantee. It can also be achieved through the promotion of competition among data intermediaries. More competition in the small business credit reporting space can be achieved if participants broaden the coverage of their data or by the entrance of new small business credit repositories.

Further investigation is warranted on the implications of merger-induced branch decline on the regulation-competition nexus. On the face of it, branch closures mean more borrowers for each remaining bank branches which, based on the above discussion, should increase the efficiency of credit allocation. However, this seemingly beneficial development is likely to exacerbate the information opacity of the local communities. Consequently, the incentives of higher earned rents among the remaining bank branches, could be offset by the increase in the information production cost due to lessened information externalities. Thus, empirical investigation is needed to reveal the net effect of merger induced branch consolidation in local credit markets. Moreover, the empirical findings of this study support the notion of moderate information technology costs although our data did not allow for measurement of the screening technology. Future research could focus on measuring screening technology explicitly to examine if the decreasing information technology costs has helped programs like CRA to become more successful in promoting the efficient credit allocation.

Finally, further research is warranted on the CRA performance. Because the regulation does not induce screening everywhere there is “money left on the table”. Measuring this inefficiency is critical for the overall evaluation of the program. More generally, our empirical investigation sets the stage for more research on the assessment of programs similar to CRA based on their economic outcomes, such as unemployment or economic growth, rather than their outputs like the volume of loans.

A Appendix: Detailed Proofs of Lemmas and Propositions

A.1 Derivations of the unilateral incentives to invest in an unregulated market

We derive a bank’s profit difference between when it invests and when it does not invest, both when all rival banks do not invest and when they invest.

A.1.1 Unilateral incentives to invest when all rival banks do not invest

We assume that $\phi = 0$ and banks, following Assumption 1, only fund T entrepreneurs. We substitute (3.8) and (3.2) into (3.5) to obtain

$$\frac{1 - \frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}q}{1 - \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q} = \frac{B_H^d}{E_H^d} \Rightarrow \frac{1 - \frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}q}{1 - \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q} = \frac{(1+i) - (1+\bar{i})}{H - (1+i)}. \tag{A.1}$$

Let

$$A_T(\theta) \equiv \frac{1 - \frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}q}{1 - \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q}. \tag{A.2}$$

It can be verified that as θ increases, i.e., more banks relative to entrepreneurs, $A_T(\theta)$ decreases.[32] We solve (A.1) with respect to i to derive the equilibrium interest rate

$$i_T^\star = \frac{A_T(\theta)(H-1) + \bar{i}}{1 + A_T(\theta)}. \tag{A.3}$$

It can be easily verified that $i_T^\star$ is increasing in A_T and hence it is decreasing in θ. As demand-adjusted bank competition intensifies, the interest rate, naturally, decreases. A bank's expected profit, using the outside option B_T^o given by (3.8), is

$$E\pi^{NR}(\phi = 0) = B_T^o = \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)} q(i_T^\star - \bar{i}) = \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)} \frac{qA_T(\theta)(H - (1+\bar{i}))}{1 + A_T(\theta)}, \tag{A.4}$$

where, as we have already discussed above, the outside option captures the bank's expected value from search in this market and NR stands for No Regulation.

We now examine a bank's incentives to unilaterally deviate to $\phi = 1$. Since when a bank invests in screening technology the signals are perfect, the bank also funds high value projects in the O group, so the probability of a deal becomes $q + (1-q)\mu$. The outside option of the bank is now given by (3.9), but the outside option of the entrepreneur is still (3.8). The entrepreneur, as we argue in Section 4.1.3, knows that the current bank has unilaterally deviated from an equilibrium where all banks choose $\phi = 0$. Hence, his outside option is consistent with this belief. Also, no other bank or entrepreneur observe this deviation.

[32] As $\theta \to 0$, $A_T(\theta) \to \frac{1}{1-q}$ and as $\theta \to \infty$, $A_T(\theta) \to 1 - q$.

Following similar steps as in the derivation of $A_T(\theta)$, we now have

$$A_H^{dev}(\theta) \equiv \frac{1-\frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}q}{1-\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}(q+(1-q)\mu)}, \tag{A.5}$$

where dev stands for deviation to indicate that the bank funds all high value projects when all rival banks do not.[33] The interest rate becomes

$$i_H^{dev} = \frac{A_H^{dev}(\theta)(H-1)+\bar{i}}{1+A_H^{dev}(\theta)}. \tag{A.6}$$

Note that $A_H^{dev}(\theta) > A_T(\theta)$, which suggests that $i_H^{dev} > i_T^{\star}$. When a bank has the ability to screen high value projects in the O group, when all other banks do not, it possesses a higher outside option which is reflected in the interest rate. As with $A_T(\theta)$, $A_H^{dev}(\theta)$ is decreasing in θ. Let $\hat{\phi}$ denote the symmetric choice of all *other* banks. The expected profits when a bank unilaterally invests in the screening technology are

$$\begin{aligned} E\pi^{NR}(\phi=1|\hat{\phi}=0) &= B_H^o = \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}(q+(1-q)\mu)(i_H^{dev}-\bar{i}) \\ &= \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}\frac{(q+(1-q)\mu))A_H^{dev}(\theta)(H-(1+\bar{i}))}{1+A_H^{dev}(\theta)}. \end{aligned} \tag{A.7}$$

