A New Theory of Banking Inefficiency*

Kinda Hachem†
Booth School of Business
University of Chicago

First Version November 2010; This Version December 2011

Abstract

Unlike much of the literature on financial sector inefficiency, this paper identifies an inefficiency that does not rely on linkages between asset prices, net worth, and financing constraints. I instead focus on how banks allocate resources across intermediation activities and demonstrate that the privately optimal allocation between matching and screening is constrained inefficient. In particular, too many resources are devoted towards getting rather than vetting borrowers but, once properly vetted, not enough of these matches are actually retained. Uninformed financing is thus inefficiently high, defaults are in excess, and a mild matching tax helps remedy the situation. (JEL D62, D83, E44)

*This paper is based on the second chapter of my dissertation at the University of Toronto. It previously circulated under the title "Screening, Lending Intensity, and the Aggregate Response to a Bank Tax." I thank Jon S. Cohen, Alok Johri, Anil Kashyap, and Shouyong Shi for valuable suggestions. I have also benefitted from discussions with Miquel Faig, Pedro Gete, Gueorgui Kambourov, Jim Pesando, and seminar participants at Toronto, Chicago Booth, NYU Stern, Georgetown, FRB Philadelphia, FRB Cleveland, FRB Richmond, FRB Chicago, the Federal Reserve Board, Michigan State, UT Austin, Simon Fraser, Waterloo, McMaster, Duke Fuqua, Berkeley Haas, UBC Sauder, Ryerson, RMM 2010, and SED 2011. Financial support from the Social Sciences and Humanities Research Council of Canada as well as Shouyong Shi's Bank of Canada Fellowship and Canada Research Chair are gratefully acknowledged. Any errors are my own.

†University of Chicago, Booth School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637, USA. Tel: 1-773-834-0229. Email: kinda.hachem@chicagobooth.edu.

Toxic assets have led to huge losses over the past few years, prompting many to explain the Great Recession as a miscalculation of risk by banks. In theory, however, banks exist precisely because they are good at intermediation, providing risk sharing in incomplete markets and screening under asymmetric information. How, then, could they have gotten it so wrong? This paper contributes to a growing literature on financial sector inefficiency by investigating a new margin: inefficiencies arising from the allocation of bank resources across intermediation activities.

The two activities I focus on are matching and screening. Due to competition, we often observe banks creating financial products and advertising their loan services in order to attract borrowers. At the same time though, asymmetric information between borrowers and lenders means that banks also devote some resources to learning about who they attract. The intensity with which each of these activities is undertaken determines the quantity and quality of bank lending so any inefficiencies in the allocation of resources between matching and screening can have serious implications for the health of the financial system.

Consider the case of mortgage-backed securities. Increased bank involvement in the MBS market following the repeal of Glass-Steagall fostered a credit boom as banks packaged loans into securities, sold them, and used the proceeds to support new originations. Fallout from the ensuing bust, however, made it clear that not enough was known about the underlying mortgages or how they would interact once packaged. Securitization was thus associated with too many originations (matches) and too little information (screening). Although the ability to offload risk by selling MBSs certainly did not stoke the incentive to screen, another interpretation is that the increase in matches afforded by securitization was overly appealing to banks because of a more fundamental problem in how they trade off quality and quantity. Stated otherwise, the rapid flight to securitization may have been more a symptom than a cause of banking inefficiency.[1]

[1] See Keys et al. (2010) for the negative effect of securitization on screening. Support for the alternative interpretation - namely that banks sacrifice screening for reasons beyond

To understand the tradeoff between matching and screening and to determine whether it begets inefficiency, I build a model that formalizes the allocation decision of banks. The goal is to understand this tradeoff in its purest form so I abstract from mortgages, securitization, and anything else that may have been unique to a particular financial episode. Instead, my economy features a continuum of heterogeneous borrowers differing in production ability. Each borrower needs one unit of capital to produce but this capital can only be intermediated by a mass of ex ante identical lenders. As described above, the intermediation process consists of attracting borrowers and screening them. Matches are important because credit is needed for production. At the same time though, screening is important because low quality borrowers are more likely to destroy capital by running unprofitable projects. Although lenders may want to undertake both matching and screening, the allocation of resources across these activities will be non-trivial if it is either too costly or too time-consuming to undertake each activity until its marginal return is zero. To ease the exposition, I capture this restriction as a resource constraint which precludes lenders from making both activities succeed with probability 1.[2]

securitization - can be gleaned from the remarks of Ed Clark, chief executive officer of Toronto-Dominion Bank. In an interview with Canada's largest national newspaper, he is quoted as follows: *"If we said 'Look, we're going to be heroes and save Canada from itself, and we'll impose a whole new [mortgage] regime on everyone else,' the other four [large] banks would say 'Let's carve them up'" ... it is impossible to expect any bank to crack the whip on borrowers because "market share loss is perceived as a strategic loss, not just a numerical or dollar loss"* (The Globe and Mail, "Banks won't lead way on fixing debt problem: TD's Clark" by Tim Kiladze, 12/15/2010). While these remarks were made about Canadian bankers, they also cast a shadow on American banks as the US financial system is generally deemed the more competitive of the two.

[2] Note that the allocation decision in my model is fundamentally different from the one in Shleifer and Vishny (2010). In the latter, banks divide scarce capital between traditional lending and investment in securitized debt. This decision interacts with asset prices to transmit investor sentiment to the real economy. In contrast, my model focuses on traditional lending so there is no alternative use for capital. Instead, banks allocate resources other than this capital to determine the quantity and quality of their traditional loans. In other words, I am essentially examining how banks allocate their staff between matchers and screeners rather than how matchers (to extend the term) allocate capital between loans and securities. Moreover, as we will see below, the mechanism through which my allocation affects the real economy does not operate through asset prices.

After analyzing an individual lender's optimal division of resources between attracting and screening borrowers, I demonstrate that the decentralized equilibrium is constrained inefficient. In particular, unmatched lenders devote too many resources to forming matches and too few to screening them. The first externality operates through the distribution of available borrowers when matches can be preserved over time. Since attracting a borrower today limits the need for matching resources tomorrow, lenders who carry their clients over can devote more of tomorrow's resources towards screening if today's screening efforts are unsuccessful. The eventual rejection of unprofitable borrowers then worsens the pool that currently unmatched lenders will draw from should they try to attract someone later on. In this way, high matching effort by some lenders induces others to also over-invest in matching, creating an "attract now, screen later" motive that propels the market to a steady state with too much uninformed lending and too many defaults. The problem is exacerbated by a second externality which renders informed lenders overly selective in the types they retain. An informed lender determines the lowest type he is willing to finance by comparing the expected value from keeping that type to the value attainable as an unmatched lender. By allocating resources to maximize the latter, unmatched lenders increase the opportunity cost of being matched and prompt increased selectivity among informed lenders. To decrease the odds of rematching and thus decrease the endogenous destruction of informed financing, the efficient allocation again prescribes less matching.

A corollary of these results is that bank taxes which limit the drive to attract borrowers can improve social welfare. I investigate a simple version of this policy, namely a proportional tax on the matching activity. I find that both uninformed financing and defaults decrease with the tax. Production exhibits a hump-shaped response since the shift towards informed financing has a positive effect as long as the frequency of new matches does not become too small. I also find that a mild version of this tax can attenuate business cycle fluctuations.

To the extent that my paper emphasizes financial non-neutrality, it is related to the macroeconomic literature on credit channels: Gurley and Shaw

(1955), Williamson (1987), Bernanke and Gertler (1989), and Kiyotaki and Moore (1997) are but a few examples. It is also related to a more recent branch of this literature which builds on the asset price propagation mechanism in Kiyotaki and Moore (1997) to investigate financial sector inefficiency. In Lorenzoni (2008) and Korinek (2011), for instance, fire sales of collateralizable assets impart pecuniary externalities which can culminate in a financial crisis.[3] In contrast, the externalities identified by model arise even if balance sheets and credit contracts are decoupled from asset prices, thus providing a new justification for regulatory intervention. Since the problem I propose exists at the level of bank decision-making, my paper is also related to previous work on the microeconomics of credit markets.[4] Examples here include Broecker (1990), Cao and Shi (2001), and Direr (2008) who examine screening externalities, Parlour and Rajan (2001) who examine competition externalities with strategic default, and Becsi et al. (2009) who examine search frictions in the credit market matching process. All these studies focus on either matching or screening though so they cannot explain how banks allocate resources between the two and how this allocation then affects the macroeconomy.

The paper proceeds as follows: Section 1 details the environment and sets up the optimization problems; Section 2 analyzes the decentralized steady state; Section 3 compares this steady state to the constrained efficient allocation and discusses the externalities; Section 4 calibrates the model to illustrate additional properties of the equilibrium and how they respond to corrective taxation; Section 5 concludes. All proofs are presented in Appendix A.

1 The Model

1.1 Environment

Time is discrete. All agents are infinitely-lived, risk neutral, and endowed with a unit of effort each period. There is a continuum of firm types, $\omega \in [0, 1]$,

[3] For more on fire sales in macroeconomics, see Shleifer and Vishny (2011) and the references therein.

[4] See Freixas and Rochet (1997) for an overview of such models.

with symmetric density function $f(\cdot)$. For simplicity, I set $f(\cdot) = 1$. Each firm has private information about its type. It also has access to a risky production project that requires one unit of external capital to operate. A type ω firm that obtains the necessary capital and exerts effort e into the production project generates $\theta(\omega)$ units of output with probability e and zero units with probability $1 - e$, where $\theta'(\cdot) > 0$ and $\theta''(\cdot) < 0$. Project output includes the original capital input so unsuccessful projects destroy capital. The firm's cost of exerting effort is $-c \ln(1 - e)$, where $c > 0$ is a constant.[5]

Firms cannot store project output and they do not have direct access to capital so they must borrow from a measure of ex ante identical lenders that also populates the economy. Lenders cannot produce but, in addition to capital, they have access to two technologies that allow them to emerge as intermediaries. First, lenders can create and/or advertise financial products to match firms with capital. The greater the number of matches, the greater the lending intensity. I abstract from the exact process through which lenders generate their matches, summarizing it instead as the operation of a matching technology. Second, lenders can screen firms to determine whether facilitating such matches is indeed profitable. Although lenders may want to undertake both activities, it is either too costly or too time-consuming to undertake each one until its marginal return is zero. This restriction is captured by a unit resource constraint. In particular, a lender who devotes π units of his effort endowment (interpreted as time) to matching gets a borrower with probability π and discovers that borrower's type with probability $1 - \pi$ immediately thereafter.[6] Lenders cannot support more than one match at a time and cannot search "on the contract" so the matching technology is only available to

[5] This functional form rules out the corner choice of $e = 1$ and thus conserves on algebra. As long as they are increasing and convex in the amount of effort, other functional forms will yield similar qualitative results.

[6] The probabilities of forming informed and uninformed matches are then $\pi(1 - \pi)$ and π^2 respectively. Therefore, if one would rather think of unmatched lenders as operating a single technology that combines getting and vetting borrowers, π and $1 - \pi$ can be interpreted as inputs into a credit market "production function" that yields an informed match with probability $\pi(1 - \pi)$, an uninformed match with probability π^2, and no match with probability $1 - \pi$.

unmatched lenders. In contrast, screening can be undertaken by all lenders.

At this point, it will be useful to provide a foundation for the unit resource constraint on lenders. To do so, consider a more general problem where intertemporal resource accumulation is permitted and lenders face concave intermediation technologies. If these technologies have enough curvature - that is, if the functions which transform resources into success probabilities for each activity exhibit sufficiently diminishing returns - then lenders will have an incentive to smooth resources out across time. In other words, the total amount of resources available each period will be roughly constant and the relevant margin is the fraction of these resources that go to matching rather than screening. Normalizing available resources in each period to one and interpreting π and $1-\pi$ as the matching and screening fractions thus proxies for the more general problem. Now, since decisions in the proxy problem are made subject to a fixed resource constraint each period, linear intermediation technologies will provide just as much intuition as concave ones. The former can then be used in conjunction with the unit constraint to further simplify the exposition.

To understand the implications of a lender's resource allocation decision, let us examine how lenders evolve over time. Begin with a lender who is unmatched at the end of date $t-1$. At the beginning of date t, the lender chooses π. If he fails to attract a borrower, then he stays unmatched throughout t and must try again in $t+1$. If, however, he succeeds in forming a match, then he exerts screening effort $1-\pi$ right after getting that match. Successful screening means that the lender's information set contains the borrower's true type whereas unsuccessful screening means that it only contains the lender's beliefs about the pool of borrowers from which he drew the match. To keep the analysis tractable, I assume that these beliefs cannot be conditioned on credit ratings if screening fails.[7]

Given his information set, the newly matched lender must make two more

[7] That is, discovering accurate credit histories requires screening to be at least partly successful. I emphasize accuracy as the recent crisis revealed several instances where "off the shelf" ratings were problematic.

decisions at the beginning of date t. First, he must decide whether to finance the borrower he just attracted or whether to let him go and try for another borrower in $t+1$. Information is clearly important here because only lenders who have successfully screened will be able to gauge how profitable their matches really are. In contrast, lenders who must rely on their beliefs about the borrower pool can only gauge average profitability across types. In what follows, I denote the retention strategy of a matched and informed lender by $a(\omega)$, where $a(\omega)$ is an indicator function that equals 1 if and only if the lender accepts to finance a type ω borrower.[8] Conditional on him keeping the borrower, the lender's second decision is what contract terms to offer. I assume no intertemporal commitment so each contract is defined by a one-period loan rate. This rate includes the borrowed unit of capital and must be paid to the lender if the project succeeds. Lenders cannot observe the exact outcome of a project but can detect the presence of positive output so borrowers repay if and only if their projects are successful. The information on which the lender conditions his loan rate is again important. Since the same rate can induce different ω's to exert different production effort, the lender's offer affects whether the borrower's project will fail and, therefore, whether capital will be destroyed.

Once retention decisions have been made and loan rates set, matched borrowers undertake production. The output of a successful project is then split so that, given loan rate R, the borrower gets $\theta(\omega) - R$ and the lender gets R. Borrowers consume their entire cut. In contrast, lenders save $(1-\delta) R$ as capital for future financing and deplete the rest, δR where $\delta \in [0,1)$, as an operating expense.[9] In what follows, I assume the existence of a competitive

[8] To ease exposition, assume that creating and/or advertising financial products is enforceable in the sense that a lender who has exerted $\pi > 0$ in the current period and attracted a match cannot reject that match unless he can prove that the applicant's ω is too low (i.e., can only discriminate based on ω).

