Equilibrium Spillover of Big Data

November 18, 2025

Preliminary; Comments welcome
For latest version click here

Abstract

We develop a parsimonious model of credit market competition where ex-ante identical lenders endogenously choose screening technologies and interest rates. Two key spillovers shape equilibrium outcomes: rejected borrowers reapply at other lenders, affecting application pool quality, and screening signals across lenders are correlated rather than conditionally independent. The equilibrium features a *hockey stick interest rate schedule*—a segmented market structure with varying degrees of fragmentation across borrower opacity levels. Lenders in different segments come to resemble traditional banks, fintech firms, or credit card issuers. We apply our framework to study how technological progress spills over across market segments. We show that whilst AI adoption increases financial inclusion, mandatory data sharing regulation counterintuitively may not benefit underserved populations and can increase inequality in financial access.

1 Introduction

The competitive landscape on today's' credit markets is rapidly changing and increasingly complex. The buyer of a house might obtain a mortgage from a traditional bank or a fin-tech lender. A small business owner might apply for a bank loan, credit from fin-tech firms, or for a high-interest rate credit card. Lenders might use different screening processes based on similar or shared data, using algorithms of various sophistication and charge different interest rates for a given borrower.

In this paper, we build a parsimonious model of credit market competition. We allow ex-ante identical lenders to choose both their screening technology and the interest rate they charge. We focus on two possible sources of spill-overs across lenders. First, we let rejected borrowers by some lenders to apply for credit at other lenders. This implies that lenders' choice of screening technology affects other lenders' application pool. Second, screening signals across lenders are correlated rather than conditionally independent. That is, given the same underlying data and similar algorithms, borrowers (mistakenly) rejected by some lenders' screening process, perhaps face rejection with higher probability by other lenders too. In equilibrium, lenders' individually optimal choices collectively determine the quality of loan applications, rejection rates, default rates and interest rates across the economy. A rich, heterogeneous structure of competitive landscape arises where ex-post some lenders are more reminiscent to traditional banks, while others to fin-tech firms or credit card issuers.

As a main application of our framework, we study how technological progress in some segments spills over to other segments. In particular, we study the impact of AI adoption and mandatory data sharing regulation on the credit market and show that while AI adoption leads to more financial inclusion, data sharing does not necessarily benefit the under-served population and counter-intuitively, can increase the inequality in financial access.

In our baseline model, borrowers type is two dimensional. Each is either good or bad which determines whether they eventually pay back or default. This is their creditworthiness. Additionally, they are heterogeneous in their opacity— determining how difficult it is for lenders to recognize them as creditworthy. Each is seeking to borrow at the lowest rate possible a quantity decreasing in the offered interest rate. Rejected borrowers at lower rates proceed to apply at higher rates. The mass of ex-ante identical lenders active in the economy proxies for the level of competition. Each lender can choose the precision of their screening technology for a cost. A higher precision technology is more costly, but can identify borrowers as good or bad up to a higher opacity threshold. Lenders advertise an interest rate recognizing the Type I and II errors their chosen screening technology implies given the quality of the pool of borrowers they serve in equilibrium.

Introducing borrowers' opacity as their non-pay-off relevant type creates the defining feature of our information structure: a specific form of correlated information across lenders. Under our assumptions, whether a lender's signal is correctly identifies the borrower's creditworthiness is fully determined by the lender's chosen precision relative to the opacity of the borrower. Hence, if two lenders choose the same precision, their signals on the same borrower will be perfectly correlated. They will be both (wrong) correct if their signal is (not) sufficiently precise compared to the borrower's opacity. The resulting information structure is *nested*. If a low precision lender correctly identifies the borrower's type as good or bad, a high precision lender certainly does so as well. Alternatively, if a high precision lender misidentifies the borrower's creditworthiness, a low precision lender will make the same mistake.[1]

In our baseline equilibrium, ex-ante homogeneous lenders choose heterogeneous levels of screening precision. Furthermore, the market structure is segmented with variable degrees of fragmentation across different level of borrower opacity.