The unilateral incentive to invest in the screening technology when the regulation is not in force is the difference between the profits when $\phi = 1$ and the profits when $\phi = 0$, which is given by

$$\begin{aligned} &\Delta^{NR}(\theta|\hat{\phi}=0) = E\pi^{NR}(\phi=1|\hat{\phi}=0) - E\pi^{NR}(\phi=0) \\ = \;& \frac{\alpha_b(\theta)\left(H-(1+\bar{i})\right)}{r+\alpha_b(\theta)}\left[\frac{q(A_H^{dev}(\theta)-A_T(\theta))}{(1+A_H^{dev}(\theta))(1+A_T(\theta))} + \frac{(1-q)\mu A_H^{dev}(\theta)}{(1+A_H^{dev}(\theta))}\right]. \end{aligned} \tag{A.8}$$

A.1.2 Unilateral incentives to invest when all rival banks invest

When all banks choose $\phi = 1$ the equilibrium interest rate, following similar steps as in Section 4.1.1, is

$$i_H^{\star} = \frac{A_H(\theta)(H-1)+\bar{i}}{1+A_H(\theta)}, \tag{A.9}$$

[33] As $\theta \to 0$, $A_H^{dev}(\theta) \to \frac{1}{1-(q+(1-q)\mu)}$ and as $\theta \to \infty$, $A_H^{dev}(\theta) \to 1-q$.

where[34,35]

$$A_H(\theta) \equiv \frac{1 - \frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}(q + (1-q)\mu)}{1 - \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}(q + (1-q)\mu)}. \tag{A.10}$$

The bank's ex-ante expected profit is

$$\begin{aligned} E\pi^{NR}(\phi = 1) &= B_H^o = \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)} \left[(q + (1-q)\mu)(i_H^\star - \bar{i})\right] \\ &= \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)} \left[\frac{(q + (1-q)\mu)A_H(\theta)(H - (1 + \bar{i}))}{1 + A_H(\theta)}\right]. \end{aligned} \tag{A.11}$$

When a bank unilaterally deviates to $\phi = 0$ it funds T entrepreneurs only. Its market bargaining power becomes[36,37]

$$A_T^{dev}(\theta) \equiv \frac{1 - \frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}(q + (1-q)\mu)}{1 - \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q}. \tag{A.12}$$

The interest rate is given by

$$i_T^{dev} = \frac{A_T^{dev}(\theta)(H-1) + \bar{i}}{1 + A_T^{dev}(\theta)}.$$

Its expected deviation profits are

$$\begin{aligned} E\pi^{NR}(\phi = 0|\hat{\phi} = 1) &= B_T^o = \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)} q(i_T^{dev} - \bar{i}) \\ &= \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)} \frac{qA_T^{dev}(\theta)(H - (1 + \bar{i}))}{1 + A_T^{dev}(\theta)}. \end{aligned} \tag{A.13}$$

The reduction in expected profits when a bank unilateral does not invest in the screening technology when the regulation is not in force is given by

$$\begin{aligned} &\Delta^{NR}(\theta|\hat{\phi} = 1) = E\pi^{NR}(\phi = 1) - E\pi^{NR}(\phi = 0|\hat{\phi} = 1) \\ = \quad &\frac{\alpha_b(\theta)\left(H - (1 + \bar{i})\right)}{r + \alpha_b(\theta)} \left[\frac{q(A_H(\theta) - A_T^{dev}(\theta))}{(1 + A_H(\theta))(1 + A_T^{dev}(\theta))} + \frac{(1-q)\mu A_H(\theta)}{(1 + A_H(\theta))}\right]. \end{aligned}$$

[34]Observe that $A_H^{dev}(\theta) > A_H(\theta)$, because although in both cases the bank funds all high value projects, under a deviation it is the only one that does so. Hence, its market power is higher.

[35]As $\theta \to 0$, $A_H(\theta) \to \frac{1}{1-(q+(1-q)\mu)}$ and as $\theta \to \infty$, $A_H(\theta) \to 1 - (q + (1-q)\mu)$.

[36]Note that $A_T^{dev}(\theta) < A_T(\theta)$.

[37]As $\theta \to 0$, $A_T^{dev}(\theta) \to \frac{1}{1-q}$ and as $\theta \to \infty$, $A_T^{dev}(\theta) \to 1 - (q + (1-q)\mu)$.

References

Agarwal, S., E. Benmelech, N. Bergman, and A. Seru (2012). Did the community reinvestment act (CRA) lead to risky lending? Technical report, National Bureau of Economic Research.

Agarwal, S. and R. Hauswald (2010). Distance and private information in lending. *The Review of Financial Studies 23*(7), 2757–2788.