[9] Since the measure of borrowers is 1 and each borrower gets at most 1 unit of capital, capital demand is bounded above by 1. There are no such restrictions on lender revenue so δ is introduced to help ensure that capital supply is also bounded. Assumption 1 in Section 2.3 provides a sufficient condition for bounded capital accumulation. While $\delta = 0$ is certainly permissible, lower values of $\theta(\cdot)$ and/or higher values of c would be required for market clearing.

interbank market for capital and denote its market clearing cost by r. A lender who does not have enough capital to finance his match must borrow at rate r while a lender who does have enough capital interprets r as the opportunity cost of accepting his match. A lender's gross cost of funds is thus $1 + r$, where 1 represents the loan made to the borrower. Combined with the fact that each borrower needs only one unit of capital and the fact that lenders can finance only one borrower at a time, the interbank market allows us to abstract from the distribution of capital across lenders and focus instead on aggregate capital accumulation.[10]

At the end of date t, matches are subject to an exogenous separation probability $\mu \in (0, 1)$. Separation implies that the lender starts $t + 1$ unmatched. Non-separation implies that he carries his match into $t + 1$ and thus cannot operate the matching technology that period. Since screening is still available to all lenders and a lender's effort endowment is not intertemporally transferable, it then follows that any matched lender who enters $t + 1$ without full information about his borrower's type will undertake complete screening. As a result, uninformedness lasts for at most one period and within-lender credit history is rendered irrelevant.[11] The lender's problem is now that of a matched lender who enters $t + 1$ with full information: at the beginning of the period

[10] In the spirit of Diamond (1984) and Rajan (1994), my model focuses on the lending function of banks. To justify the abstraction from deposit-taking decisions, one could imagine something akin to deposit insurance. By making banks look equally riskless to depositors, such insurance would dull both the need to compete for deposits and the fear of a bank run, allowing all the action to come from the lending side. Interbank capital could then be interpreted as the insured deposits held by the banking system and δ could be interpreted as the exogenous fraction of deposits withdrawn each period. That banks funded entirely by insured deposits would then just maximize expected interest income is consistent with Stein (1998).

[11] That lenders are uninformed for at most one period also simplifies the inference problem for unmatched lenders and keeps the model analytically tractable. In particular, if uninformed lenders did not fully screen in the second period, then inferences about the pool of available borrowers would also depend on how long borrowers stay in their matches before being let go and what the lenders that let them go may have learned in that time. In a way, type discovery by the second period of a match is consistent with the literature on relationship lending which argues that repeated interactions with the same borrower hasten the rate at which a lender learns about that borrower. For more on relationship lending, see Boot (2000), Hachem (2011), and the references therein.

and conditional on ω, he decides whether to finance the borrower again and, if he accepts to finance, then he also chooses a one-period loan rate. If he rejects, then he enters $t+2$ unmatched.

1.2 Optimization Problems

1.2.1 Borrowers

Consider a type ω borrower who has obtained financing at loan rate R. Without credit ratings or intertemporal incentives, the borrower's problem is a static one: given R, he chooses how much effort to put into the production project so as to maximize his one-period expected utility. Recall that a type ω who exerts production effort e succeeds with probability e, in which case the project generates $\theta(\omega)$ and the borrower repays his lender R. Taking into account the disutility of effort, the borrower thus chooses $e \in [0,1]$ to maximize $e[\theta(\omega) - R] + c \ln(1-e)$. Conditional on ω and R, this problem yields the following optimal strategy:

$$e(\omega, R) = \begin{cases} 0 & if \quad R > \theta(\omega) - c \\ 1 - \frac{c}{\theta(\omega) - R} & if \quad R \leq \theta(\omega) - c \end{cases} \tag{1}$$

If the loan rate is higher than the choke rate $\theta(\omega) - c$, then ω's project fails with certainty because he has no incentive to exert production effort. If the loan rate is lower than the choke rate, then ω's effort is positive but strictly decreasing in R. Since $\theta'(\cdot) > 0$, a higher type is more likely to exert positive effort and his effort in this case will be higher for any given loan rate.

1.2.2 Lenders

As described earlier, a lender's problem depends on whether he is matched or unmatched and, if matched, it also depends on whether he is informed or uninformed about his borrower's type. Since a lender's choices affect how he evolves over time, I formulate the problem as a dynamic program. Let J denote the value function of an informed lender, X the value function of an uninformed lender, U the value function of an unmatched lender, and $\beta \in (0,1)$ the common discount factor. The aggregate state is summarized by $S \equiv \left\{K, V(\cdot), \widetilde{V}(\cdot), \lambda_{-1}(\cdot), \phi_{-1}(\cdot)\right\}$, where K is the beginning-of-period capital stock, $V(\omega)$ is the value of type ω under informed financing, $\widetilde{V}(\omega)$ is the value of ω if unmatched, $\lambda_{-1}(\omega)$ is the proportion of ω's financed by informed lenders last period, and $\phi_{-1}(\omega)$ is the proportion financed by uninformed lenders.

Consider first an informed lender matched with a type ω borrower. The lender takes as given S and his individual state $\{\omega, v\}$, where v is the value attained by his borrower. In turn, he must choose whether to keep the borrower (a), what loan rate to charge if he does keep him (R), and what continuation value to offer (v_{+1}). Since the borrower has the option of turning down the contract and hoping for a new lender next period, R and v_{+1} must make the borrower want to participate (i.e., the present discounted value of staying with an informed lender cannot be less than $\beta\widetilde{V}_{+1}(\omega)$). The informed problem can be written as:

$$J(\omega, v, S) = \max_{a,R,v_{+1}} \left\{ \begin{array}{l} (1-a)\,\beta U\left(S_{+1}, \psi_{+1}\right) \\ +a \left[\begin{array}{l} \left(1 - \frac{c}{\theta(\omega)-R}\right) R - (1 + r(S)) \\ +\beta\left[(1-\mu)\, J(\omega, v_{+1}, S_{+1}) + \mu U\left(S_{+1}, \psi_{+1}\right)\right] \end{array} \right] \end{array} \right\}$$

$$\begin{gathered} \textit{subject to} \\ a \in [0,1]\,,\ R \in [0, \theta(\omega) - c] \\ v = \theta(\omega) - R - c + c\ln\left(\frac{c}{\theta(\omega)-R}\right) + \beta\left[(1-\mu)\, v_{+1} + \mu\widetilde{V}_{+1}(\omega)\right] \geq \beta\widetilde{V}_{+1}(\omega) \\ S_{+1} = \Gamma(S)\,,\ \psi_{+1} = \mathcal{G}(S_{+1}) \end{gathered} \tag{2}$$

Let us work through equation (2). If the lender rejects the borrower, then he gets the discounted value of being unmatched next period (ψ, the individual state of an unmatched or uninformed lender, is defined in the next paragraph). In contrast, if he accepts the borrower, then his expected profit in the current period is $e(\omega, R)\,R$ less the gross cost of funds $1 + r(S)$. The lender's future value is then $J(\omega, v_{+1}, S_{+1})$ if the match is not exogenously destroyed and $U\left(S_{+1}, \psi_{+1}\right)$ otherwise. Note that $e(\omega, R)$ as in (1) means the lender will never want to charge above $\theta(\omega) - c$. Moreover, although higher values of R increase the lender's revenue if repaid, they also decrease the probability of repayment so an informed lender will not want to monopolize his borrower and the participation constraint may not actually bind. To complete the problem, the lender's beliefs about the evolution of S and ψ are governed by laws of motion which, as will be discussed in Section 2.1, must be consistent with aggregate behaviour.

Consider now an uninformed lender. Without knowledge of ω, this lender can only offer a pooled rate $\overline{R}$ which induces $e\left(\omega, \overline{R}\right) > 0$ if and only if $\overline{R} < \theta(\omega) - c$. Let $\eta\left(\overline{R}\right)$ denote the highest type that does not exert effort under $\overline{R}$ and let $\psi(\cdot)$ denote the density function that the lender believes characterizes the pool of borrowers from which he drew. The expected one-period profit of an uninformed lender is then $\int_{\eta(\overline{R})}^{1} \left(1 - \frac{c}{\theta(\omega) - \overline{R}}\right) \overline{R}\psi(\omega)\, d\omega$ less the gross cost of funds. Recall from Section 1.1 that uninformedness lasts for at most one period so, if the match is not exogenously destroyed at the end of this period, the lender's future value is $J(\omega, V_{+1}(\omega), S_{+1})$. Since ω is not known at the time of the uninformed problem though, this value must be weighted by $\psi(\omega)$ and integrated. If the match is exogenously destroyed, then the lender's future value is $U\left(S_{+1}, \psi_{+1}\right)$. The uninformed problem is thus:

$$X(S, \psi) = \max_{\overline{R}} \left\{ \begin{array}{l} \displaystyle\int_{\eta(\overline{R})}^{1} \left(1 - \frac{c}{\theta(\omega) - \overline{R}}\right) \overline{R}\psi(\omega)\, d\omega - (1 + r(S)) \\ \displaystyle +\beta(1-\mu) \int_{0}^{1} J(\omega, V_{+1}(\omega), S_{+1})\, \psi(\omega)\, d\omega \\ +\beta\mu U\left(S_{+1}, \psi_{+1}\right) \end{array} \right\} \tag{3}$$

$$\textit{subject to}$$

$$\overline{R} \in [0, \theta(1) - c],\ \eta\left(\overline{R}\right) = \arg\min_{w \in [0,1]} \left|\theta(w) - c - \overline{R}\right|$$

$$S_{+1} = \Gamma(S),\ \psi_{+1} = \mathcal{G}(S_{+1})$$

Finally, consider an unmatched lender who must decide how to divide resources between getting matches and screening applicants. A lender who devotes π units to attracting a borrower becomes matched and informed with probability $\pi(1-\pi)$, matched and uninformed with probability π^2, and stays unmatched with probability $1-\pi$. We thus have the following problem for an unmatched lender:

$$U(S,\psi) = \max_{\pi} \left\{ \begin{array}{l} \pi(1-\pi)\int_0^1 J(\omega, V(\omega), S)\,\psi(\omega)\,d\omega \\ +\pi^2 X(S,\psi) + (1-\pi)\beta U\left(S_{+1}, \psi_{+1}\right) \end{array} \right\} \tag{4}$$

$$\textit{subject to}$$

$$\pi \in [0,1]\,,\ S_{+1} = \Gamma(S)\,,\ \psi_{+1} = \mathcal{G}(S_{+1})$$

1.3 Laws of Motion

1.3.1 Capital

The evolution of the capital stock is governed by:

$$K_{+1} = (1-\delta)\left[\int_0^1 e(\omega, R(\omega))\,R(\omega)\,\lambda(\omega)\,d\omega + \int_0^1 e\left(\omega, \overline{R}\right)\overline{R}\phi(\omega)\,d\omega\right] \tag{5}$$

Starting from K, K_{+1} is calculated by subtracting the amount of capital put into production then adding the share of output saved by lenders. Each loan transfers one unit of capital to the borrower so the amount of capital put into production equals the measure of borrowers financed. This is essentially capital demand and it is given by $\tilde{K} \equiv \int_0^1 [\lambda(\omega) + \phi(\omega)]\,d\omega$. In equilibrium though, the cost of funds $r(S)$ adjusts to yield aggregate versions of $a(\cdot)$ and π that clear the capital market.[12] Therefore, $\tilde{K} = K$ and, as per equation (5), K_{+1} just equals the output saved by lenders during the previous period.

1.3.2 Distributions

I now present the laws of motion for the proportion of type ω's with informed financing and the proportion with uninformed financing, $\lambda(\omega)$ and $\phi(\omega)$ re-

[12] The role of r is entirely indirect. If r is interpreted as an opportunity cost that the lender must be compensated for, then it does not enter into aggregate accounting. If it is instead interpreted as a direct cost - namely the cost of borrowing the required unit from another lender on the interbank market - then r is subtracted from the revenues of the borrowing lender and added to the revenues of the lending lender, effectively washing out.

spectively. Let Π denote the aggregate lending intensity of unmatched lenders and $A(\cdot)$ the aggregate retention strategy of informed lenders. The evolution of $\lambda(\omega)$ follows:

$$\lambda(\omega) = A(\omega)\left[\begin{array}{l} (1-\mu)\left[\lambda_{-1}(\omega)+\phi_{-1}(\omega)\right] \\ +\left[1-(1-\mu)\left[\lambda_{-1}(\omega)+\phi_{-1}(\omega)\right]\right]\Pi(1-\Pi) \end{array}\right] \tag{6}$$

If $A(\omega)=0$, then all ω's are rejected by informed lenders so $\lambda(\omega)=0$. Consider now $A(\omega)=1$. Borrowers who were financed by informed lenders last period and who are still matched with their lenders at the beginning of this period will again obtain informed financing. Since uninformedness lasts for at most one period, borrowers who were financed by uninformed lenders last period and who are still matched this period will also obtain informed financing. These two statements explain $(1-\mu)\left[\lambda_{-1}(\omega)+\phi_{-1}(\omega)\right]$ in equation (6). To see where the second term comes from, note that some borrowers who start the current period unmatched may also obtain informed financing. The mass of ω's available to be matched is the total mass of ω's, 1, minus the mass that is already in matches, $(1-\mu)\left[\lambda_{-1}(\omega)+\phi_{-1}(\omega)\right]$. Fraction Π of these available borrowers are drawn into new matches and, out of these new matches, fraction $1-\Pi$ are informed.

Turning to the law of motion for $\phi(\omega)$:

$$\phi(\omega) = \left[1-A(\omega)(1-\mu)\left[\lambda_{-1}(\omega)+\phi_{-1}(\omega)\right]\right]\Pi^2 \tag{7}$$

Since borrowers previously matched with uninformed lenders are either separated or discovered, there is no carry-over of uninformed matches. Instead, $\phi(\omega)$ is composed entirely of new matches that were unsuccessfully screened. If $A(\omega)=0$, then all ω's are available for new matches at the start of this period and, if $A(\omega)=1$, then only $1-(1-\mu)\left[\lambda_{-1}(\omega)+\phi_{-1}(\omega)\right]$ are available. Fraction Π of these available borrowers are drawn into new matches and, out of these new matches, fraction Π are uninformed.