In particular, there is a segment which resembles a traditional credit market where a homogeneous low interest rate is advertised, but not sufficiently transparent borrowers are rejected. It is so, because this market segment served by low-skilled lenders who cannot assess opaque borrowers. The second segment resembles a high-tech lending. Lenders present in this market make a large investment into their screening technology to be able to serve good borrowers whom are hard to assess. This is expensive, therefore lenders ask for a high interest rate which is increasing in borrowers opacity. The final segment features indiscriminate lending at the highest interest rate. Depending on the context, this region resembles a market of loan-sharks, of high-rate credit cards, or of low-documentation mortgages. Lenders do not invest in screening at all, instead ask for a very high interest rate as compensation for severe adverse selection.

Given the shape of interest rate schedule across these three segments as a function of good borrowers opacity, we refer to this structure as the *hockey stick interest rate schedule*. Note that the equilibrium matching between borrowers and lenders is non-assortative: the lowest skilled lenders serve the most opaque borrowers at the highest rate.

A crucial aspect of our model we discuss is the cross-dependence of market conditions across segments. Both high-tech lending and indiscriminate lending benefits from the presence of the traditional banking sector as the latter, by mistake, serves some of the hard-to-recognize bad borrowers which otherwise would end up in the borrower pool of other segments. This cross-dependence is crucial for the intuition of spill-overs we focus in the second part of the paper.

In the second part of the paper we introduce new entrants and study their effect on various segments of borrowers. That is, just right after incumbents choose their technology and become active, unexpectedly a new group of lenders arrive who can raise capital at potentially different cost and has a different degree of technological development compared to incumbents. At that point, incumbents are stuck with their technology choice. They can only adjust the interest rate they are lending at as a response. We use this set up to study the short-term effect of the impact of improvement in big data technologies as well as adoption of data sharing policies on credit markets.

First we show that the hockey stick interest rate schedule is robust to new entry. While its shape might change, the three segments remain. The robustness of the equilibrium market structure to new entry and exogenous changes in the economic environment is interesting, because despite being highly stylized, the hockey stick equilibrium structure of the model matches credit market outcomes in various contexts. For instance, it is inline the broad features of the small business lending market. Berger and Udell (2006) provide a detailed description of different instruments through which SMEs raise credit by reviewing the lit-

[1]In Section 5.2 we present a variant our model with the more standard conditionally iid signal structure to emphasize the critical role of nestedness in our insights.

erature, and argue that firms with higher financial transparency benefit from clearer risk assessments and lower costs of credit due to reduced information asymmetry. This leads to a convergence of interest rates for these firms, making their borrowing costs more uniform compared to less transparent firms. Strahan (1999) also documents that borrowers that are harder for outside investors to value pay more for their loans.

Alternatively, a similar pattern is documented in the mortgage market. For borrowers with strong financial profiles—e.g. available financial data and stable income, segmentation is less pronounced. These borrowers are generally offered the most competitive and uniform rates, especially when there is competition among lenders. While there may be some variation based on factors like the size of the down payment or loan term, the differences tend to be smaller for more transparent mortgage borrowers.

Our model with entry is especially suited to investigate the presence of spill-overs: That is, can borrowers who are served by the same incumbent lenders benefit or be harmed by entry in other segments? Using a serious of intuitive examples, we show that the answer is affirmative.

In particular, we use our model to study the consequences of growth in big data technologies and adoption of policies related to consumer data on the credit market equilibrium. We will interpret a reduction in the cost of screening precision as an improvement in data processing technology or improved access to consumer data. In particular, a "directed" change in the screening cost affects the interest rates borrowers in different segments are offered.

We focus on two particular exercises where spill-overs are present: Open Banking, and innovation in AI and Machine Learning.

We model Open Banking by endowing new entrants with reduced cost to screen relatively transparent borrowers, but not most opaque ones. We show that this might lead to positive or negative spill-over in the segment of the most opaque borrowers. The positive spill-over manifests when only the data of most transparent good borrowers is shared among the institutions. In other words, the cost for screening for new entrants reduces only for very low levels of skill. Alternatively, the financially excluded, most opaque borrowers are harmed by adoption of Open Banking when the data of a wider range of borrowers is shared among the institutions.