Allen, F. and D. Gale (2004). Competition and financial stability. *Journal of Money, Credit and Banking*, 453–480.

Avery, R. B., R. W. Bostic, and G. B. Canner (2005). Assessing the necessity and efficiency of the community reinvestment act. *Housing Policy Debate 16*(1), 143–172.

Avery, R. B. and K. P. Brevoort (2015). The subprime crisis: Is government housing policy to blame? *Review of Economics and Statistics 97*(2), 352–363.

Begley, T. A. and A. Purnanandam (2021). Color and credit: Race, regulation, and the quality of financial services. *Journal of Financial Economics*.

Berger, A. N. and L. K. Black (2011). Bank size, lending technologies, and small business finance. *Journal of Banking & Finance 35*(3), 724–735.

Berger, A. N., N. H. Miller, M. A. Petersen, R. G. Rajan, and J. C. Stein (2005). Does function follow organizational form? evidence from the lending practices of large and small banks. *Journal of Financial Economics 76*(2), 237–269.

Berger, A. N., A. Saunders, J. M. Scalise, and G. F. Udell (1998). The effects of bank mergers and acquisitions on small business lending. *Journal of Financial Economics 50*(2), 187–229.

Berger, A. N. and G. F. Udell (1998). The economics of small business finance: The roles of private equity and debt markets in the financial growth cycle. *Journal of Banking & Finance 22*(6-8), 613–673.

Berger, A. N. and G. F. Udell (2006). A more complete conceptual framework for sme finance. *Journal of Banking & Finance 30*(11), 2945–2966.

Bhutta, N. (2011). The community reinvestment act and mortgage lending to lower income borrowers and neighborhoods. *The Journal of Law and Economics 54*(4), 953–983.

Bianco, M., M. E. Bontempi, R. Golinelli, and G. Parigi (2013). Family firms' investments, uncertainty and opacity. *Small Business Economics 40*(4), 1035–1058.

Binmore, K., A. Rubinstein, and A. Wolinsky (1986). The Nash bargaining solution in economic modelling. *The RAND Journal of Economics*, 176–188.

Black, S. E. and P. E. Strahan (2002). Entrepreneurship and bank credit availability. *The Journal of Finance 57*(6), 2807–2833.

Bostic, R. W. and H. Lee (2017). Small business lending under the community reinvestment act. *Cityscape 19*(2), 63–84.

Boyd, J. H. and G. De Nicolo (2005). The theory of bank risk taking and competition revisited. *The Journal of Finance 60*(3), 1329–1343.

Broecker, T. (1990). Credit-worthiness tests and interbank competition. *Econometrica*, 429–452.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica 82*(6), 2295–2326.

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2019). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press.

Cetorelli, N. and P. E. Strahan (2006). Finance as a barrier to entry: Bank competition and industry structure in local us markets. *The Journal of Finance 61*(1), 437–461.

Cole, R. A., L. G. Goldberg, and L. J. White (2004). Cookie cutter vs. character: The micro structure of small business lending by large and small banks. *Journal of Financial and Quantitative Analysis*, 227–251.

Dahl, D., D. D. Evanoff, and M. F. Spivey (2010). The community reinvestment act and targeted mortgage lending. *Journal of Money, Credit and Banking 42*(7), 1351–1372.

Dell'Ariccia, G. and R. Marquez (2004). Information and bank credit allocation. *Journal of Financial Economics 72*(1), 185–214.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *The Review of Economic Studies 51*(3), 393–414.

Ding, L., H. Lee, and R. Bostic (2018). Effects of the community reinvestment act (CRA) on small business lending. *FRB of Philadelphia Working Paper*.

Ding, L. and L. Nakamura (2020). Don't know what you got till it's gone - the community reinvestment act in a changing financial landscape. *The Community Reinvestment Act in a Changing Financial Landscape (February, 2020). FRB of Philadelphia Working Paper* (20-08).

Greenstone, M., A. Mas, and H.-L. Nguyen (2020). Do credit market shocks affect the real economy? Quasi-experimental evidence from the great recession and "normal" economic times. *American Economic Journal: Economic Policy 12*(1), 200–225.

Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient markets. *The American Economic Review 70*(3), 393–408.

Hauswald, R. and R. Marquez (2003). Information technology and financial services competition. *The Review of Financial Studies 16*(3), 921–948.

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

Heckman, J. J., R. J. LaLonde, and J. A. Smith (1999). The economics and econometrics of active labor market programs. In *Handbook of Labor Economics*, Volume 3, pp. 1865–2097. Elsevier.

Hellmann, T. F., K. C. Murdock, and J. E. Stiglitz (2000). Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough? *American Economic Review 90*(1), 147–165.

Hyytinen, A. and M. Pajarinen (2008). Opacity of young businesses: Evidence from rating disagreements. *Journal of Banking & Finance 32*(7), 1234–1241.

Inderst, R. and H. M. Müller (2004). The effect of capital market characteristics on the value of start-up firms. *Journal of Financial Economics 72*(2), 319–356.