With $\lambda(\omega)$ and $\phi(\omega)$ characterized, we can also formalize $\psi(\omega)$. For any ω, the proportion of unmatched borrowers at the beginning of the current period can be written as $1 - A(\omega)(1-\mu)\left[\lambda_{-1}(\omega) + \phi_{-1}(\omega)\right]$ so equilibrium beliefs about the composition of available borrowers must satisfy:

$$\psi(\omega) = \frac{1 - A(\omega)(1-\mu)\left[\lambda_{-1}(\omega) + \phi_{-1}(\omega)\right]}{\int_0^1 \left[1 - A(x)(1-\mu)\left[\lambda_{-1}(x) + \phi_{-1}(x)\right]\right] dx} \tag{8}$$

2 Decentralized Equilibrium

2.1 Definition of Equilibrium

A symmetric equilibrium in this model is a set of lender value functions $\{J, X, U\}$ and sequences of borrower continuation values $\left\{V, \widetilde{V}\right\}$, individual decision rules $\left\{a, \pi, R, \overline{R}, v_{+1}\right\}$, aggregate decision rules $\{A, \Pi\}$, distributions $\{\lambda, \phi\}$, financing capital $\{K_{+1}\}$, costs of funds $\{r\}$, and beliefs $\{\psi, \Gamma, \mathcal{G}\}$ satisfying:

1. Lender optimality as per the optimization problems in Section 1.2.2.
2. Symmetry (i.e., $A = a$, $\Pi = \pi$, and $V = v$).
3. Capital market clearing as described in Section 1.3.1.
4. Laws of motion (5), (6), and (7).
5. Functional equations for V and $\widetilde{V}$.[13]
6. Consistency of beliefs and, in particular, ψ as given by (8).

[13] Recall that $V(\omega)$ and $\widetilde{V}(\omega)$ are just used to construct the participation constraint that an informed lender must satisfy in order to retain ω. The functional equations for the types that the lender does indeed want to keep are:

$$\begin{aligned}
V(\omega) &= \theta(\omega) - R(\omega) - c + c\ln\left(\tfrac{c}{\theta(\omega)-R(\omega)}\right) + \beta\left[(1-\mu)V_{+1}(\omega) + \mu\widetilde{V}_{+1}(\omega)\right] \\
\widetilde{V}(\omega) &= \Pi^2\left[\max\left\{\theta(\omega) - \overline{R} - c + c\ln\left(\tfrac{c}{\theta(\omega)-\overline{R}}\right), 0\right\} + \beta\left[(1-\mu)V_{+1}(\omega) + \mu\widetilde{V}_{+1}(\omega)\right]\right] \\
&\quad + \Pi(1-\Pi)V(\omega) + (1-\Pi)\beta\widetilde{V}_{+1}(\omega)
\end{aligned}$$

The rest of this paper focuses on symmetric equilibria with non-binding borrower participation constraints. In effect, I am restricting attention to equilibria where $R(\cdot)$ is independent of Π for analytical tractability. Participation constraints are thus ignored until Section 2.3. At that point, I will present conditions under which the resulting choices do indeed satisfy the constraints.

2.2 Optimal Lending Intensity

The key allocation decision in my model is the choice of π so I start by characterizing the best response function of an unmatched lender. That is, for any state of affairs on the informed side of the market, how does π respond to aggregate lending intensity? Proposition 1 eases notation by reducing the informed retention strategy from an indicator function to a cutoff type. Proposition 2 then establishes that the steady state best response function of an unmatched lender is decreasing in Π when interior:

Proposition 1 *There is a scalar ξ such that $A(\omega) = 1$ if and only if $\omega \geq \xi$.*

Proposition 2 *Let $\pi_l(\Pi|\xi)$ denote the steady state best response of π to Π for a particular value of ξ. There exists a scalar $\widehat{\xi}$ and a function $\widehat{\Pi}(\cdot)$ such that $\widehat{\Pi}'(\cdot) \leq 0$ and:*

1. *If $\xi < \widehat{\xi}$, then $\pi_l(\cdot|\xi) = 1$.*
2. *If $\xi \in \left[\widehat{\xi}, 1\right)$, then $\pi_l(\Pi|\xi) = 1$ for all $\Pi < \widehat{\Pi}\left(\widehat{\xi}\right)$ and $\pi_l(\Pi|\xi) \in (0,1)$ with $\frac{\partial \pi_l(\Pi|\xi)}{\partial \Pi} < 0$ for all $\Pi \in \left[\widehat{\Pi}\left(\widehat{\xi}\right), 1\right]$*
3. *If $\xi = 1$, then $\pi_l(\cdot|\xi) = 0$.*

To better understand the content of Proposition 2, notice that a lender has two incentives to learn his borrower's type. First, the information will help him reject unprofitable applicants and, second, it will help him determine how much surplus he can extract from the profitable ones. Very high values of ξ mean that only a small group of borrowers are profitable so the desire to identify them drives lending intensity down and, in the extreme case, we get $\pi_l(\cdot|1) = 0$. On the other hand, very low values of ξ mean that almost all types are profitable so the first incentive is diminished. Moreover, for ξ sufficiently low, the risk of not forming a match this period outweighs the second incentive and lending intensity tends to 1.

The interesting case is $\xi \in \left[\widehat{\xi}, 1\right)$. If $\Pi = 0$, then any unmatched lender who successfully expends $\pi > 0$ will have drawn from the initial distribution of types. As long as this distribution yields profitable expectations (which it must in order for the credit market to get off the ground), the lender will indeed choose $\pi > 0$. What happens if Π is slightly positive? Although other lenders are only getting a few borrowers, they screen them so intensely that at least some good types are pulled off the market while almost all the bad types remain. The average quality of available borrowers is thus lower relative to the case with $\Pi = 0$, increasing any individual lender's incentive to screen and decreasing the choice of π. Consider now a high value of Π, denoted by Π_H. A lot of matches are being formed but immediate type discovery is not common among other lenders so both good and bad borrowers are pulled off the market. This effect will be even more pronounced at $\Pi_H + \varepsilon$ so, if uninformed matches were to stay uninformed, beliefs under $\Pi_H + \varepsilon$ would be closer to the initial distribution than beliefs under Π_H, increasing the choice of π and delivering a U-shaped best response function. Recall, however, that uninformedness is eventually resolved (and bad borrowers thus released) when lenders can preserve their matches across periods. In turn, higher values of Π do not translate into better steady state beliefs and the best response function slopes downwards.[14]

[14] Three comments are in order. First, the deterioration in beliefs would not be eliminated by the entry of new borrowers. To see why, suppose that exogenously separated borrowers are replaced by new draws from $f(\cdot)$. Exogenous separations will still inject good and bad types into the available pool whereas endogenous separations will still only inject bad types. Once again then, the steady state pool will be characterized by a higher density of bad types than the initial distribution. Second, the deterioration will also be robust to multiple matches per lender. Suppose, for example, uninformed lenders can split resources between three activities: attracting a second borrower, screening the second borrower, and learning about the first borrower. Discovery of the first borrower's type may thus get delayed (potentially destroying more capital) but it cannot be delayed indefinitely for all lenders so

2.3 Existence and Uniqueness

With the best response non-increasing in Π, each value of ξ yields at most one symmetric equilibrium in the game between unmatched lenders. The question is now whether there is a unique ξ which, when combined with the solution to $\Pi = \pi_l(\Pi|\xi)$, clears the capital market. To establish the existence and uniqueness of such a steady state, I assume the following:

Assumption 1 $\int_0^1 \left(\sqrt{\theta(\omega)} - \sqrt{c}\right)^2 d\omega < \frac{1}{1-\delta}$

Assumption 2 $\theta(1) < \left(1 + \sqrt{\theta(0)}\right)^2$

Assumption 1 ensures that capital cannot be accumulated unboundedly while Assumption 2 regulates the worst borrower by putting a lowerbound on his output if successful. Proposition 3 summarizes the results:

Proposition 3 *There exist scalars* $\underline{c}$ *and* $\overline{c}$ *such that* $0 < \underline{c} < \overline{c} \leq \theta(0)$ *and:*

1. *For* $c \in (\underline{c}, \overline{c})$*, the set of symmetric equilibria with non-binding borrower participation constraints contains a trivial steady state (i.e.,* $\Pi = 0$*,* $\xi = 1$*, and* $K = 0$*) and a unique non-trivial steady state.*

2. *For* $c \geq \overline{c}$*, the aforementioned set only contains the trivial steady state.*

To understand the role of c here, recall the borrower strategy in equation (1). Without c, borrower effort is unaffected by the loan rate as long as the latter does not exceed the entire output of the project. Higher values of c introduce a moral hazard problem and, the higher the c, the more surplus an informed

an eventual rejection of bad borrowers still occurs. Third, if one wants to interpret higher Π as "more banks," then my downward sloping best response function is consistent with the screening literature's prediction that more banks can decrease any one bank's desire to provide credit. Unlike this literature, however, my mechanism evolves from an endogenous resource allocation decision, not from the one-period winner's curse faced by bidding banks who incur no extra costs to attract borrowers. Moreover, as I show below, the best response that results here is actually less downward sloping than the planner's prescription (i.e., in equilibrium, the desire to get borrowers is inefficiently high and the desire to screen them is inefficiently low, which is not the usual prediction of screening models).

lender will have to concede to any ω in order to incentivize him. Provided c is not so high that it shuts down the market, this surplus-sharing motive makes the participation constraint redundant and produces an equilibrium where $R(\cdot)$ is independent of Π. Going forward, I thus focus on $c \in (\underline{c}, \overline{c})$ and the non-trivial steady state.

3 Constrained Efficiency

3.1 Benchmark for Comparison

Consider a steady state social planner who holds the entire capital stock and who must allocate it to firms in order to achieve production. He faces the same constraints and intermediation technologies as lenders in the decentralized economy. In particular, Π determines the fraction of firms the planner reaches, $1-\Pi$ determines how many of these firms he is informed about, and ξ is the worst firm he allocates to if informed. Aggregate feasibility requires that the total amount of capital he allocates does not exceed the total amount he holds. Subject to aggregate feasibility, the planner chooses Π, ξ, $R(\cdot)$, and $\overline{R}$ to maximize the total present discounted value of capital. The R's are now interpreted as the division of project output into capital and consumption. A similar interpretation is valid for the decentralized economy since borrowers only consume while lenders only save. Loan rates were thus more than just prices so, to shut down any distortions stemming from how lenders versus planners want to divide output into its components, I start with a planner who uses capital rather than net output in his objective function. This yields:

Proposition 4 *Letting γ denote the multiplier on the aggregate feasibility constraint, the Lagrangian for the constrained efficiency problem is:*

$$\begin{aligned}
\mathcal{L} &= [1+\gamma(1-\beta)(1-\delta)]\left[\Pi^2 \int_{\eta(\overline{R})}^{\xi} e\left(\omega, \overline{R}\right) d\omega + \tfrac{\mu\Pi^2}{\mu+(1-\mu)\Pi} \int_{\xi}^{1} e\left(\omega, \overline{R}\right) d\omega\right] \overline{R} \\
&\quad + [1+\gamma(1-\beta)(1-\delta)] \tfrac{\Pi(1-\mu\Pi)}{\mu+(1-\mu)\Pi} \int_{\xi}^{1} e(\omega, R(\omega))\, R(\omega)\, d\omega \\
&\quad - \gamma(1-\beta)\left[\tfrac{\Pi(1-\xi)}{\mu+(1-\mu)\Pi} + \Pi^2 \xi\right]
\end{aligned}$$

Whether the decentralized equilibrium is efficient can now be addressed. Comparing the equilibrium conditions to the planner's first order conditions yields the following steady state results:

Proposition 5 *Denote the equilibrium allocation by* (ξ^*, Π^*) *and the constrained efficient allocation by* (ξ', Π')*.*

1. *If* $\mu = 1$*, then the equilibrium is constrained efficient.*

2. *If* $\mu \neq 1$ *and* β *is high, then the equilibrium is inefficient. For high values of* c *not exceeding* $\bar{c}$*, the direction of inefficiency is* $\Pi^* > \Pi'$ *and* $\xi^* > \xi'$*.*

Relative to the decentralized market, the second part of Proposition 5 says that the planner would devote more resources to screening new matches but be less restrictive in his cutoff once informed. Unmatched lenders are thus too liberal in their provision of credit while informed lenders are too conservative. As we will see later on, this leads to an inefficiently large market for uninformed financing and, on aggregate, an inefficiently high rate of delinquencies.

3.2 Externalities

Before proceeding, it will be useful to establish some intuition for Proposition 5. When $\mu = 1$, all matches are destroyed at the end of every period so the pool of available borrowers next period is always the initial distribution. That the equilibrium is efficient in this case suggests the externalities underlying the inefficiency when $\mu \neq 1$ stem from the intertemporal preservation of at least some matches. Indeed, high values of β confirm the intertemporal nature of the externality by ensuring that lenders care enough about the future implications of $\mu \neq 1$.

Starting from any given distribution, these implications are as follows. Since attracting a borrower today limits the need for matching resources tomorrow, lenders who carry their clients over can devote more of tomorrow's resources towards screening if today's screening efforts are unsuccessful. The eventual rejection of unprofitable borrowers then worsens the pool that currently unmatched lenders will draw from should they try to attract someone later on. In this way, high matching effort by some lenders will induce other lenders to also over-invest in matching, creating an "attract now, screen later" motive that propels the market to an inefficient steady state. In the language of Section 2.2, the externality is that unmatched lenders do not account for the fact that their lending intensity adversely affects the steady state pool of available borrowers. Just as Section 2.2 established that any $\Pi > 0$ worsens this pool relative to the initial distribution, the same line of reasoning establishes that any $\pi_l(\Pi|\xi) > 0$ also worsens the pool when adopted by all unmatched lenders. Each lender, however, sets $\pi_l(\Pi|\xi)$ without internalizing this negative feedback effect, making aggregate lending intensity inefficiently high.[15]

Unmatched lenders also fail to internalize their effect on the informed problem. Recall from equation (2) that an informed lender only retains borrowers who yield him at least as much as his outside option, βU. Unmatched lenders, however, choose individual lending intensity to maximize U, not recognizing that it then feeds back into the choice of ξ. Proposition 5 demonstrates that ξ is inefficiently high in the decentralized equilibrium, suggesting that the outside option of an informed lender is too large. The planner thus prescribes a lower value of Π both because he internalizes the effect on the distribution of available borrowers and because he internalizes the effect on the endogenous destruction of informed matches. To see how implementing the planner's allocation would make informed lenders less selective, note the two competing effects on U: while lowering Π may give informed lenders an incentive to hold out for better borrowers by mitigating the eventual deterioration in the quality of the available pool (the intensive effect), prescribing lower lending intensity to everyone also means a lower rematching rate should any informed lender decide to dissolve his current match (the extensive effect). Lower lending intensity thus lowers ξ through the extensive margin and results in more

[15] I emphasize that this externality operates through the quality of available borrowers, not the quantity. Indeed, by not making Π an argument in the matching technology of an individual lender, I have expressly shut down the type of congestion externality that the search literature is concerned with.

informed matches being preserved.