As AI enables lenders to screen the less-traditional borrowers better and less traditional borrowers are highly opaque, we interpret this innovation as new entrants' directed cost reduction to screen the most opaque borrowers but not the transparent ones. We show that this leads to a positive spill-over for the financially excluded. This is so, because high-tech lenders still some of business from indiscriminate lenders. As the latter group is stuck with their capital and technology, they end up reducing their rates to their remaining pool to be able to lend out all their capital.

Literature Review We contribute to the extensive theoretical literature that argues that adverse selection is an important in financial markets. Some of these models consider a market structure in which all trades must take place at one price (Eisfeldt, 2004; Daley and Green, 2012; Tirole, 2012; Chari et al., 2014). More generally, Gale (1992) provides a Walrasian theory of markets with one sided adverse selection with exclusive markets.

Adverse selection has also been introduced into models of random search (Lauermann and Wolinsky, 2016; Kaya and Kim, 2018; Lockwood, 1991). In these random search models, early selection dilutes the applicant pool at later firms. In our model, this ordering is exactly reversed due to the nestedness of information structure. As such, our model features creme skimming which is a prevalent feature of financial markets. Another strand of literature with one-sided adverse selection is competitive search. Different from us, these models deliver (almost) fully separating equilibria (Guerrieri et al., 2010; Guerrieri and Shimer, 2014; Chang, 2018). These models also assume market exclusivity, which allows signaling by sellers.

Most directly related to our work are models of adverse selection with two-sided heterogeneity (Kurlat, 2016; Li and Shimer, 2019; Board et al., 2017). We are similar to Kurlat (2016) in that the seller side of the market has a two-dimensional type in order to allow a nested information structure among the buyers. We allow non-exclusive markets as well. Li and Shimer (2019) and Board et al. (2017) both feature two-sided heterogeneity and thus share many features of our model, except the nestedness of the information structure. As such, the equilibrium market structures are different.

In particular, Board et al. (2017) are similar to us in that markets are non-exclusive, but they have an iid information structure and implies a fully separating equilibrium, while our equilibrium features both pooling and separating segments. Li and Shimer (2019) also features an iid information structure but with false positive error rates only. Furthermore, they have exclusive markets. Their equilibrium features pooling on the sellers' side and separation on buyers' side. This outcome is also different from our equilibrium structure.

Our paper builds on the previous work by the coauthors (Kurlat, 2016; Farboodi and Kondor, 2022). The information and market structure builds on Kurlat (2016), while the demand elasticity borrows from Farboodi and Kondor (2022). We generalize the theoretical contribution of both of these models in two dimensions. Both of these papers only consider the extreme cases where lenders make only type I or only type II errors. In this paper, we allow an arbitrary constant relative rate of type I versus type II error. As such, we illustrate that the two seemingly unrelated equilibria featured in both Kurlat (2016) and Farboodi and Kondor (2022) are in fact the two end points of the spectrum of a continuous set of equilibria.

Second, both of these papers take the wealth distribution of lenders as exogenous. One of our main contributions is to show that when ex-ante identical lenders choose their type, a unique heterogeneous wealth distribution emerges endogenously. This further enables us to consider entry of new lenders in our framework.

One of the applications of the model that we consider is the impact of adoption of Open Banking regulation in the credit market. We contribute to a small but growing literature that considers this question, including but not limited to Goldstein et al. (2022), He et al. (2023), and Babina et al. (2025).

The rest of the paper is organized as follows. Section 2 presents our baseline model, provides a construction for the equilibrium and highlights its main properties. Section 3 introduces new entrants and provides a characterization of their effect on the equilibrium. Section 4 investigates the impact of big data technological growth and related policies on the credit market equilibrium. Section 5.2 provides a benchmark with non-nested information structure. Section 6 concludes.