James, C. (1987). Some evidence on the uniqueness of bank loans. *Journal of Financial Economics 19*(2), 217–235.

Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature 48*(2), 281–355.

Levine, R., C. Lin, Q. Peng, and W. Xie (2020). Communication within banking organizations and small business lending. *The Review of Financial Studies 33*(12), 5750–5783.

Marquez, R. (2002). Competition, adverse selection, and information dispersion in the banking industry. *The Review of Financial Studies 15*(3), 901–926.

Martinez-Miera, D. and R. Repullo (2010). Does competition reduce the risk of bank failure? *The Review of Financial Studies 23*(10), 3638–3664.

Petersen, M. A. and R. G. Rajan (1995). The effect of credit market competition on lending relationships. *The Quarterly Journal of Economics 110*(2), 407–443.

Petersen, M. A. and R. G. Rajan (2002). Does distance still matter? The information revolution in small business lending. *The Journal of Finance 57*(6), 2533–2570.

Pissarides, C. A. (2000). *Equilibrium unemployment theory*. MIT press.

Rajan, R. G. (1992). Insiders and outsiders: The choice between informed and arm's-length debt. *The Journal of Finance 47*(4), 1367–1400.

Ramakrishnan, R. T. and A. V. Thakor (1984). Information reliability and a theory of financial intermediation. *The Review of Economic Studies 51*(3), 415–432.

Rice, T. and P. E. Strahan (2010). Does credit competition affect small-firm finance? *The Journal of Finance 65*(3), 861–889.

Rogerson, R., R. Shimer, and R. Wright (2005). Search-theoretic models of the labor market: A survey. *Journal of Economic Literature 43*(4), 959–988.

Ruckes, M. (2004). Bank competition and credit standards. *The Review of Financial Studies 17*(4), 1073–1102.

Sharpe, S. A. (1990). Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *The Journal of Finance 45*(4), 1069–1087.

Silveira, R. and R. Wright (2016). Venture capital: A model of search and bargaining. *Review of Economic Dynamics 19*, 232–246.

B Internet appendix

B.1 Proof of Lemma 2

We examine how $\Delta^{NR}(\theta|\hat{\phi}=0)$, given by (4.1), is affected by θ.[38] We begin by noting that the $\frac{A(\theta)}{1+A(\theta)}$ terms are decreasing in θ. The $\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}$ term is also decreasing in θ. The only ambiguity comes from the $\frac{A_H^{dev}(\theta)}{1+A_H^{dev}(\theta)} - \frac{A_T(\theta)}{1+A_T(\theta)}$ term, since both of the term's components are decreasing in θ and the question is which one is decreasing faster.

When $q=0$, the derivative of $\frac{A_T(\theta)}{1+A_T(\theta)}$ with respect to θ is 0, while the derivative of $\frac{A_H^{dev}(\theta)}{1+A_H^{dev}(\theta)}$ is strictly negative as long as $\mu>0$. This implies that $\Delta E\pi^{NR}(\theta|\hat{\phi}=0)$ is decreasing in θ. By continuity, this should hold at least in the neighborhood of $q=0$.

When $q=1$, on the other hand, the slopes of $\frac{A_T(\theta)}{1+A_T(\theta)}$ and $\frac{A_H^{dev}(\theta)}{1+A_H^{dev}(\theta)}$ with respect to θ are equal. However, for q's very close to 1 and at $\theta=0$, which implies that $\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}$ is close to 1 and $\frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}$ is close to 0, the derivative of $\frac{A_H^{dev}(\theta)}{1+A_H^{dev}(\theta)} - \frac{A_T(\theta)}{1+A_T(\theta)}$ with respect to θ is positive. To demonstrate this, we differentiate $\frac{A_H^{dev}(\theta)}{1+A_H^{dev}(\theta)} - \frac{A_T(\theta)}{1+A_T(\theta)}$ with respect to θ and q and we evaluate the derivative at $q=1$. This yields,

$$\begin{aligned}
&\left.\frac{\partial^2\left(\frac{A_H^{dev}(\theta)}{1+A_H^{dev}(\theta)} - \frac{A_T(\theta)}{1+A_T(\theta)}\right)}{\partial\theta\partial q}\right|_{q=1} = -\mu\left(3\frac{d\left(\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}\right)}{d\theta} + \frac{d\left(\frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}\right)}{d\theta}\right) \\
= \;& -\mu\left(3\frac{r}{(r+\alpha_b(\theta))^2}\frac{d\alpha_b(\theta)}{d\theta} + \frac{r}{(r+\alpha_e(\theta))^2}\frac{d\alpha_e(\theta)}{d\theta}\right) \\
& \text{letting } \alpha_b \to \infty \& \alpha_e \to 0 \text{ as } \theta \to 0 \\
= \;& -\mu\left(\frac{d\alpha_e(\theta)}{d\theta}\right) < 0,
\end{aligned} \tag{B.1}$$

since $\frac{d\alpha_e(\theta)}{d\theta} > 0$. Recall that the derivative of $\frac{A_H(\theta)}{1+A_H(\theta)} - \frac{A_T(\theta)}{1+A_T(\theta)}$ with respect to θ at $q=1$ is zero. Then, the negative derivative with respect to q suggests that for q's close to 1 the derivative of $\frac{A_H^{dev}(\theta)}{1+A_H^{dev}(\theta)} - \frac{A_T(\theta)}{1+A_T(\theta)}$ with respect to θ is positive at $\theta=0$. By continuity, it should be positive in the neighborhood of zero. This in turn can make the derivative of $\Delta E\pi^{NR}(\theta)$ with respect to θ positive.