3.3 Corrective Taxation

The direction of the inefficiency identified in Proposition 5 motivates a tax on lending intensity. Consider a linear tax which makes activities designed to attract borrowers more costly. The tax rate is denoted by τ and only affects unmatched lenders. In particular, the maximization problem on the right-hand side of equation (4) now includes the term $-\tau\pi$. The tax revenues are then added back to aggregate capital so that all other equations are unchanged. Proposition 6 establishes that this tax does indeed have the desired effect:

Proposition 6 *Under the conditions that guarantee* $\Pi^* > \Pi'$ *and* $\xi^* > \xi'$ *in Proposition 5,* $\frac{d\Pi^*}{d\tau} < 0$ *and* $\frac{d\xi^*}{d\tau} < 0$.

Since τ makes lending intensity more costly, the negative response of Π^* is straightforward. The negative response of ξ^* then follows from the fact that higher taxes and lower rematching probabilities decrease the outside option of informed lenders, making them less restrictive in their retention of borrowers. Although alternative specifications of τ are certainly possible, I begin with the simple version described here to fix ideas. This τ can be interpreted as either a direct tax on the number of loans or a regulation which increases the cost of engaging in the matching activity.[16]

4 A Calibrated Example

Recall from Proposition 5 that high values of c are a sufficient condition for $\Pi^* > \Pi'$ and $\xi^* > \xi'$ when the decentralized equilibrium is inefficient. All

[16] The tax could also be implemented through the activities of bank examiners. As argued in Kashyap et al. (2008), the regulatory response to the recent crisis could include monitoring bank decision-making subject to the caveat that centralizing governance and restructuring employee compensation may distort the "search for performance" that allows banks to allocate resources. As I have shown, however, the decentralized allocation of internal bank resources is itself distorted so there is indeed scope for examiners to monitor the composition of a bank's workforce and levy costs accordingly.

else constant, higher c increases the probability of capital destruction by increasing the disutility of borrower effort. Economies with high c are thus ones where we might have actually expected careful lending practices, making their inefficiently low screening all the more interesting. In this section, I conduct a set of numerical exercises to investigate how high c has to be in order to guarantee both non-bindingness of the borrower participation constraints and $\Pi^* > \Pi'$. I also use the calibrated model to illustrate additional properties of the decentralized equilibrium and its response to the corrective tax proposed above.

4.1 Parameterization

For different values of c, I calibrate the model's steady state to match features of the US credit market over the period 1995-2005. Although the Gramm-Leach-Bliley Act did not officially institute broad banking until 1999, the Fed began easing Glass-Steagall in the late 1980s, effectively expanding the range of activities that banks could engage in. As such, I calibrate the model under $\tau = 0$ before considering policy experiments where $\tau \neq 0$.

I employ production functions of the form $\theta(\omega) = y_0 + y_1\omega^\alpha$, normalizing $y_0 = 1$ so that every successful project returns enough to cover its capital input. I also define the model period to be a quarter and use the standard discount factor $\beta = 0.99$. The parameters left to be calibrated are: the exogenous separation probability μ, the depletion parameter δ, and the production parameters y_1 and α.

From the 1997 Census of Manufactures, Dziczek et al. (2008) estimate that the difference between the log labour productivity of the 90th and 10th percentile manufacturing plants is 1.62 so I use this figure to target the dispersion of production among successful borrowers. I also target K/Y to match the ratio of net business loans to GDP. Defining net loans as the difference between the credit market debt and the credit market assets of non-farm non-financial businesses, FRED data yields a ratio of 0.57. To help pin down μ, I use the estimate of Bharath et al. (2011) that 71% of business loans come from lenders who recently provided the firm with another loan. The target in my model is the proportion of loans not in their first period. Finally, I use the capacity utilization rate for manufacturing, roughly 0.78 in the FRED database, to target the model's ratio of actual production to capacity. Here, capacity is defined as the production that could be achieved if, all else constant, borrowers exerted effort $e(\cdot, 0)$. The resulting values of μ, δ, α, and y_1 are listed in Table 1.

4.2 Steady State Results

Consider first the effect of c. The black squares in Figure 1 plot Π^* and ξ^* when the model is calibrated using the indicated c. The lines emanating from each black square then plot Π^* and ξ^* when c is varied but the other parameters are not. Each line starts at the lowest c for which the borrower participation constraints do not bind (i.e., the lowest $\underline{c}$). The c at which $\Pi^* = 0$ then represents $\bar{c}$. Figure 1 also plots in red the planner's solution for each parameterization. As we can see, the direction of the inefficiency in Proposition 5 is robust to a wide range of c's.

Turn now to the macroeconomic implications of $\Pi^* > \Pi'$ and $\xi^* > \xi'$. To conduct the remaining analysis, I calibrate c in addition to μ, δ, α, and y_1. BEA Economic Accounts report that the value-added of the financial industry as a fraction of GDP is 0.075 for the period under consideration. Value-added sums compensation to employees, production taxes, and gross operating surplus (at least part of which is distributed as dividends and thus not available for future loans) so I interpret δK as the model's counterpart and use $\delta K/Y$ as the additional target. The resulting parameters are $c = 0.285$, $\mu = 0.14$, $\delta = 0.13$, $\alpha = 0.5$, and $y_1 = 2.05$.

Table 2 shows how the decentralized equilibrium (MARKET) and the constrained efficiency benchmark (K-MAX) differ for several variables. It also compares the allocations of a planner who maximizes the total present discounted value of capital (K-MAX) to those of a planner who maximizes the total present discounted value of net output (W-MAX).[17] Whether the planner

[17] Given R, the net output of type ω is $e(\omega, R)\,\theta(\omega) + c\ln(1 - e(\omega, R))$. The total present

cares about capital or net output, the conclusions are the same: uninformed lending in the decentralized economy is too high, informed lending is too low, defaults are in excess, and a welfare loss obtains.[18]

4.3 Effect of the Corrective Tax

Figure 2 illustrates how the decentralized steady state varies with the lending intensity tax τ. The top two panels confirm the content of Proposition 6, namely that higher values of τ lead to lower values of Π^* and ξ^*. There are four other noteworthy features. First, market size (the measure of borrowers financed or, equivalently, total credit) exhibits a hump-shaped response to increases in τ.[19] Two competing forces drive this result. On one hand, the decline in lending intensity decreases match formation but, on the other, the decline in informed selectivity increases match preservation. The latter effect dominates at low tax rates but is eventually overtaken by the former. Second, higher values of τ increase the average quality of the credit market. Since a borrower's default probability depends on both his type and the loan rate he is charged, one of the advantages of informed lenders is that they can give borrowers better incentives to run successful projects. Although the decline in ξ^* lowers the average type financed, it (along with the increase in screening, $1-\Pi^*$) increases the proportion of financing that is informed and thus decreases the average delinquency rate. Third, production exhibits a hump-shaped response to increases in the tax. In particular, the shift towards informed financing has a positive effect as long as the frequency of new matches does not become too small. Finally, welfare increases as ξ^* and Π^* approach the efficient allocation.

Figure 3 illustrates the effect of τ on dynamics. Suppose a borrower who exerts effort e now succeeds with probability $(1+z)\,e$, where $z \in (-\varepsilon, \varepsilon)$ is an unanticipated mean-zero aggregate productivity shock that is IID over time. z is realized after all decisions have been made and is not contractible. The

discounted value of net output is thus a measure of aggregate welfare.

[18] Since uninformed loans are the main source of new lending, new lending is also too high.

[19] Note that market clearing also makes total credit equivalent to aggregate capital.

shock I consider in Figure 3 is negative and temporary, with $z_1 = -0.01$ and $z_t = 0$ for all $t \geq 2$. Appendix B describes the algorithm used to compute the responses. A negative shock to the probability of project success increases capital destruction in $t = 1$. This then implies a higher cost of funds in $t = 2$, reducing the incentive to lend and prompting both a decline in lending intensity and an increase in the informed cutoff. There is, however, a countervailing force acting on the informed cutoff: by lowering the value of being unmatched, a higher r_2 and a lower rematching probability also deteriorate the outside option of an informed lender. A small lending intensity tax reinforces the deterioration both by introducing a tangible cost to being unmatched and by prolonging the recovery path of Π_t. Although total credit still falls under this tax, a short-term decline in ξ_t combines with the increase in $1 - \Pi_t$ to limit the contraction of informed financing, thereby hastening the reaccumulation of capital and fostering a faster recovery in production.

Figure 3 also provides some insight into the effects of contractionary monetary policy. Since the fall in z occurred after all $t = 1$ decisions were made, K_2 was the only $t = 2$ state variable affected. The lower value of K_2 was entirely captured by the higher value of r_2 so the dynamics from $t = 2$ onwards can also be viewed as the dynamics following an open market operation that increases the interbank rate to r_2 at the beginning of $t = 2$.

5 Conclusion

This paper has examined the allocation of bank resources across two important intermediation activities: creating credit market matches and screening the borrowers in those matches. I began by constructing a model to disentangle the implications of this allocation decision in an environment with private information and competing lenders. I then showed that banks are inefficient at allocating resources between matching and screening, leading to too much low-quality credit. There are two externalities at play. The first operates through the distribution of available borrowers when matches can be preserved over time while the second arises because unmatched lenders also fail to internalize their effect on the informed problem. From a policy perspective, these results contribute to the current debate on bank taxes. In particular, the inefficiencies identified by my model suggest that taxing and/or regulating matching activities would be more effective than imposing a general profit tax. Indeed, steady state results show that production exhibits a hump-shaped response to increases in a matching tax and the model's dynamics indicate that a mild such tax can also attenuate business cycle fluctuations. Extending the model to evaluate different implementations of this matching tax is therefore an interesting avenue for future research.

References

Becsi, Zsolt, Victor E. Li, and Ping Wang. 2009. "Credit Mismatch and Breakdown." http://pingwang.wustl.edu/Working%20Papers/Moral-12-12-09.pdf.

Bernanke, Ben, and Mark Gertler. 1989. "Agency Costs, Net Worth and Business Fluctuations." *American Economic Review*, 79(1): 14-31.

Bharath, Sreedhar T., Sandeep Dahiya, Anthony Saunders, and Anand Srinivasan. 2011. "Lending Relationships and Loan Contract Terms." *Review of Financial Studies*, 24(4): 1141-1203.

Boot, Arnoud W.A. 2000. "Relationship Banking: What Do We Know?" *Journal of Financial Intermediation*, 9(1): 7-25.

Broecker, Thorsten. 1990. "Credit-Worthiness Tests and Interbank Competition." *Econometrica*, 58(2): 429-452.

Cao, Melanie, and Shouyong Shi. 2001. "Screening, Bidding, and the Loan Market Tightness." *European Finance Review*, 5(1-2): 21-61.

Diamond, Douglas. 1984. "Financial Intermediation and Delegated Monitoring." *Review of Economic Studies*, 51(3): 393–414.

Direr, Alexis. 2008. "Multiple Equilibria in Markets with Screening." *Journal of Money, Credit and Banking*, 40(4): 791-798.

Dziczek, Kristin M., Mark C. Long, Daniel D. Luria, and Edith A. Wiarda. 2008. "Wage and Productivity Stability in U.S. Manufacturing Plants." *Monthly Labor Review*, 131(5): 24-36.

Freixas, Xavier, and Jean-Charles Rochet. 1997. *Microeconomics of Banking*. Cambridge: MIT Press.

Gurley, John G., and E. S. Shaw. 1955. "Financial Aspects of Economic Development." *American Economic Review*, 45(4): 515-538.

Hachem, Kinda. 2011. "Relationship Lending and the Transmission of Monetary Policy." Forthcoming in *Journal of Monetary Economics*, doi:10.1016/j.jmoneco.2011.12.001.

Kashyap, Anil K., Raghuram G. Rajan, and Jeremy C. Stein. 2008. "Rethinking Capital Regulation." In: *Maintaining Stability in a Changing Financial System*, Federal Reserve Bank of Kansas City Symposium, 431-471.

Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vikrant Vig. 2010. "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans." *Quarterly Journal of Economics*, 125(1): 307-362.

Kiyotaki, Nobuhiro, and John Moore. 1997. "Credit Cycles." *Journal of Political Economy*, 105(2): 211-248.

Korinek, Anton. 2011. "Systemic Risk-Taking: Amplification Effects, Externalities, and Regulatory Responses." European Central Bank Working Paper 1345.

Lorenzoni, Guido. 2008. "Inefficient Credit Booms." *Review of Economic Studies*, 75(3): 809-833.

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

Rajan, Raghuram G. 1994. "Why Bank Credit Policies Fluctuate: A Theory and Some Evidence." *Quarterly Journal of Economics*, 109(2): 399-441.

Shleifer, Andrei, and Robert W. Vishny. 2010. "Unstable Banking." *Journal of Financial Economics*, 97(3): 306–318.

Shleifer, Andrei, and Robert W. Vishny. 2011. "Fire Sales in Finance and Macroeconomics." *Journal of Economic Perspectives*, 25(1): 29-48.

Stein, Jeremy C. 1998. "An Adverse-Selection Model of Bank Asset and Liability Management with Implications for the Transmission of Monetary Policy." *RAND Journal of Economics*, 29(3): 466-486.

Williamson, Stephen D. 1987. "Financial Intermediation, Business Failures, and Real Business Cycles." *Journal of Political Economy*, 95(6): 1196-1216.