2 Baseline Economy

We model a credit market with lenders with heterogeneous skill to learn their borrowers' type. In our baseline economy presented in this section, lenders are either endowed with this skill, or obtain it for a cost. In enter case, they simultaneously choose the interest rate at which they are willing to lend at.

In section 3, we will introduce an additional group of lenders who can enter and choose their skill at a different cost taking the skill distribution of incumbents as given. This will make it possible to study how the equilibrium is affected by new entrants with potentially different technology.

2.1 Set-Up

There are two dates, $t = 1, 2$ and there is no discounting. There are two types of agents, lenders and borrowers. Borrowers borrow an endogenous quantity at an endogenous interest rate from lenders at $t = 1$ promising to pay back in period $t = 2$.

There is a continuum of heterogeneous borrowers. Each borrower has a two-dimensional type, (τ, ω). The first dimension, $\tau \in \{G, B\}$ controls borrower performance vis-a-vis the lender. A good borrower ($\tau = G$) pays back fully, while a bad borrower ($\tau = B$) defaults. The second dimension $\omega \in [0, 1]$, is the opacity of the borrower which, as we will specify shortly, refers to the difficulty to be recognized by a lender as a type τ. The dimension of opacity is not pay-off relevant for the lender. The continuum of borrowers is distributed with measure/density $G(\omega)/g(\omega)$ and $B(\omega)/b(\omega)$ on $\omega \in [0, 1]$. We represent borrowers preferences by the following reduced form assumption.

Assumption 1. *Each borrower wishes to borrow at the lowest interest rate possible. If the lowest rate at which she can obtain any credit is r, she demands $D(r)$ units where $D(r)$ is a strictly decreasing function.*

Observe that each borrower's demand function is identical, independently of her type, τ. We make this assumption to focus the reader's attention on the lender's side of the market where the engine of our mechanism is. The most straightforward interpretation is that the borrowers do not know their own type. We follow this interpretation in the main text. In contrast, Appendix B presents a micro-foundation where borrowers' know their type but a collateral constraint determines the same borrowing limit for each type.

There is a mass $\bar{W}$ of ex-ante identical lenders. Each lender is endowed with a common *basic screening technology* which we parameterize with $\beta \in [0, 1]$. Additionally, for a cost $C(\alpha)$ each can add precision $\alpha \in [0, 1]$ to her screening technology at the beginning of $t = 1$ where $C(\alpha)$ is strictly increasing, continuous with $C(0) = 0$. We will refer to lenders choosing a higher precision as more skilled.

Each lender, given their precision, α chooses an interest rate, r at which they wish to lend. Each interest rate in their choice set $r \in [0, \infty]$ defines a *market*. The lender who chooses the given interest rate is active on that market. Borrowers can apply for loans in any subset of markets. A lender active in a given market observes a signal of each borrower who applied for loans at that market. This signal depends on the precision of the lender, α, the type of the borrower τ, ω and on the common basic screening technology β as follows:

Definition 1 (**Nested Information structure).** *Conditional whether a borrower with opacity ω is good, $\tau = G$, or bad, $\tau = B$, a lender with a screening technology β and precision α will observe a signal*

$$s(\tau = G, \omega, \alpha; \beta) = \begin{cases} g & \text{if } \omega < \omega_g(\alpha) \\ b & \text{otherwise} \end{cases} \tag{1}$$

or

$$s(\tau = B, \omega, \alpha; \beta) = \begin{cases} b & \text{if } \omega < \omega_b(\alpha) \\ g & \text{otherwise} \end{cases} \tag{2}$$

on that borrower, respectively, where $\omega_g(\alpha) \equiv \beta + \alpha(1-\beta)$ is the most opaque good borrower a lender with precision α recognizes as good, while $\omega_b(\alpha) \equiv (1-\beta) + \alpha\beta$ is the most opaque bad borrower a lender with precision α recognizes as bad.