[38] The proof for $\Delta^{NR}(\theta|\hat{\phi}=1)$, given by (4.2), is similar and is omitted.

When, on the other hand, $\theta \to \infty$, $\alpha_e \to \infty$ and $\alpha_b \to 0$. Following similar steps as above

$$\left.\frac{\partial^2 \left(\frac{A_H^{dev}(\theta)}{1+A_H^{dev}(\theta)} - \frac{A_T(\theta)}{1+A_T(\theta)}\right)}{\partial\theta\partial q}\right|_{q=1\&\theta\to\infty} = -\mu\left(3\frac{d\alpha_b(\theta)}{d\theta}\right) > 0,$$

since $\frac{d\alpha_b(\theta)}{d\theta} < 0$. Hence, for high enough θ, $\Delta E\pi^{NR}(\theta|\hat{\phi} = 0)$ is definitely decreasing in θ.

B.2 Derivations of the unilateral incentives to invest in a regulated market

We derive a bank's profit difference between when it invests and when it does not invest, both when all rival banks do not invest and when they invest. The derivations below follow the same steps as the ones in Section A.1, the only difference is the banks' profits when $\phi = 0$, since in this case they are forced by the regulation to fund projects that are unprofitable in expectation.

B.2.1 Unilateral incentives to invest when all rival banks do not invest

We begin with the $\phi = 0$ equilibrium, so the signal s is uninformative. Suppose the entrepreneur is in the O group. Given that the surplus in this case is negative, see Assumption 1, the Nash bargaining problem is not valid. The interest rate must yield zero expected profit to the entrepreneur (equal to his participation constraint) and negative to the bank, that is, $i_O^R = H - 1$, where R stands for regulation. When $v = L = 0$, the entrepreneur defaults. This happens with probability $(1-q)(1-\mu)$. The bank's deal value is $B_O^d = \mu H - (1+\bar{i}) < 0$ and the entrepreneur's is $E_O^d = 0$. Now let's assume the entrepreneur is in the T group. We substitute (3.6), (3.7), $B_O^d = \mu H - (1+\bar{i}) < 0$, $E_O^d = 0$ and (3.2) into (3.5) to obtain

$$\begin{aligned}
\left(1 - \frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}q\right) E_H^d &= \left(1 - \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q\right) B_H^d - (1-q)\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)} B_O^d \\
\Rightarrow \quad \frac{\left(1 - \frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}q\right) + (1-q)\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}(\mu H - (1+\bar{i}))}{1 - \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q} &= \frac{B_H^d}{E_H^d} \\
\Rightarrow \quad \frac{\left(1 - \frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}q\right) + (1-q)\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}(\mu H - (1+\bar{i}))}{1 - \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q} &= \frac{i - \bar{i}}{H - (1+i)}.
\end{aligned} \tag{B.2}$$

Let

$$
\begin{aligned}
A(\theta) &\equiv \frac{\left(1-\frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}q\right)}{1-\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q}+\frac{(1-q)\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}(\mu H-(1+\bar{i}))}{1-\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q} \\
&= A_T(\theta)+\frac{(1-q)\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}(\mu H-(1+\bar{i}))}{1-\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q}. \qquad \text{(B.3)}
\end{aligned}
$$

Therefore, and given that the second term in the above expression is negative, $A_H(\theta) > A_T(\theta) > A(\theta)$. Then, solving (B.2) with respect to i, yields the interest rate for the T group is

$$
i_T^R = \frac{A(\theta)(H-1)+\bar{i}}{1+A(\theta)}. \qquad \text{(B.4)}
$$

By comparing (A.3) with (B.4), we can verify that $i_T^R < i_T^\star$, because $A(\theta) < A_T(\theta)$. The bank has a worse outside option when it must fund the O group and it cannot screen the high from the low value projects, than when such a requirement is absent. As a result it lowers the interest rate to the T group. The bank's ex-ante expected profit, assuming all banks lend to all entrepreneurs, is

$$
\begin{aligned}
E\pi^R(\phi=0) &= B^o = \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}\left[q(i_T^R-\bar{i})+(1-q)(\mu(1+i_O^R)-(1+\bar{i}))\right] \\
&= \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}\left[\frac{qA(\theta)(H-(1+\bar{i}))}{1+A(\theta)}+(1-q)(\mu H-(1+\bar{i}))\right]. \qquad \text{(B.5)}
\end{aligned}
$$