Table 1: Calibration Results for Different Values of c

	μ	δ	α	y_1
$c = 0.1$	0.20	0.260	0.30	1.365
$c = 0.2$	0.16	0.175	0.45	1.765
$c = 0.3$	0.13	0.120	0.55	2.140
$c = 0.4$	0.11	0.060	0.60	2.415
$c = 0.5$	0.10	0.020	0.63	2.700

Figure 1: Inefficiency Results for Different Values of c

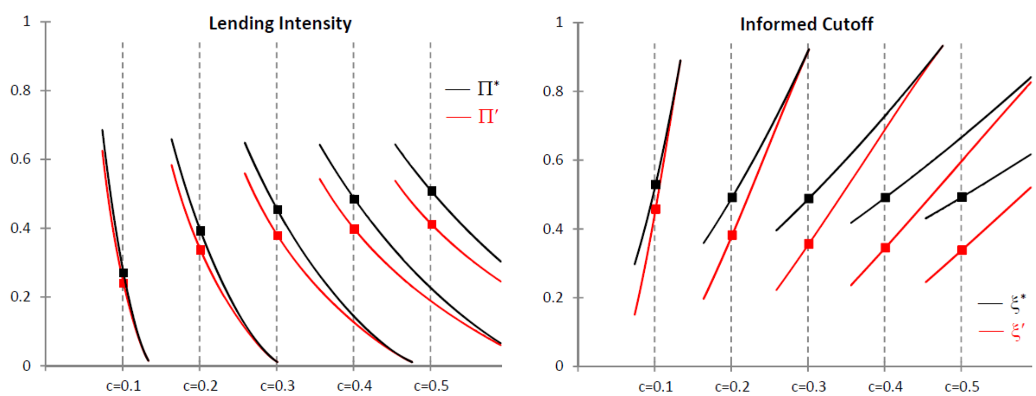


Table 2: Steady State Comparison Using $c = 0.285$

	MARKET	K-MAX	W-MAX
Lending Intensity (Π)	0.4309	0.3634	0.3637
Informed Cutoff (ξ)	0.4901	0.3667	0.3770
Amount of Informed Credit	0.4044	0.4827	0.4750
Amount of Uninformed Credit	0.1169	0.0743	0.0753
Total Credit	0.5213	0.5570	0.5503
Total New Credit	0.1512	0.1196	0.1199
Average Type Financed	0.6578	0.6399	0.6432
Average Delinquency Rate	0.3426	0.3388	0.3201
Aggregate Welfare	75.100	79.634	80.551

Figure 2: Effect of Corrective Tax on Steady State

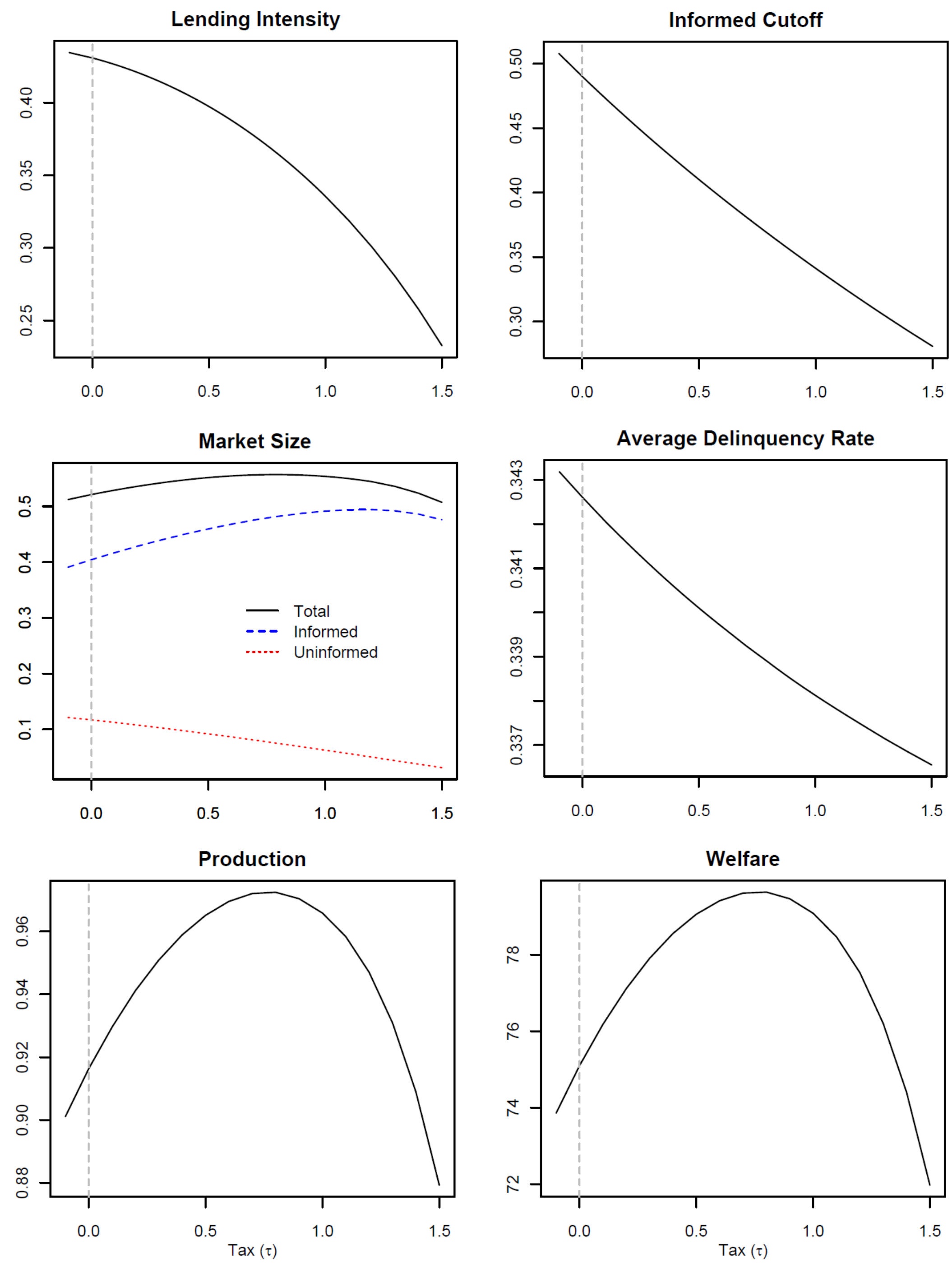
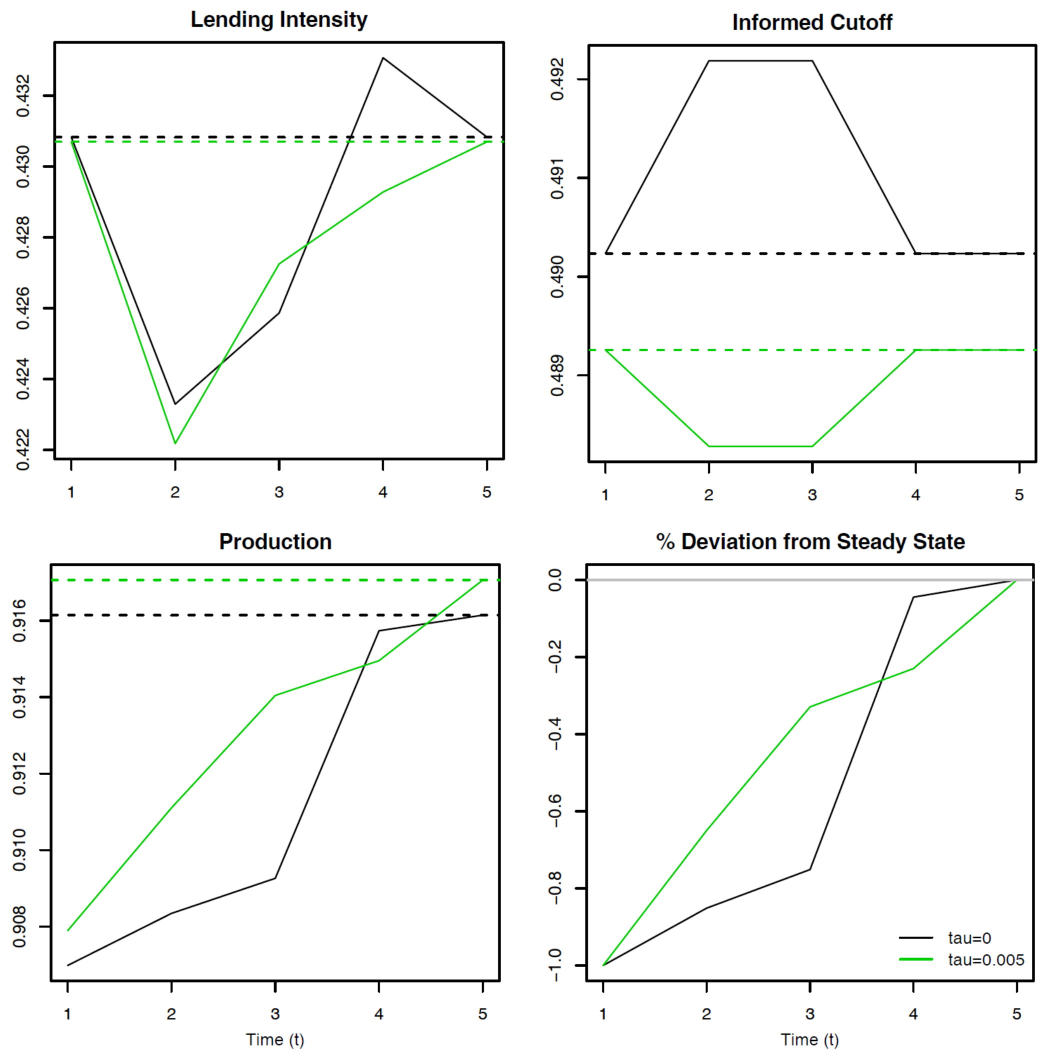


Figure 3: Effect of Corrective Tax on Dynamics



Appendix A - Proofs

Proof of Proposition 1

The proof amounts to showing that J, the informed value function, is increasing in ω so I first establish that J exists. Define indicator i and value function D such that $TD(S,\omega,v,\psi,i) = i \times D(S,\omega,v,\psi,1) + (1-i) \times D(S,0,0,\psi,0)$ where $D(S,\omega,v,\psi,1) \equiv J(\omega,v,S)$ and $D(S,0,0,\psi,0) \equiv U(S,\psi)$. Now suppose D exists in the set of bounded and continuous functions $(\mathcal{C})$. By the Theorem of the Maximum, the right-hand side of equation (2) produces $D(\cdot,1) \in \mathcal{C}$ while the right-hand side of (4) with X as per (3) produces $D(\cdot,0) \in \mathcal{C}$. Therefore, $TD \in \mathcal{C}$. We can also show that Blackwell's sufficient conditions for a contraction are satisfied so, by the Contraction Mapping Theorem, there is indeed a unique $D \in \mathcal{C}$. By implication, J and U exist and are unique, bounded, and continuous. A similar contraction mapping argument can then be used to show that J lies in the set of increasing-in-ω functions. ■

Proof of Proposition 2

In what follows, $J(\omega,v,S)$ and $U(S,\psi)$ are shortened to $J(\omega)$ and U while r is used in place of $r(S)$. I start with the value of an informed lender, ignoring the borrower's participation constraint. If ω is kept, then the optimal loan rate is $R(\omega) = \theta(\omega) - \sqrt{c\theta(\omega)}$. Defining $g(\omega) \equiv \left(\sqrt{\theta(\omega)} - \sqrt{c}\right)^2$ then gives:

$$J(\omega) = \begin{cases} \beta U & if \quad \omega < \xi \\ \frac{g(\omega)-(1+r)+\beta\mu U}{1-\beta(1-\mu)} & if \quad \omega \geq \xi \end{cases}$$

$$where$$

$$\xi = \arg\min_{x\in[0,1]} \left| g(x) - (1+r) - \beta(1-\beta)(1-\mu)U \right|$$

Turn now to the distribution of borrowers across financing class. In steady state, equations (6) and (7) simplify to:

$$\lambda(\omega) = \begin{cases} 0 & if \quad \omega < \xi \\ \frac{\Pi(1-\mu\Pi)}{\mu+(1-\mu)\Pi} & if \quad \omega \geq \xi \end{cases} \quad \text{and} \quad \phi(\omega) = \begin{cases} \Pi^2 & if \quad \omega < \xi \\ \frac{\mu\Pi^2}{\mu+(1-\mu)\Pi} & if \quad \omega \geq \xi \end{cases}$$

Equation (8) then implies the following beliefs:

$$\psi\left(\omega|\xi,\Pi\right)=\begin{cases}\frac{\mu+(1-\mu)\Pi}{\mu+(1-\mu)\Pi\xi} & if \quad \omega<\xi\\ \frac{\mu}{\mu+(1-\mu)\Pi\xi} & if \quad \omega\geq\xi\end{cases}$$

Going forward, let $\psi_L(\xi,\Pi)$ denote the first row of the above equation and let $\psi_H(\xi,\Pi)$ denote the second. Also let $\Gamma(\xi,\Pi)$ denote the expected one-period profit of an uninformed lender evaluated at the optimal choice of $\overline{R}$. Defining $h\left(\omega,\overline{R}\right)\equiv\left(1-\frac{c}{\theta(\omega)-\overline{R}}\right)\overline{R}$, we can write $\Gamma(\xi,\Pi)$ as:

$$\Gamma(\xi,\Pi)=\psi_L(\xi,\Pi)\int_{\eta\left(\overline{R}\right)}^{\xi}h\left(\omega,\overline{R}\right)d\omega+\psi_H(\xi,\Pi)\int_{\xi}^{1}h\left(\omega,\overline{R}\right)d\omega$$

The value of an unmatched lender pursuing optimal strategy π^* is then:

$$\begin{aligned} U &= \pi^{*2}\left[\Gamma(\xi,\Pi)-(1+r)+\beta(1-\mu)\int_0^1 J(\omega)\,\psi(\omega|\xi,\Pi)\,d\omega+\beta\mu U\right]\\ &\quad +\pi^*(1-\pi^*)\int_0^1 J(\omega)\,\psi(\omega|\xi,\Pi)\,d\omega+(1-\pi^*)\,\beta U\end{aligned}$$

Substituting in for $J(\omega)$, we can rearrange the above expression to isolate U. With π^* optimal, $dU/d\pi^*=0$ so differentiating the isolated expression yields an implicit definition of π^*. The definition can be simplified by combining terms to reconstitute U then using the definition of ξ to sub U out. After some algebra, we get the following best response:

$$\pi^*=\min\left\{\frac{1}{2[1-\beta(1-\mu)]}\left(\frac{\int_\xi^1[g(\omega)-g(\xi)]d\omega}{\int_\xi^1[g(\omega)-g(\xi)]d\omega+\left(\frac{g(\xi)-\Gamma(\xi,\Pi)}{\psi_H(\xi,\Pi)}\right)}\right),1\right\}\equiv\pi_l(\Pi|\xi) \qquad \text{(A.1)}$$