To understand the implied information structure consider first the case with $\beta = 1$. In that case, a lender with precision α gets a signal g both on each good applicant and on those bad applicants which are sufficiently opaque compared to the lender's precision ($\omega > \alpha$). At the same time, she will get a signal b only on (sufficiently transparent) bad borrowers. That is, the lender makes only false positive mistakes. In contrast, if $\beta = 0$, a lender with precision α gets a signal b on each bad applicant and those good applicants which are sufficiently opaque compared to her precision ($\omega > \alpha$). That is, the lender will make only false negative mistakes. In general, under an interior basic screening technology, β, the lender makes both false positive and false negative mistakes on opaque borrowers. In fact, our parametrization implies that the fraction of false positive to false negative mistakes is driven only by β:

$$\frac{\text{type I error rate}}{\text{type II error rate}} = \frac{\beta}{1-\beta}.$$

In Farboodi and Kondor (2023) we connect (the extreme values of) parameter β with aggregate business cycle conditions. In good times lenders tend to follow more lax lending standards corresponding to more false positive and less false negative mistakes, a high β, while lending standards tend to be tighter in bad times. In this paper, we keep β fixed. Our results require only that it is interior.

Importantly, lenders make correlated mistakes. Note that the signal a lender observes on a particular borrower in her application pool depends solely on the lender's chosen precision and the borrower's two dimensional type. That is, two lenders choosing the same precision necessarily identical signals on the same borrower. Now consider two lenders with $\alpha' < \alpha''$. All the bad borrowers for whom the more skilled lender would receives a signal g, the less skilled lender will also receive a signal g, and symmetrically for Type II errors. At the same time, there are always a set of good borrowers which only the more skilled lenders can identify as good, and likewise for bad borrowers. We call this property *nestedness* and it plays a crucial role in our analysis. In section 5.2 we illustrate the force of this assumption by solving our model with the non-nested, conditionally iid version of our information structure.

Note that opacity is not an observable characteristic on which the lender can condition her decision. By definition, it purely a characterizes the required precision lenders need to asses the creditworthiness of that borrower. In fact, opaque good borrowers and opaque bad borrowers do not need to look alike. Intuitively, an opaque bad borrower might be able to provide rich documentation which for a low skilled lender looks immaculate. This is why she makes the false positive mistake. At the same time, an opaque good borrower might have irregular documentation and this is why low skilled lenders mistakenly take them as bad.

2.2 Incumbent Equilibrium

In our baseline economy, equilibrium works as follows. First, borrowers simultaneously choose their precision α. Then the lending markets open. Each possible interest rate r defines a different market. Borrowers submit applications to various markets sequentially, starting from the lowest interest rate. If their application is accepted in market r, they borrow $D(r)$ and exit; if it is rejected, they continue to apply to higher interest rates. Lenders who choose to lend in market r have to decide whether to be selective, in which case they only lend to applicants for whom they observe $s = g$, or non-selective, in which case they lend to anyone. In either case, they lend one unit of capital to a randomly selected acceptable applicant.

If many lenders lend in the same market, the pool of applicants each faces depends on the order in which they lend, since borrowers who have already been served exit the pool. We will assume that lenders are queued in order of increasing α, so those with lower precision go first (and non-selective lenders before everyone else). We later show that, perhaps surprisingly, all lenders prefer this ordering, so that if we generalized our definition of equilibrium to encompass an endogenous ordering, as in Kurlat (2016), this is the ordering that would emerge.

We use the following notation. The functions $r(\alpha)$ and $z(\alpha)$ denote, respectively, the choice of market and selectivity by a lender with precision α, with $z(\alpha) = 1$ representing the decision to be selective. The function $\gamma(r, z, \alpha)$ denotes the probability that a borrower faced by a lender with precision α and selectivity z in market r is a good borrower. The measures $G(\cdot; r, z, \alpha)$ and $B(\cdot; r, z, \alpha)$ (defined over the space of opacity $\omega \in [0, 1]$) denote how many good and bad borrowers respectively of each opacity are in the pool of applicants in market r by the time it's the turn of lender α with selectivity z. The measure $W(\cdot)$ (defined over the space of precision $\alpha \in [0, 1]$) denotes how many lenders choose each precision.