The deviation profit, $E\pi^R(\phi=1|\hat{\phi}=0)$, is the same as $E\pi^{NR}(\phi=1|\hat{\phi}=0)$ given by (A.7). This is because the outside options are the same, which follows from: i) the T group is not affected and ii) under regulation $E_O^d=0$, while under no regulation the O group does not get funded, when $\hat{\phi}=0$, so expected profit is still zero.[39] The incentive to unilaterally invest in the screening technology when the regulation is in force is given by

$$
\begin{aligned}
& \Delta^R(\theta|\hat{\phi}=0) = E\pi^R(\phi=1|\hat{\phi}=0) - E\pi^R(\phi=0) \\
= \; & \frac{\alpha_b(\theta)(H-(1+\bar{i}))}{r+\alpha_b(\theta)}\left[q\left(\frac{A_H^{dev}(\theta)}{1+A_H^{dev}(\theta)} - \frac{A(\theta)}{1+A(\theta)}\right) + (1-q)\left(\frac{\mu A_H^{dev}(\theta)}{1+A_H^{dev}(\theta)} - \frac{\mu H-(1+\bar{i})}{H-(1+\bar{i})}\right)\right] \\
= \; & \frac{\alpha_b(\theta)(H-(1+\bar{i}))}{r+\alpha_b(\theta)}\left[\frac{q(A_H^{dev}(\theta)-A(\theta))}{(1+A_H^{dev}(\theta))(1+A(\theta))} + (1-q)\left(\frac{\mu A_H^{dev}(\theta)}{1+A_H^{dev}(\theta)} - \frac{\mu H-(1+\bar{i})}{H-(1+\bar{i})}\right)\right].
\end{aligned}
$$

B.2.2 Unilateral incentives to invest when all rival banks invest

We now examine the conditions required to sustain $\phi = 1$ as a symmetric equilibrium. Assume all banks choose $\phi = 1$. Because regulation is not binding when $\phi = 1$, the equilibrium profit $E\pi^R(\phi = 1)$ is the same as $E\pi^{NR}(\phi = 1)$ given by (A.11). The interest rate is derived as follows. We substitute (3.6), (3.9), $B_O^d = \mu H - (1+\bar{i}) < 0$, $E_O^d = 0$ and (3.2) into (3.5) to obtain

$$
\begin{aligned}
& \left(1-\frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}(q+(1-q)\mu)\right)E_H^d = \left(1-\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q\right)B_H^d - (1-q)\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}B_O^d \\
\Rightarrow \; & \frac{\left(1-\frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}(q+(1-q)\mu)\right) + (1-q)\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}(\mu H-(1+\bar{i}))}{1-\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q} = \frac{B_H^d}{E_H^d} \\
\Rightarrow \; & \frac{\left(1-\frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}(q+(1-q)\mu)\right) + (1-q)\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}(\mu H-(1+\bar{i}))}{1-\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q} = \frac{i-\bar{i}}{H-(1+i)}.
\end{aligned} \tag{B.6}
$$

Let

$$
\begin{aligned}
A^{dev}(\theta) \;&\equiv\; \frac{\left(1-\frac{\alpha_e(\theta)}{r+\alpha_e(\theta)}(q+(1-q)\mu)\right)}{1-\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q} + \frac{(1-q)\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}(\mu H-(1+\bar{i}))}{1-\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q} \\
&=\; A_T^{dev}(\theta) + \frac{(1-q)\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}(\mu H-(1+\bar{i}))}{1-\frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}q}.
\end{aligned} \tag{B.7}
$$

[39] The entrepreneur's alternative to bank funding can also yield a positive expected profit. In this case the bank's profit will be even more negative and $E_O^d = k > 0$. This would suggest that $E\pi^R(\phi = 1|\hat{\phi} = 0) > E\pi^{NR}(\phi = 1|\hat{\phi} = 0)$, but the results would not not be affected qualitatively. Bank incentives to invest under regulation would be even stronger.

It is easy to verify that, $A(\theta) > A^{dev}(\theta)$. Then, solving (B.6) with respect to i, yields the interest rate for the T group when a bank has unilaterally deviated to $\phi = 0$

$$i_T^{R,dev} = \frac{A^{dev}(\theta)(H-1)+\bar{i}}{1+A^{dev}(\theta)}.$$

The profits from a unilateral deviation to $\phi = 0$ are given as follows

$$\begin{aligned} E\pi^R(\phi=0|\hat{\phi}=1) &= B^o = \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}\left[q(i_T^{R,dev}-\bar{i})+(1-q)(\mu(1+i_O^R)-(1+\bar{i}))\right] \\ &= \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}\left[\frac{qA^{dev}(\theta)(H-(1+\bar{i}))}{1+A^{dev}(\theta)}+(1-q)(\mu H-(1+\bar{i}))\right]. \end{aligned}$$