Note that $\pi_l(\Pi|\xi)=0$ if and only if $\xi=1$. Consider now $\xi<1$ and the interior solution for $\pi_l(\Pi|\xi)$. Taking derivatives and using the envelope theorem to replace $\Gamma_\Pi(\xi,\Pi)$ and $\Gamma_\xi(\xi,\Pi)$ yields $\frac{\partial\pi_l(\Pi|\xi)}{\partial\Pi}\propto -g(\xi)\,\xi+\int_{\overline{\eta}\left(\overline{R}\right)}^{\xi}h\left(\omega,\overline{R}\right)d\omega$ and $\frac{\partial\pi_l(\Pi|\xi)}{\partial\xi}\propto -\frac{(1-\mu)\Pi\left[g(\xi)-h\left(\xi,\overline{R}\right)\right]\int_\xi^1[g(\omega)-g(\xi)]d\omega}{\mu+(1-\mu)\Pi\xi}-\left[\int_\xi^1 g(\omega)\,d\omega-(1-\xi)\,\Gamma(\xi,\Pi)\right]g'(\xi)$. Since an informed lender can always charge type ω loan rate $\overline{R}$, $g(\omega)$ must be greater than or equal to $h\left(\omega,\overline{R}\right)$ for all ω. This inequality will be strict for at least some types, implying $\int_{\xi}^{1} g(\omega)\, d\omega > (1-\xi)\,\Gamma(\xi,\Pi)$. Combining with $g'(\omega) > 0$ and $h_{\omega}(\omega,\cdot) > 0$, we can now conclude that $\frac{\partial \pi_l(\Pi|\xi)}{\partial \xi} < 0$ and $\frac{\partial \pi_l(\Pi|\xi)}{\partial \Pi} < 0$ whenever $\pi_l(\Pi|\xi)$ is interior. To complete the proof, notice that $\pi_l(\Pi|0)$ is independent of Π. If $\pi_l(\cdot|0) < 1$, then $\frac{\partial \pi_l(\Pi|\xi)}{\partial \xi}$ and $\frac{\partial \pi_l(\Pi|\xi)}{\partial \Pi}$ imply $\widehat{\xi} = \widehat{\Pi}(\cdot) = 0$. In contrast, if $\pi_l(\cdot|0) = 1$, then the derivatives imply $\widehat{\xi} > 0$ and $\widehat{\Pi}(\cdot) \geq 0$ with $\widehat{\Pi}'(\cdot) \leq 0$. ■

Proof of Proposition 3

Abstracting again from the participation constraint, the revenue of an informed lender is given by $g(\omega)$ as defined in the proof of Proposition 2. The steady state market clearing equation can then be written as:

$$\Pi = \max\left\{\frac{1}{\mu}\left(\frac{\int_{\xi}^{1}\left[g(\omega)-\frac{1}{1-\delta}\right]d\omega}{\int_{\xi}^{1}\left[g(\omega)-\frac{1}{1-\delta}\right]d\omega+\left(\frac{\frac{1}{1-\delta}-\Gamma(\xi,\Pi)}{\psi_H(\xi,\Pi)}\right)}\right),0\right\} \equiv \pi_m(\Pi|\xi) \qquad \text{(A.2)}$$

Proving that a steady state exists amounts to proving the existence of a ξ and Π that satisfy $\Pi = \pi_l(\Pi|\xi)$ and $\Pi = \pi_m(\Pi|\xi)$, where $\pi_l(\Pi|\xi)$ and $\pi_m(\Pi|\xi)$ are as defined in equations (A.1) and (A.2) respectively. Let $\Pi_l(\xi)$ denote the solution to $\Pi = \pi_l(\Pi|\xi)$ and let $\Pi_m(\xi)$ denote the solution to $\Pi = \pi_m(\Pi|\xi)$.

If $\xi = 1$, then $\pi_l(\cdot|1) = \pi_m(\cdot|1) = 0$ and thus $\Pi_l(1) = \Pi_m(1) = 0$. Stated otherwise, there exists a trivial equilibrium with $\xi = 1$ and $\Pi = 0$ for any value of c. Focus now on $\xi < 1$. Notice that $\partial g(\omega)/\partial c < 0$ and $\partial \Gamma(\xi,\Pi)/\partial c < 0$ so $\partial \pi_m(\Pi|\xi)/\partial c < 0$ and thus $\partial \Pi_m(\xi)/\partial c < 0$ when $\int_{\xi}^{1}\left[g(\omega) - \frac{1}{1-\delta}\right] d\omega > 0$. In other words, higher values of c decrease the feasible values of Π. We can also show that $\Pi_m(\cdot) = 0$ if $c = \theta(0)$. To see how, note that $\Pi_m(\xi) > 0$ requires $\pi_m(\Pi_m(\xi)|\xi) > 0$ and a necessary condition for the latter is $g(1) > \frac{1}{1-\delta}$ or, equivalently, $\theta(1) > \left(\sqrt{\frac{1}{1-\delta}} + \sqrt{c}\right)^2$. This requires at least $\theta(1) > (1+\sqrt{c})^2$ which, given Assumption 2, cannot be true for $c = \theta(0)$. Since $\partial g(\omega)/\partial c < 0$, we can further conclude that there exists a $\overline{c} \leq \theta(0)$ such that $\Pi_m(\xi) = 0$ for all ξ if and only if $c \geq \overline{c}$. As a result, there is no non-trivial steady state with non-binding borrower participation constraints when $c \geq \overline{c}$.

Consider then $\xi < 1$ and $c < \overline{c}$. That there exists one and only one point such that $\Pi_l(\xi) = \Pi_m(\xi) > 0$ is established in a series of lemmas.

Lemma 1 *Any non-trivial steady state has* $\Pi \in (0, 1)$*, allowing us to use the interior solutions for* $\pi_l(\Pi|\xi)$ *and* $\pi_m(\Pi|\xi)$ *from hereon in.*

Proof. Non-triviality rules out $\Pi = 0$ so focus on $\Pi = 1$. (A.2) reduces to $(1-\mu)\int_\xi^1 g(\omega)\,d\omega + [\mu + (1-\mu)\xi]\,\Gamma(\xi, 1) = \frac{1}{1-\delta}$. Since $g(\cdot) > 0$, we have $\int_\xi^1 g(\omega)\,d\omega < \int_0^1 g(\omega)\,d\omega$. Moreover, $h\left(\omega, \overline{R}\right) \leq g(\omega)$, $\eta\left(\overline{R}\right) \geq 0$, $\psi_L(\xi, \Pi) \geq \psi_H(\xi, \Pi)$, and $\psi_H(\xi, \Pi) \leq 1$ imply $\Gamma(\xi, \Pi) < \int_0^1 g(\omega)\,d\omega$. Therefore, given Assumption 1, $\Pi = 1$ cannot satisfy market clearing. □

Lemma 2 $\Pi_l(\xi)$ *and* $\Pi_m(\xi)$ *intersect at least once.*

Proof. Define $\underline{\xi}$ and $\overline{\xi}$ such that $\int_{\underline{\xi}}^1 \left[g(\omega) - \frac{1}{1-\delta}\right] d\omega = 0$ and $g\left(\overline{\xi}\right) = \frac{1}{1-\delta}$ respectively. Some algebra then yields $\Pi_l\left(\underline{\xi}\right) > \Pi_m\left(\underline{\xi}\right)$ and $\Pi_l(\xi) < \Pi_m(\xi)$ for all $\xi \geq \overline{\xi}$. □

Lemma 3 *All intersections between* $\Pi_l(\xi)$ *and* $\Pi_m(\xi)$ *are associated with the same value of* Π*, labelled* Π_0*.*

Proof. Rearrange $\Pi = \pi_l(\Pi|\xi)$ and $\Pi = \pi_m(\Pi|\xi)$ to get two expressions for $\frac{\Gamma(\xi,\Pi)}{\psi_H(\xi,\Pi)}$. Equating these expressions and rearranging again yields a quadratic in Π, where the roots of this quadratic determine the values Π can achieve at an intersection. Recall from the proof of Lemma 2 that intersections require $\xi < \overline{\xi}$ which, given $g'(\cdot) > 0$, is equivalent to $g(\xi) < \frac{1}{1-\delta}$. This fact combined with $\Pi > 0$ can be used to eliminate one of the roots, implying that any intersection achieves the same value of Π. □

Lemma 4 $\Pi_l'(\xi) < 0$ *so there is only one value of* ξ *such that* $\Pi_l(\xi) = \Pi_0$*.*

Proof. Totally differentiate $\Pi = \pi_l(\Pi|\xi)$. Based on the resulting expression, a sufficient condition for $\Pi_l'(\xi) < 0$ is $\Gamma(\xi, \Pi) \leq \frac{1}{1-\xi}\int_\xi^1 g(\omega)\,d\omega$. Since $h\left(\omega, \overline{R}\right) \leq g(\omega)$, $\eta\left(\overline{R}\right) \geq 0$, $\psi_L(\xi, \Pi) \geq \psi_H(\xi, \Pi)$, and $\psi_H(\xi, \Pi) \leq 1$, this condition is satisfied. □

Conditional on the participation constraints not binding, we can now conclude that there is a unique point such that $\Pi_l(\xi) = \Pi_m(\xi) > 0$ when $c < \overline{c}$. Completing the proof thus requires finding values of c for which the unconstrained informed loan rate, $R(\omega) = \theta(\omega) - \sqrt{c\theta(\omega)}$, does indeed induce ω to participate. The steady state participation constraint for type $\omega \geq \xi$ simplifies to $\Theta(\omega, R(\omega)) \geq \left(\frac{\beta(1-\mu)\Pi^2}{1+\beta(1-\mu)\Pi^2}\right)\Theta\left(\omega, \overline{R}\right)$ where $\Theta(\omega, R) \equiv \theta(\omega) - R - c + c\ln\left(\frac{c}{\theta(\omega)-R}\right)$. Note that $\Theta(\omega, R(\omega)) > 0$ for all $c \in (0, \theta(0))$ while $\Pi = 0$ in the extreme case of $c = \overline{c}$. Therefore, there exists a $\underline{c} < \overline{c}$ such that the participation constraint is indeed satisfied when $c \in (\underline{c}, \overline{c})$. ■

Proof of Proposition 4

Let $W_I(\omega)$ denote the present discounted value of putting ω in an informed match, $W_U(\omega)$ the value of putting him in an uninformed match, and $W_N(\omega)$ the value of keeping him unmatched. As before, $\lambda(\omega)$ denotes the proportion of ω's in informed matches and $\phi(\omega)$ the proportion in uninformed matches. For a given Π and ξ, the proportions are still governed by equations (6) and (7), with steady state values as in the proof of Proposition 2. Measured just prior to production, the planner's objective function is then:

$$\mathcal{W} = \int_0^1 W_I(\omega)\lambda(\omega)\,d\omega + \int_0^1 W_U(\omega)\phi(\omega)\,d\omega + \int_0^1 \beta W_N(\omega)\left[1 - \lambda(\omega) - \phi(\omega)\right]d\omega$$

Suppressing their dependence on the choice variables, the steady state functional equations are:

$$W_I(\omega) = \begin{cases} \beta W_N(\omega) & if \quad \omega < \xi \\ e(\omega, R(\omega))\,R(\omega) + \beta(1-\mu)W_I(\omega) + \beta\mu W_N(\omega) & if \quad \omega \geq \xi \end{cases}$$

$$W_U(\omega) = \begin{cases} e\left(\omega, \overline{R}\right)\overline{R} + \beta W_N(\omega) & if \quad \omega < \xi \\ e\left(\omega, \overline{R}\right)\overline{R} + \beta(1-\mu)W_I(\omega) + \beta\mu W_N(\omega) & if \quad \omega \geq \xi \end{cases}$$

$$W_N(\omega) = \begin{cases} \Pi^2 W_U(\omega) + (1-\Pi^2)\beta W_N(\omega) & if \quad \omega < \xi \\ \Pi^2 W_U(\omega) + \Pi(1-\Pi)W_I(\omega) + (1-\Pi)\beta W_N(\omega) & if \quad \omega \geq \xi \end{cases}$$

Solving for $W_I(\omega)$, $W_U(\omega)$, and $W_N(\omega)$ then substituting into the objective function (along with the steady state versions of $\lambda(\omega)$ and $\phi(\omega)$) yields:

$$\mathcal{W} = \Pi^2 \int_{\eta\left(\overline{R}\right)}^{\xi} \frac{e\left(\omega,\overline{R}\right)\overline{R}}{1-\beta} d\omega + \frac{\mu\Pi^2}{\mu+(1-\mu)\Pi} \int_{\xi}^{1} \frac{e\left(\omega,\overline{R}\right)\overline{R}}{1-\beta} d\omega + \frac{\Pi(1-\mu\Pi)}{\mu+(1-\mu)\Pi} \int_{\xi}^{1} \frac{e(\omega,R(\omega))R(\omega)}{1-\beta} d\omega$$

Recall that the planner's problem is to choose Π, ξ, $\overline{R}$, and $R(\cdot)$ in order to maximize $\mathcal{W}$ subject to an aggregate feasibility constraint. The constraint requires that the amount of capital allocated to firms equals the amount of capital available to the planner each period. It is thus equivalent to the market clearing equation presented earlier. Letting γ denote the Lagrange multiplier, the planner's Lagrangian can now be written as in the statement of Proposition 4. ■

Proof of Proposition 5

I start by deriving the planner's first order conditions from the Lagrangian in Proposition 4. Optimizing with respect to $R(\cdot)$ and $\overline{R}$ yields:

$$R(\omega) = \theta(\omega) - \sqrt{c\theta(\omega)}$$

$$\int_{\eta\left(\overline{R}\right)}^{\xi} \left(1 - \frac{c\theta(\omega)}{\left[\theta(\omega)-\overline{R}\right]^2}\right) d\omega + \frac{\mu}{\mu+(1-\mu)\Pi} \int_{\xi}^{1} \left(1 - \frac{c\theta(\omega)}{\left[\theta(\omega)-\overline{R}\right]^2}\right) d\omega = 0$$

Note that these are the same equations generated by the decentralized economy so, as desired, there are no direct inefficiencies stemming from $R(\cdot)$ and $\overline{R}$. Stated otherwise, the market's choice of $\overline{R}$ will be inefficient if and only if ξ and/or Π is inefficient. Consider now the first order conditions for ξ and Π. Optimizing with respect to ξ yields:

$$\frac{(1-\mu\Pi)}{1-\mu\Pi-(1-\mu)\Pi^2} g(\xi) - \frac{(1-\mu)\Pi^2}{1-\mu\Pi-(1-\mu)\Pi^2} h\left(\xi,\overline{R}\right) = \frac{\gamma(1-\beta)}{1+\gamma(1-\beta)(1-\delta)}$$

where $g(\cdot)$ and $h\left(\cdot,\overline{R}\right)$ are as defined in the proof of Proposition 2. After some algebra, optimization with respect to Π yields:

$$2\Pi\left(\int_{\xi}^{1}\left[g\left(\omega\right)-\tfrac{\gamma(1-\beta)}{1+\gamma(1-\beta)(1-\delta)}\right]d\omega+\frac{\frac{\gamma(1-\beta)}{1+\gamma(1-\beta)(1-\delta)}-\Gamma(\xi,\Pi)}{\psi_H(\xi,\Pi)}\right)$$

$$=\quad \tfrac{1}{\mu+(1-\mu)\Pi}\int_{\xi}^{1}\left[g\left(\omega\right)-\tfrac{\gamma(1-\beta)}{1+\gamma(1-\beta)(1-\delta)}\right]d\omega+\tfrac{(1-\mu)\Pi^2}{\mu+(1-\mu)\Pi}\int_{\xi}^{1}\left[g\left(\omega\right)-h\left(\omega,\overline{R}\right)\right]d\omega$$