The problem of a lender can be divided into two parts. Conditional on a given precision α, the lender must choose a market r and selectivity z to solve:

$$\tilde{\Pi}(\alpha) = \max_{r,z} \gamma(r, z; \alpha)(1 + r) - 1 \tag{3}$$

The lender lends out 1 and, with probability $\gamma(r, z; \alpha)$, gets $1 + r$ in return, so the expected gross profit is $\tilde{\Pi}(\alpha)$. Then, here is also a choice-of-precision problem to maximize net profit, that is, gross profit minus the cost of precision:

$$\max_{\alpha} \tilde{\Pi}(\alpha) - C(\alpha). \tag{4}$$

We expect, and later verify, that any precision α chosen in equilibrium by a positive measure of lenders must imply the same profit Π which is decreasing in the total mass of lenders, $\bar{W}$. In this sense, $\bar{W}$ is a measure of the intensity of lender's competition in our economy.

The quality $\gamma(r, z; \alpha)$ faced by the lender can be computed as follows. Define

$$I^G(\alpha, z) = \begin{cases} [0, \omega_g(\alpha)] & \text{if } z = 1 \\ [0,1] & \text{if } z = 0 \end{cases}$$

$$I^B(\alpha, z) = \begin{cases} [\omega_b(\alpha), 1] & \text{if } z = 1 \\ [0,1] & \text{if } z = 0 \end{cases}$$

I^G and I^B represent, respectively, the subsets of good and bad assets that the lender accepts, depending on their information α and their selectivity z. Let Ω^G and Ω^B be any subsets of good and bad borrowers. If lender α is chooses z in market r, the probability of getting a borrower who belongs in one of these subsets is, respectively:

$$\Pr_G\left(\Omega^G; r, z, \alpha\right) = \frac{G\left(\Omega^G \cap I^G(\alpha, z); r, z, \alpha\right)}{G\left(I^G(\alpha, z); r, z, \alpha\right) + B\left(I^B(\alpha, z); r, z, \alpha\right)} \tag{5}$$

$$\Pr_B\left(\Omega^B; r, z, \alpha\right) = \frac{B\left(\Omega^B \cap I^B(\alpha, z); r, z, \alpha\right)}{G\left(I^G(\alpha, z); r, z, \alpha\right) + B\left(I^B(\alpha, z); r, z, \alpha\right)} \tag{6}$$

The denominators in (5) and (6) are the measure of all borrowers that are acceptable to lender α, and the numerators are the measures in subsets Ω^G and Ω^B respectively. Using (5) and (6), the probability that a lender α with selectivity z in market r gets a good borrower is:

$$\gamma(r, z; \alpha) = \frac{\Pr_G([0,1]; r, z, \alpha)}{\Pr_G([0,1]; r, z, \alpha) + \Pr_B([0,1]; r, z, \alpha)} \tag{7}$$

when the denominator is positive, and zero otherwise. The numerator in (7) is the total measure of good borrowers that are acceptable to an α lender with selectivity z in market r, while the denominator is sum of the total measure of acceptable good and bad borrowers.

It remains to compute the measures $G(\cdot; r, z, \alpha)$ and $B(\cdot; r, z, \alpha)$. For this we need to subtract from the original pool of borrowers those who have been served in lower-r markets or in market r by lower-α or non-selective lenders. Let

$$A(r, z, \alpha) = \{\tilde{\alpha} : r(\tilde{\alpha}) < r\} \cup \{\tilde{\alpha} : r(\tilde{\alpha}) = r, z(\tilde{\alpha})\tilde{\alpha} < z(\alpha)\alpha\} \tag{8}$$

be the set of lenders that choose a lower-interest-rate market than r or choose r but pick before α. Each of these lenders lends to $1/D(r(\alpha))$ borrowers in the market they visit, distributed across opacity levels according to (5) and (6). Hence, the distributions faced by lender α with selectivity z in market r are:

$$G\left(\Omega^G; r, z, \alpha\right) = G\left(\Omega^G\right) - \int_{A(r,z,\alpha)} \Pr{}_G\left(\Omega^G; r(\alpha), z(\alpha), \alpha\right) \frac{1}{D(r(\alpha))} dW(\alpha) \tag{9}$$

and

$$B\left(\Omega^B; r, z, \alpha\right) = B\left(\Omega^B\right) - \int_{A(r,z,\alpha)} \Pr{}_B\left(\Omega^B; r(\alpha), z(\alpha), \alpha\right) \frac{1}{D(r(\alpha))} dW(\alpha) \tag{10}$$