The unilateral incentive not to invest in the screening technology when the regulation is in force is given by

$$\begin{aligned} & \Delta^R(\theta|\hat{\phi}=1) = E\pi^R(\phi=1) - E\pi^R(\phi=0|\hat{\phi}=1) \\ = & \frac{\alpha_b(\theta)(H-(1+\bar{i}))}{r+\alpha_b(\theta)}\left[q\left(\frac{A_H(\theta)}{1+A_H(\theta)}-\frac{A^{dev}(\theta)}{1+A^{dev}(\theta)}\right)+(1-q)\left(\frac{\mu A_H(\theta)}{1+A_H(\theta)}-\frac{\mu H-(1+\bar{i})}{H-(1+\bar{i})}\right)\right] \\ = & \frac{\alpha_b(\theta)(H-(1+\bar{i}))}{r+\alpha_b(\theta)}\left[\frac{q(A_H(\theta)-A^{dev}(\theta))}{(1+A_H(\theta))(1+A^{dev}(\theta))}+(1-q)\left(\frac{\mu A_H(\theta)}{1+A_H(\theta)}-\frac{\mu H-(1+\bar{i})}{H-(1+\bar{i})}\right)\right]. \end{aligned}$$

B.3 Proof of Lemma 6

The first inequality

$$\Delta^R(\theta|\hat{\phi}=1) > \Delta^{NR}(\theta|\hat{\phi}=1)$$

follows from $A^{dev}(\theta) < A_T^{dev}(\theta)$, see (B.7), and the second inequality

$$\Delta^R(\theta|\hat{\phi}=0) > \Delta^{NR}(\theta|\hat{\phi}=0)$$

follows from $A(\theta) < A_T(\theta)$, see (B.3).

Moreover, when $\mu = 0$, $\Delta^{NR}(\theta|\hat{\phi}=1) = \Delta^{NR}(\theta|\hat{\phi}=0) = 0 < \Delta^R(\theta|\hat{\phi}=0) = \Delta^R(\theta|\hat{\phi}=1) = \frac{\alpha_b(\theta)}{r+\alpha_b(\theta)}(1-q)(1+\bar{i}) > 0$. Therefore, by continuity,

$$\max\left\{\Delta^{NR}(\theta|\hat{\phi}=0), \Delta^{NR}(\theta|\hat{\phi}=0)\right\} < \min\left\{\Delta^R(\theta|\hat{\phi}=0), \Delta^R(\theta|\hat{\phi}=0)\right\}$$

is satisfied for μ at least in the neighborhood of zero.

B.4 Proof of Proposition 3

Assumption 2 ensures that $\phi = 0$ is an equilibrium under no regulation. Then we know from Lemma 2 that a bank's unilateral incentives to invest are either decreasing or inverted U-shaped in θ. But in either case they are decreasing for high enough θ. Finally, from Lemma 6, we know that the unilateral incentives to invest under regulation are stronger than under no regulation. The rest of the proof follows by noting that for the range of values of c that satisfy (4.5), i.e., intermediate cost, $\phi = 0$ is the unique equilibrium under regulation, while under no regulation, $\phi = 1$ is the unique equilibrium for low θ and $\phi = 0$ is the unique equilibrium for high θ.

B.5 Asymmetric equilibrium in information production under no regulation

We assume that a fraction x_0 of banks chooses $\phi = 0$ and the remaining fraction, $1-x_0$, chooses $\phi = 1$. Let $\theta_0 \equiv \frac{x_0 N_b}{N_e}$ and $\theta_1 \equiv \frac{(1-x_0)N_b}{N_e} = \theta - \theta_0$. The banks that have chosen not to invest in information production will only fund T entrepreneurs, while banks who have invested in information production fund all high value projects.

We first compute an entrepreneur's outside option. Consider the value of not having a deal with a bank for a short period Δt. The flow benefit is zero. The Poisson arrival rate of a deal for the entrepreneur is $\alpha_e(\theta)$. At the end of $t + \Delta t$ the entrepreneur either continues with no deal with probability $e^{-\alpha_e(\theta)\Delta t}$, or strikes a deal with a bank that has made no investment in information acquisition and the entrepreneur is T, which happens with probability $1 - e^{-\alpha_e(\theta_0)q\Delta t}$, or strikes a deal with a bank that has made an investment in information acquisition and the entrepreneur's project is of high value, which happens with probability $1 - e^{-\alpha_e(\theta_1)(q+\mu(1-q))\Delta t}$. In either case the value of the deal is E_H^d. Thus, the value of the outside option is expressed as follows

$$E^o = e^{-r\Delta t}\left(e^{-\alpha_e(\theta)\Delta t}E^o + \left(1 - e^{-(\alpha_e(\theta_0)q)\Delta t}\right)E_H^d + \left(1 - e^{-(\alpha_b(\theta_1)(q+\mu(1-q))\Delta t}\right)E_H^d\right).$$

Solving for E^o and letting $\Delta t \to 0$, using l'Hopital's rule, we obtain

$$E^o = \frac{\alpha_e(\theta_0)q + \alpha_e(\theta_1)(q + \mu(1-q))}{r + \alpha_e(\theta)} E_H^d. \tag{B.8}$$