Combining the first order conditions for ξ and Π then produces:

$$\Pi=\tfrac{1+(1-\mu)\Pi^2}{2[\mu+(1-\mu)\Pi]}\left(\frac{\int_{\xi}^{1}[g(\omega)-g(\xi)]d\omega-Q(\xi,\Pi)}{\int_{\xi}^{1}[g(\omega)-g(\xi)]d\omega+\left(\frac{g(\xi)-\Gamma(\xi,\Pi)}{\psi_H(\xi,\Pi)}\right)}\right)\equiv\pi_e\left(\Pi|\xi\right) \tag{A.3}$$

$$where$$

$$Q\left(\xi,\Pi\right)\equiv\tfrac{(1-\mu)\Pi^2[\mu+(1-\mu)\Pi]}{1+(1-\mu)\Pi^2}\left[\frac{\Pi[\mu(1+\xi)+2(1-\mu)\Pi\xi]\left[g(\xi)-h\left(\xi,\overline{R}\right)\right]}{\mu[1-\mu\Pi-(1-\mu)\Pi^2]}+\frac{\int_{\xi}^{1}\left[h\left(\omega,\overline{R}\right)-h\left(\xi,\overline{R}\right)\right]d\omega}{[\mu+(1-\mu)\Pi]}\right]$$

The first order condition for γ just returns the aggregate feasibility constraint which, as explained in the proof of Proposition 4, is equivalent to the market clearing equation. Therefore, the constrained efficient solution is implicitly defined by $\Pi=\pi_e\left(\Pi|\xi\right)$ and $\Pi=\pi_m\left(\Pi|\xi\right)$ whereas the decentralized equilibrium was implicitly defined by $\Pi=\pi_l\left(\Pi|\xi\right)$ and $\Pi=\pi_m\left(\Pi|\xi\right)$. The two parts of Proposition 5 can now be proven:

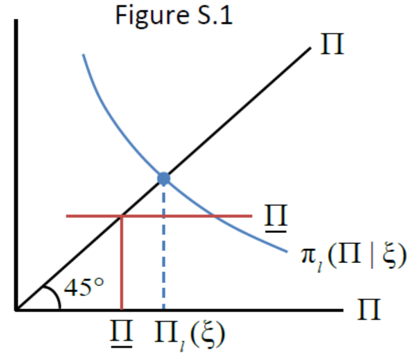
1. If $\mu=1$, then $\pi_e\left(\Pi|\xi\right)=\pi_l\left(\Pi|\xi\right)$. The decentralized equilibrium and the constrained efficient allocation are thus defined by the same system of equations and, from Proposition 3, this system has a unique solution. ■

2. *Inefficiency*: Let $\Pi_j\left(\xi\right)$ denote the solution to $\Pi=\pi_j\left(\Pi|\xi\right)$ for $j\in\{e,l,m\}$. The decentralized equilibrium is defined by $\Pi_l\left(\xi^*\right)=\Pi_m\left(\xi^*\right)$ while the constrained efficient allocation is defined by $\Pi_e\left(\xi'\right)=\Pi_m\left(\xi'\right)$ so showing that $\Pi_e\left(\cdot\right)$ never intersects $\Pi_l\left(\cdot\right)$ will be enough to show $\xi^*\neq\xi'$ and thus $\left(\xi^*,\Pi^*\right)\neq\left(\xi',\Pi'\right)$. In fact, the lack of an intersection just needs to be established over the interval $\xi\in\left[\widehat{\xi},1\right)$ since Proposition 2 and Lemma 1 imply $\xi^*\geq\widehat{\xi}$.[20] More-

[20]To see why, consider any $\xi_0<\widehat{\xi}$. From Proposition 2, $\pi_l\left(\Pi|\xi_0\right)=1$ for all Π so

over, when $\xi \in \left[\widehat{\xi}, 1\right)$, Proposition 2 says that $\pi_l(\Pi|\xi)$ is downward-sloping in Π so $\pi_e(\Pi|\xi) < \pi_l(\Pi|\xi)$ for all $\Pi \geq \Pi_l(\xi)$ would guarantee $\Pi_e(\xi) < \Pi_l(\xi)$. Figure S.1 illustrates why (ignore $\underline{\Pi}$ for the moment). If $\pi_e(\Pi|\xi) < \pi_l(\Pi|\xi)$ for all $\Pi \geq \Pi_l(\xi)$, then $\pi_e(\Pi|\xi) < \Pi$ for all $\Pi \geq \Pi_l(\xi)$ so any Π that yields $\pi_e(\Pi|\xi) = \Pi$ – that is, any $\Pi_e(\xi)$ – must be less than $\Pi_l(\xi)$. A sufficient condition for inefficiency of the decentralized equilibrium is thus $\pi_e(\Pi|\xi) < \pi_l(\Pi|\xi)$ for all $\Pi \geq \Pi_l(\xi)$ and all $\xi \in \left[\widehat{\xi}, 1\right)$.[21]

Figure S.1



Fix any such ξ. With $Q(\xi, \Pi) \geq 0$, a sufficient condition for $\pi_e(\Pi|\xi) < \pi_l(\Pi|\xi)$ is $\frac{1+(1-\mu)\Pi^2}{\mu+(1-\mu)\Pi} < \frac{1}{1-\beta(1-\mu)}$ or, equivalently, $\Pi \in \left(\frac{1-\sqrt{S(\beta)}}{2[1-\beta(1-\mu)]}, \frac{1+\sqrt{S(\beta)}}{2[1-\beta(1-\mu)]}\right) \equiv (\underline{\Pi}, \overline{\Pi})$ where $S(\beta) \equiv 1 - 4(1-\beta)[1-\beta(1-\mu)]$. Note that $\beta > \frac{1}{2-\mu}$ ensures both $S(\beta) > 0$ and $\overline{\Pi} > 1$. Therefore, the sufficient condition for $\pi_e(\Pi|\xi) < \pi_l(\Pi|\xi)$ is just $\Pi > \underline{\Pi}$ and it will be enough to show $\Pi_l(\xi) > \underline{\Pi}$ in order to show $\pi_e(\Pi|\xi) < \pi_l(\Pi|\xi)$ for all $\Pi \geq \Pi_l(\xi)$. Return to Figure S.1 and notice that $\Pi_l(\xi) > \underline{\Pi}$ will certainly be true if $\pi_l(\Pi|\xi) > \underline{\Pi}$ for all Π. The latter reduces to $\sqrt{S(\beta)} > \frac{\frac{g(\xi)-\Gamma(\xi,\Pi)}{f_H(\xi,\Pi)}}{\int_\xi^1 [g(\omega)-g(\xi)]d\omega + \frac{g(\xi)-\Gamma(\xi,\Pi)}{f_H(\xi,\Pi)}}$, abbreviated as $LHS > RHS$ and satisfied by $g(\xi) < \Gamma(\xi, \Pi)$.[22] If instead $g(\xi) > \Gamma(\xi, \Pi)$, then RHS is positive but less

$\Pi_l(\xi_0) = 1$. From Lemma 1 though, $\Pi = 1$ violates market clearing – that is, $\pi_m(1|\cdot) \neq 1$ – so $\Pi_m(\xi_0) \neq 1$. As a result, $\Pi_l(\xi_0) \neq \Pi_m(\xi_0)$ and $\xi^* \neq \xi_0$.

[21] For an arbitrary ξ, note that $\pi_e(0|\xi) = \frac{1-\beta(1-\mu)}{\mu}\pi_l(0|\xi) > \pi_l(0|\xi) > 0$ so $\pi_e(\Pi_l(\xi)|\xi) < \pi_l(\Pi_l(\xi)|\xi) \equiv \Pi_l(\xi)$ will also establish $\Pi = \pi_e(\Pi|\xi)$ for some $\Pi \in (0, \Pi_l(\xi))$. In other words, it will also establish the existence of $\Pi_e(\xi)$.

[22] $\int_\xi^1 [g(\omega) - g(\xi)]\, d\omega + \frac{g(\xi)-X(\xi,\Pi)}{f_H(\xi,\Pi)}$ is always positive since $g(\omega) \geq h(\omega, \overline{R})$ and $g'(\omega) > 0$.

than 1 for $\xi < 1$ while LHS increases with β and goes to 1 as β goes to 1. Since RHS is independent of β, high values of β thus ensure $LHS > RHS$ and, therefore, $\Pi_l(\xi) > \underline{\Pi}$. As explained above, this then ensures $\pi_e(\Pi|\xi) < \pi_l(\Pi|\xi)$ for all $\Pi \geq \Pi_l(\xi)$. Recall that ξ was set arbitrarily so we can now conclude that $\pi_e(\Pi|\xi) < \pi_l(\Pi|\xi)$ for all $\Pi \geq \Pi_l(\xi)$ and all $\xi \in \left[\widehat{\xi}, 1\right)$. ■

Direction of inefficiency: Recall that $\Pi_l\left(\underline{\xi}\right) > \Pi_m\left(\underline{\xi}\right)$ and $\Pi_l\left(\overline{\xi}\right) < \Pi_m\left(\overline{\xi}\right)$ from the proof of Lemma 2. Uniqueness implies only one point satisfying $\Pi_l(\xi^*) = \Pi_m(\xi^*)$ so it follows that $\Pi_l(\xi) < \Pi_m(\xi)$ for all $\xi \in (\xi^*, 1)$. Combining with $\widehat{\xi} \leq \xi^*$ and $\Pi_e(\xi) < \Pi_l(\xi)$ for all $\xi \in \left[\widehat{\xi}, 1\right)$ as shown above, we can also conclude that $\Pi_e(\xi) < \Pi_m(\xi)$ for all $\xi \in [\xi^*, 1)$. In other words, only $\xi' < \xi^*$ can satisfy $\Pi_e(\xi') = \Pi_m(\xi')$ and be part of the efficient allocation. Turn now to Π'. Both (ξ^*, Π^*) and (ξ', Π') satisfy $\Pi = \Pi_m(\xi)$ so $\Pi_m'(\xi) > 0$ for all $\xi \in \left(\underline{\xi}, \xi^*\right]$ would ensure $\Pi' < \Pi^*$ given $\xi' < \xi^*$.[23] Differentiating $\Pi_m(\xi) = \pi_m(\Pi_m(\xi)|\xi)$ yields $\Pi_m'(\xi) = \frac{\partial \pi_m(\Pi|\xi)/\partial \xi}{1 - \partial \pi_m(\Pi|\xi)/\partial \Pi} \propto A(\xi, \Pi)$ where $A(\xi, \Pi) \equiv (1 - \mu\Pi)\left[\frac{1}{1-\delta} - g(\xi)\right] - (1-\mu)\Pi^2\left[\frac{1}{1-\delta} - h\left(\xi, \overline{R}(\xi, \Pi)\right)\right]$ and Π is evaluated at $\Pi_m(\xi)$. Any critical point of $\Pi_m(\xi)$, denoted as $\left(\widetilde{\xi}, \widetilde{\Pi}\right)$, is implicitly defined by $A\left(\widetilde{\xi}, \widetilde{\Pi}\right) = 0$ and $\widetilde{\Pi} = \pi_m\left(\widetilde{\Pi}|\widetilde{\xi}\right)$ which combine to give:

$$\mu\widetilde{\Pi} = \frac{\left[\frac{1}{1-\delta} - h\left(\widetilde{\xi}, \overline{R}\left(\widetilde{\xi}, \widetilde{\Pi}\right)\right)\right] \int_{\widetilde{\xi}}^{1}\left[g(\omega) - \frac{1}{1-\delta}\right]d\omega - \left[\frac{1}{1-\delta} - g\left(\widetilde{\xi}\right)\right]\left[\frac{1}{1-\delta} - \Gamma\left(\widetilde{\xi}, \widetilde{\Pi}\right)\right]\widetilde{\xi}}{\left[\frac{1}{1-\delta} - h\left(\widetilde{\xi}, \overline{R}\left(\widetilde{\xi}, \widetilde{\Pi}\right)\right)\right]\left(\int_{\widetilde{\xi}}^{1}\left[g(\omega) - \frac{1}{1-\delta}\right]d\omega + \left[\frac{1}{1-\delta} - \Gamma\left(\widetilde{\xi}, \widetilde{\Pi}\right)\right]\right) - \left[\frac{1}{1-\delta} - g\left(\widetilde{\xi}\right)\right]\left[\frac{1}{1-\delta} - \Gamma\left(\widetilde{\xi}, \widetilde{\Pi}\right)\right]\widetilde{\xi}} \equiv p\left(\widetilde{\xi}, \widetilde{\Pi}\right)$$

Consider now (ξ^*, Π^*) which, as described earlier, is defined by $\Pi^* = \pi_l(\Pi^*|\xi^*)$ and $\Pi^* = \pi_m(\Pi^*|\xi^*)$. Combining these two equations yields:

$$\mu\Pi^* = \frac{\frac{\mu}{2[1-\beta(1-\mu)]}\left[\frac{1}{1-\delta} - \Gamma(\xi^*, \Pi^*)\right] \int_{\xi^*}^{1}[g(\omega) - g(\xi^*)]d\omega - [g(\xi^*) - \Gamma(\xi^*, \Pi^*)] \int_{\xi^*}^{1}\left[g(\omega) - \frac{1}{1-\delta}\right]d\omega}{\left[\frac{1}{1-\delta} - \Gamma(\xi^*, \Pi^*)\right] \int_{\xi^*}^{1}[g(\omega) - g(\xi^*)]d\omega - [g(\xi^*) - \Gamma(\xi^*, \Pi^*)] \int_{\xi^*}^{1}\left[g(\omega) - \frac{1}{1-\delta}\right]d\omega} \equiv q(\xi^*, \Pi^*)$$

We can now say that $\widetilde{\Pi} = \pi_m\left(\widetilde{\Pi}|\widetilde{\xi}\right)$ and $\mu\widetilde{\Pi} = p\left(\widetilde{\xi}, \widetilde{\Pi}\right)$ implicitly define $\left(\widetilde{\xi}, \widetilde{\Pi}\right)$ while $\Pi^* = \pi_l(\Pi^*|\xi^*)$ and $\mu\Pi^* = q(\xi^*, \Pi^*)$ implicitly define (ξ^*, Π^*).

Lemma 5 *If* $q\left(\widetilde{\xi}, \cdot\right) < p\left(\widetilde{\xi}, \cdot\right)$, *then* $\Pi_m'(\xi) > 0$ *for all* $\xi \in \left(\underline{\xi}, \xi^*\right]$.