We can now formally define an equilibrium:

Definition 2 (Incumbent Equilibrium). *The equilibrium consists of*

1. *A measure W over lender screening precision such that $W([0,1]) = \bar{W}$,*

2. *choice-of-market function $r(\alpha)$ and a (binary) choice-of selectiveness function $z(\alpha)$ for each lender α in the support of W,*

3. *measures of good and bad borrowers available to lender α with selectivity z in market r: $G(\cdot; r, z, \alpha)$ and $B(\cdot; r, z, \alpha)$*

such that

1. *Given α, $r(\alpha)$ and $z(\alpha)$ solve the lender's problem* (3)*, with γ defined by* (5), (6) *and* (7)*,*

2. *Every α in the support of W solves* (4)*,*

3. *The measures $G(\cdot; r, z, \alpha)$ and $B(\cdot; r, z, \alpha)$ satisfy* (9) *and* (10) *respectively.*

For comparison, we will also look at what happens for an exogenous distribution of α in Section 5.1. In that case, the definition of equilibrium is the same, except that we take W as given and do not require every α in the support of W to solve problem (4).

We sometimes refer to the equilibrium where lenders enter in period 1 only as the Incumbent Equilibrium. In section 3 we introduce new entrants and define the Entry Equilibrium to compare.

2.3 Equilibrium Construction

In this section, we construct the equilibrium in our economy. Our strategy is to provide the formal argument for the steps of construction in this section and highlight the economic intuition and properties in Section 2.4.

The next proposition states the main properties of the equilibrium. We emphasize that the emerging market structure is segmented characterized by a "hockey stick interest rate schedule"[2].

Proposition 1 (Incumbent Equilibrium: The hockey stick interest rate schedule). *There is a exist a unique equilibrium defined by the endogenous thresholds on lender screening precision: α_0, α_1, α_2 satisfying $0 \leq \alpha_0 \leq \alpha_1 \leq \alpha_2 \leq 1$ and the endogenous lenders density $w(\alpha)$, corresponding to measure $W(\cdot)$, which has an atom $w(0) = w^{NS}$ at 0 and a continuous support in the range $[\alpha_0, \alpha_2]$.*

The equilibrium interest rate schedule consists of (at most) three segments, ordered by increasing interest rates:

1. **Region I:** *A low interest rate r_p where both good and bad borrowers borrow.*

   Every easy-to-recognize good borrower with $\omega \leq \omega_g(\alpha_1)$, every hard-to-recognize bad borrower, (b, ω) with $\omega > \omega_b(\alpha_1)$ borrow at r_p.

[2]Under a loose interpretation of what a hockey stick looks like.