The bank's outside option depends on the information acquisition strategy it has adopted and is given as follows

$$B_0^o = \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)} q B_H^d \quad \text{or} \quad B_1^o = \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)} (q + \mu(1-q)) B_H^d. \tag{B.9}$$

Following similar steps as in the derivation of (A.2), and using (B.8) and (B.9), we derive the market power expression for a bank that has made no investment in information acquisition

$$A_0(\theta_0, \theta) \equiv \frac{1 - \frac{\alpha_e(\theta_0)q + \alpha_e(\theta_1)(q + \mu(1-q))}{r + \alpha_e(\theta)}}{1 - \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)} q}$$

and the market power expression for a bank that has made an investment in information acquisition

$$A_1(\theta_0, \theta) \equiv \frac{1 - \frac{\alpha_e(\theta_0)q + \alpha_e(\theta_1)(q + \mu(1-q))}{r + \alpha_e(\theta)}}{1 - \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)} (q + \mu(1-q))},$$

where $A_1(\theta_0, \theta) > A_0(\theta_0, \theta)$.

Observe that as $\theta_0 \to 0$, $A_0 \to A_T^{dev}$ given by (A.12) and $A_1 \to A_H$ given by (A.10). Also, observe that as $\theta_0 \to \theta$, $A_1 \to A_H^{dev}$ given by (A.5) and $A_0 \to A_T$ given by (A.2).

The equilibrium interest rates, as a function of a bank's information acquisition strategy, are given by

$$i_0^\star = \frac{A_0(\theta_0, \theta)(H - 1) + \bar{i}}{1 + A_0(\theta_0, \theta)} \quad \text{or} \quad i_1^\star = \frac{A_1(\theta_0, \theta)(H - 1) + \bar{i}}{1 + A_1(\theta_0, \theta)}.$$

A bank that chooses $\phi = 1$ charges a higher interest rate than a bank that chooses $\phi = 0$, i.e., $i_1^\star > i_0^\star$. This is because a bank that has invested in information acquisition has a better outside option. When a bank does not invest in information production its expected profit, using the outside option B_0^o, is

$$E\pi^{NR}(\phi = 0; \theta_0, \theta) = B_0^o = \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)} q (i_0^\star - \bar{i}) = \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)} \frac{q A_0(\theta_0, \theta)(H - (1 + \bar{i}))}{1 + A_0(\theta_0, \theta)}.$$

When a bank invests in information production its expected profit excluding the cost c, using the outside option B_1^o, is

$$\begin{aligned} E\pi^{NR}(\phi = 1; \theta_0, \theta) &= B_1^o = \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)}(q + \mu(1-q))(i_1^\star - \bar{i}) \\ &= \frac{\alpha_b(\theta)}{r + \alpha_b(\theta)} \frac{(q + \mu(1-q))A_1(\theta_0, \theta)(H - (1 + \bar{i}))}{1 + A_1(\theta_0, \theta)}. \end{aligned}$$

It follows easily that $E\pi^{NR}(\phi = 1; \theta_0, \theta) > E\pi^{NR}(\phi = 0; \theta_0, \theta)$. Let $\Delta^{NR}(\theta_0, \theta) \equiv E\pi^{NR}(\phi = 1; \theta_0, \theta) - E\pi^{NR}(\phi = 0; \theta_0, \theta)$. An asymmetric equilibrium for any θ is characterized by a θ_0, such that $\Delta^{NR}(\theta_0, \theta) = c$.

It can be readily verified that $\Delta^{NR}(\theta_0 = \theta, \theta)$ is equal to $\Delta^{NR}(\theta|\hat{\phi} = 0)$ given by (4.1) and $\Delta^{NR}(\theta_0 = 0, \theta)$ is equal to $\Delta^{NR}(\theta|\hat{\phi} = 1)$ given by (4.2). Moreover, as θ_0 moves from 0 to θ, $\Delta^{NR}(\theta_0, \theta)$ moves continuously. Therefore, when $\Delta^{NR}(\theta|\hat{\phi} = 0) > \Delta^{NR}(\theta|\hat{\phi} = 1)$ and $c \in (\Delta^{NR}(\theta|\hat{\phi} = 1), \Delta^{NR}(\theta|\hat{\phi} = 0))$, or $\Delta^{NR}(\theta|\hat{\phi} = 0) < \Delta^{NR}(\theta|\hat{\phi} = 1)$ and $c \in (\Delta^{NR}(\theta|\hat{\phi} = 0), \Delta^{NR}(\theta|\hat{\phi} = 1))$, by the Intermediate Value Theorem there exists a $\hat{\theta}_0 \in (0, \theta)$ such that $\Delta^{NR}(\hat{\theta}_0, \theta) = c$. The fraction of banks that chooses $\phi = 0$ is then $\hat{x}_0 = \frac{\hat{\theta}_0}{\theta}$. This is the asymmetric equilibrium in Lemma 1.