[23] Recall that $\Pi_m(\xi)$ is only defined for $\xi \in \left[\underline{\xi}, 1\right]$, with $\Pi_m\left(\underline{\xi}\right) = 0$.

Proof. Let $\Pi_q(\xi)$ denote the solution to $\mu\Pi = q(\xi, \Pi)$ and let $\Pi_p(\xi)$ denote the solution to $\mu\Pi = p(\xi, \Pi)$. If $q\left(\widetilde{\xi}, \cdot\right) < p\left(\widetilde{\xi}, \cdot\right)$, then $\Pi_q\left(\widetilde{\xi}\right) < \Pi_p\left(\widetilde{\xi}\right)$. By definition though, $\Pi_p\left(\widetilde{\xi}\right) = \widetilde{\Pi} = \Pi_m\left(\widetilde{\xi}\right)$ so $\Pi_q\left(\widetilde{\xi}\right) < \Pi_p\left(\widetilde{\xi}\right)$ implies $\Pi_q\left(\widetilde{\xi}\right) < \Pi_m\left(\widetilde{\xi}\right)$. In contrast, $\Pi_q\left(\underline{\xi}\right) = \frac{1}{2[1-\beta(1-\mu)]} > 0 = \Pi_m\left(\underline{\xi}\right)$. We can thus conclude that $\Pi_q(\cdot)$ intersects $\Pi_m(\cdot)$ for some $\xi \in \left(\underline{\xi}, \widetilde{\xi}\right)$. Since any such intersection constitutes the unique decentralized equilibrium, we have just shown $\xi^* \in \left(\underline{\xi}, \widetilde{\xi}\right)$. The analysis applies for any critical point $\left(\widetilde{\xi}, \widetilde{\Pi}\right)$ so it applies in particular for the first critical point. Given $\Pi_m\left(\underline{\xi}\right) = 0$ and $\Pi_m\left(\underline{\xi} + \varepsilon\right) > 0$, the first critical point is a maximum and, therefore, $\Pi_m'(\xi) > 0$ for all $\xi \in \left(\underline{\xi}, \xi^*\right]$. $\square$

With Lemma 5 in hand, showing $p\left(\widetilde{\xi}, \cdot\right) > q\left(\widetilde{\xi}, \cdot\right)$ will be enough to establish $\Pi' < \Pi^*$. Since $h\left(\omega, \overline{R}\right) \leq g(\omega)$ for all ω, a lowerbound for $p\left(\widetilde{\xi}, \cdot\right)$ is obtained by replacing $h\left(\widetilde{\xi}, \overline{R}\left(\widetilde{\xi}, \cdot\right)\right)$ with $g\left(\widetilde{\xi}\right)$ in the definition of $p\left(\widetilde{\xi}, \cdot\right)$. Moreover, since $\frac{\mu}{2[1-\beta(1-\mu)]} < \frac{1}{2}$, an upperbound for $q\left(\widetilde{\xi}, \cdot\right)$ is obtained by replacing $\frac{\mu}{2[1-\beta(1-\mu)]}$ with $\frac{1}{2}$ in the definition of $q\left(\widetilde{\xi}, \cdot\right)$. A sufficient condition for $p\left(\widetilde{\xi}, \cdot\right) > q\left(\widetilde{\xi}, \cdot\right)$ is then that the lowerbound of $p\left(\widetilde{\xi}, \cdot\right)$ exceeds the upperbound of $q\left(\widetilde{\xi}, \cdot\right)$. Defining $B(\xi) \equiv \frac{2}{1-\delta} - (1+\xi)\, g(\xi) - \int_{\xi}^{1} g(\omega)\, d\omega$, this condition simplifies to $B\left(\widetilde{\xi}\right)\left(1 - \widetilde{\xi}\right)\Gamma\left(\widetilde{\xi}, \Pi\right) \geq B\left(\widetilde{\xi}\right)\int_{\widetilde{\xi}}^{1} g(\omega)\, d\omega$. Having $h\left(\omega, \overline{R}\right) \leq g(\omega)$ for all ω (with strict inequality for at least some ω) implies $\left(1 - \widetilde{\xi}\right)\Gamma\left(\widetilde{\xi}, \Pi\right) < \int_{\widetilde{\xi}}^{1} g(\omega)\, d\omega$ so the sufficient condition is true if and only if $B\left(\widetilde{\xi}\right) \leq 0$ or, equivalently:

$$\int_{\widetilde{\xi}}^{1} \left[g(\omega) - \tfrac{1}{1-\delta}\right] d\omega \geq \left(1 + \widetilde{\xi}\right)\left[\tfrac{1}{1-\delta} - g\left(\widetilde{\xi}\right)\right] \tag{A.4}$$

Recall that differentiating $\Pi_m(\xi) = \pi_m(\Pi_m(\xi)\,|\,\xi)$ gives $\Pi_m'(\xi) = \frac{\partial\pi_m(\Pi|\xi)/\partial\xi}{1-\partial\pi_m(\Pi|\xi)/\partial\Pi}$ with Π evaluated at $\Pi_m(\xi)$. Since $\Pi_m'\left(\widetilde{\xi}\right) = 0$ at any critical point and $\pi_m(\cdot|\xi)$ is well behaved, it follows that $\left.\frac{\partial\pi_m\left(\widetilde{\Pi}|\xi\right)}{\partial\xi}\right|_{\xi=\widetilde{\xi}} = 0$. Lemma 6 completes the proof of Proposition 5 by establishing a contradiction of $\left.\frac{\partial \pi_m\left(\widetilde{\Pi}|\xi\right)}{\partial \xi}\right|_{\xi=\widetilde{\xi}} = 0$ if (A.4) is not satisfied.

Lemma 6 *If (A.4) is untrue, then* $\left.\frac{\partial \pi_m(\cdot|\xi)}{\partial \xi}\right|_{\xi=\widetilde{\xi}} > 0$ *for high* c *not exceeding* $\overline{c}$.

Proof. Write $\pi_m\left(\Pi|\xi\right) = \frac{1}{\mu}\left(\frac{1}{1+M(\Pi|\xi)}\right)$ where $M\left(\Pi|\xi\right) \equiv \frac{\frac{1}{1-\delta}-\Gamma(\xi,\Pi)}{\frac{\mu}{\mu+(1-\mu)\Pi\xi}\int_{\xi}^{1}\left[g(\omega)-\frac{1}{1-\delta}\right]d\omega}$. Note that $\frac{\partial}{\partial \xi}\left(\frac{1}{1-\delta} - \Gamma\left(\xi, \Pi\right)\right) \propto \Gamma\left(\xi, \Pi\right) - h\left(\xi, \overline{R}\right)$ is negative if and only if $-\int_0^{\eta\left(\overline{R}\right)} h\left(\xi, \overline{R}\right) d\omega + \int_{\eta\left(\overline{R}\right)}^{\xi}\left[h\left(\omega, \overline{R}\right) - h\left(\xi, \overline{R}\right)\right] d\omega + \frac{\mu\int_{\xi}^{1}\left[h\left(\omega,\overline{R}\right)-h\left(\xi,\overline{R}\right)\right]d\omega}{\mu+(1-\mu)\Pi} < 0$. Given $h_\omega\left(\omega, \cdot\right) > 0$, the latter is true at sufficiently high ξ. Differentiating the equation that defines $\underline{\xi}$, namely $\int_{\underline{\xi}}^{1}\left[g\left(\omega\right) - \frac{1}{1-\delta}\right] d\omega = 0$ where $g\left(\omega\right) = \left(\sqrt{\theta\left(\omega\right)} - \sqrt{c}\right)^2$, yields $\partial\underline{\xi}/\partial c > 0$. That is, higher values of c increase the lowest value of ξ that can be considered, prompting $\Gamma\left(\xi, \Pi\right) < h\left(\xi, \overline{R}\right)$ and thus $\frac{\partial}{\partial \xi}\left(\frac{1}{1-\delta} - \Gamma\left(\xi, \Pi\right)\right) < 0$. Consider now $\left.\frac{\partial}{\partial \xi}\left(\frac{\mu}{\mu+(1-\mu)\Pi\xi}\int_{\xi}^{1}\left[g\left(\omega\right) - \frac{1}{1-\delta}\right] d\omega\right)\right|_{\xi=\widetilde{\xi}} \propto \left(\frac{\mu}{(1-\mu)\widetilde{\Pi}} + \widetilde{\xi}\right)\left[\frac{1}{1-\delta} - g\left(\widetilde{\xi}\right)\right] - \int_{\widetilde{\xi}}^{1}\left[g\left(\omega\right) - \frac{1}{1-\delta}\right] d\omega$. If (A.4) does not hold, then low $\widetilde{\Pi}$ will guarantee that $\left.\frac{\partial}{\partial \xi}\left(\frac{\mu}{\mu+(1-\mu)\Pi\xi}\int_{\xi}^{1}\left[g\left(\omega\right) - \frac{1}{1-\delta}\right] d\omega\right)\right|_{\xi=\widetilde{\xi}}$ cannot be very negative. Recall from the proof of Proposition 3 that higher values of c decrease the feasible values of Π (in the extreme case of $c = \overline{c}$, only $\Pi = 0$ is feasible), thus completing the proof. □ ■

Proof of Proposition 6

With the tax, $\pi_l\left(\Pi|\xi\right) = \frac{1}{2[1-\beta(1-\mu)]}\left(\frac{\int_{\xi}^{1}[g(\omega)-g(\xi)]d\omega - \frac{[1-\beta(1-\mu)]}{\psi_H(\xi,\Pi)}\tau}{\int_{\xi}^{1}[g(\omega)-g(\xi)]d\omega + \left(\frac{g(\xi)-\Gamma(\xi,\Pi)}{\psi_H(\xi,\Pi)}\right)}\right)$ while $\pi_m\left(\Pi|\xi\right)$ is unchanged. It is straightforward to show that $\partial\pi_l\left(\cdot|\xi\right)/\partial\tau < 0$ and thus $\partial\Pi_l\left(\xi\right)/\partial\tau < 0$ for any relevant ξ. Recall from the proof of Proposition 5 that $\Pi_m'\left(\xi\right) > 0$ for all $\xi \in \left(\underline{\xi}, \xi^*\right]$ where ξ^* is the equilibrium ξ under $\tau = 0$. Therefore, starting from $\left(\xi^*, \Pi^*\right)$, the downward shift in $\Pi_l\left(\cdot\right)$ caused by an increase in τ yields a decrease in ξ^* and Π^*. The new informed cutoff, call it ξ_τ^*, is lower than ξ^* so we have $\Pi_m'\left(\xi\right) > 0$ for all $\xi \in \left(\underline{\xi}, \xi_\tau^*\right]$. Starting from $\left(\xi_\tau^*, \Pi_\tau^*\right)$ then, another increase in τ shifts $\Pi_l\left(\cdot\right)$ further down and yields a decrease in ξ_τ^* and Π_τ^*. Continuing in this way establishes the result. ■

Appendix B - Algorithm for Computing Dynamics

Suppose a one-time aggregate productivity shock hits at $t = 1$. Recall that $z_1 < 0$ is realized after lending and production decisions have been made so all credit market variables are still in steady state at $t = 1$. The capital available for $t = 2$ is then:

$$K_2 = (1-\delta)(1+z)\left[\begin{array}{c} \int_{\xi_{ss}}^{1}\left(1-\sqrt{\frac{c}{\theta(\omega)}}\right)R(\omega)\lambda_{ss}(\omega)\,d\omega \\ +\int_{\overline{\eta}_{ss}}^{1}\left(1-\frac{c}{\theta(\omega)-\overline{R}_{ss}}\right)\overline{R}_{ss}\phi_{ss}(\omega)\,d\omega \end{array}\right] < K_{ss}$$

Although z returns to its expected value by $t = 2$, the effects of the $t = 1$ shock are propagated over time due to the change in capital. I start by computing the propagation in absence of the participation constraints. Let $T + 1$ denote the date at which ξ_t returns to ξ_{ss} and let $\overline{T}$ denote the date at which the entire economy returns to steady state. Note that $T+1 < \overline{T}$ since the partition of the type space implied by the evolution of ξ_t must stabilize before the distribution over that space can stabilize. The rest of the transition path is computed in four steps:

1. For $t = 2, \ldots, T$:
    - Guess Π_t.
    - Use Π_t, $\lambda_{t-1}(\cdot)$, and $\phi_{t-1}(\cdot)$ to get $\lambda_t(\cdot)$, $\phi_t(\cdot)$, $\overline{R}_t$, and K_{t+1}.
    - By bisection, find the ξ_t that equates $\tilde{K}_t$ (capital demand as defined in Section 1.3.1) to K_t.
2. For $t = T+1, \ldots, \overline{T}-1$:
    - Use $\lambda_{t-1}(\cdot)$, $\phi_{t-1}(\cdot)$, and $\xi_t = \xi_{ss}$ to get beliefs $\psi_t(\cdot)$.
    - Use $J_{t+1}(\xi_{ss}, \cdot) = \beta U_{t+2}(\cdot)$ to get an expression for r_t.

- Recursive substitution of $J_{t+1}(\omega, \cdot)$ into $J_t(\omega, \cdot)$ yields:

$$J_t(\omega, \cdot) = \frac{g(\max\{\omega, \xi_{ss}\}) - g(\xi_{ss})}{1 - \beta(1-\mu)} + \beta U_{t+1}(\cdot) \text{ for } t \in \left[T+1, \overline{T}-1\right)$$

- Use $\psi_t(\cdot)$, r_t, and the expression for $J_t(\omega, \cdot)$ to get $U_t(\cdot)$.
- Based on the first order condition for π_t, get the optimal π_t^*.

3. For $t = \overline{T} - 1, \ldots, T + 1$:

   - Recall that the value functions at date $\overline{T}$ are the steady state ones. Starting at $t = \overline{T} - 1$, use π_t^* as computed in step 2 to get $U_t(\cdot)$ and $J_t(\cdot)$.
   - Work back until $t = T + 1$.

4. For $t = T, \ldots, 2$

   - From step 3, we know the date $T + 1$ value functions. Starting at $t = T$, determine the optimal choice π_t^* then the value functions $U_t(\cdot)$ and $J_t(\cdot)$.
   - Work back until $t = 2$.

Symmetry requires $\Pi_t = \pi_t$ so compare the guess $\{\Pi_t\}_{t=2}^{T}$ with the result $\{\pi_t^*\}_{t=2}^{T}$. If the root mean squared error is not sufficiently small, then update the guess in the direction suggested by the result. Repeat until RMSE-convergence then verify that the unconstrained choice of $R(\cdot)$ does indeed satisfy the borrower participation constraint.