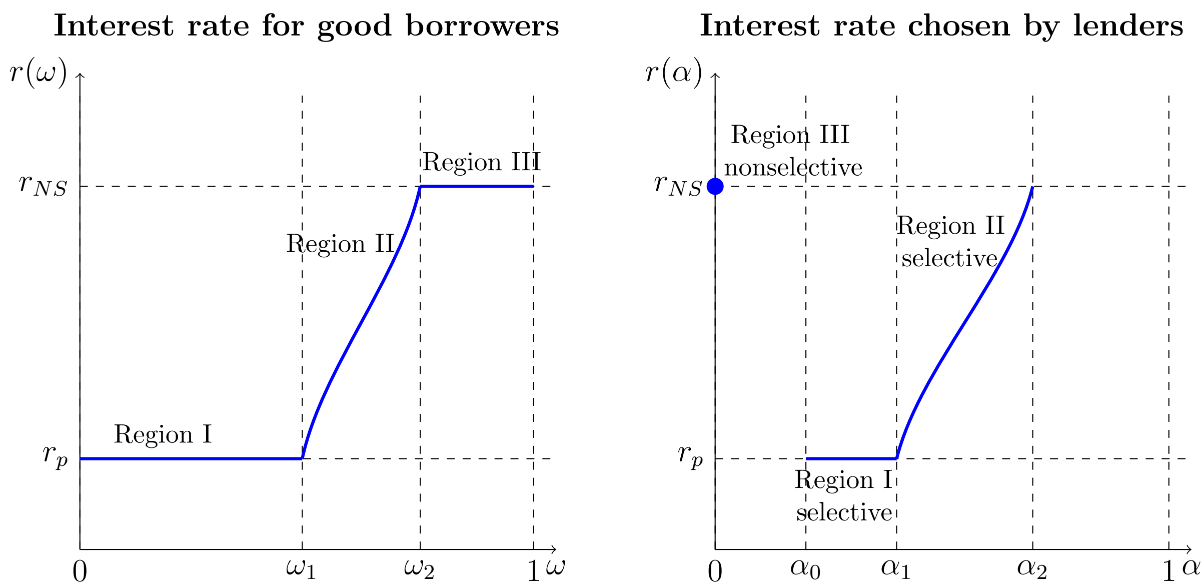

Figure 1: A hockey-stick interest rate schedule

Lenders with intermediate degrees of technological precision $\alpha \in [\alpha_0, \alpha_1]$ *lend at* r_p. *All lenders in this market are selective.*

2. **Region II:** *An increasing interest rate schedule for good borrowers only.*

   Every moderately-easy-to-recognize good borrower, with $\omega \in [\omega_g(\alpha_1), \omega_b(\alpha_2)]$ *borrows at a single interest rate* $r(\omega)$ *within this range. No bad borrower borrows in this range.*

   Lenders with a high degree of technological precision, $\alpha \in [\alpha_1, \alpha_2]$, *lend in this segment. All lenders in these markets are selective.*

3. **Region III:** *A high interest rate* r_{NS} *where both good and bad borrowers borrow.*

   All hard-to-recognize good borrowers, (g, ω) *with* $\omega > \omega_g(\alpha_2)$ *and all easy-to-recognize bad borrowers,* (b, ω) *with* $\omega < \omega_b(\alpha_2)$ *borrow at* r_{NS}.

   Lenders with the lowest technology level, $\alpha = 0$ *who are non-selective lend at* r_{NS}.

The interest rate schedule is continuous in ω *for good borrowers that is* $r(\omega_g(\alpha_1)) = r_p$ *and* $r(\omega_g(\alpha_2)) = r_{NS}$.

Figure 1 shows and example of a hockey-stick schedule. The left panel shows the interest rate at which good borrowers obtain credit, as a function of their opacity ω, which has the hockey-stick shape. The right panel shows the interest rate chosen by lenders, as a function of their precision α. In Section 2.4, we expand on the economic interpretation behind this pattern.

We construct the equilibrium as follows. We conjecture a common profit level Π. Then we follow the logic of our market clearing mechanism matching borrowers to lenders. We start at the market with the lowest offered rate, r_p. All borrowers apply to that market. We identify the lenders choosing the lowest precision α_0 offering that rate. The borrowers accepted by these lenders exit the pool of applicants, changing the density of borrowers the rest of the lenders face. We work out how these densities change as we proceed to lenders offering r_p but choosing higher precision α. This is Region I. Then we proceed to markets with higher interest rates to Region II, and then to Region III. Finally, we calculate the total mass of lenders needed to support this equilibrium. If its larger (smaller) than $\bar{W}$, we adjust the conjectured profit level Π downward (upward) until finding the fixed point. We explain why none of the lenders or borrowers would want to deviate. In the Appendix, we also prove that this construction always leads to the unique equilibrium.

Region I Let us start from a conjectured profit level Π. It will be convenient to define the adjusted cost function

$$K(\alpha) = \Pi + C(\alpha).$$

This is lender's cost of entry with precision α adjusted with their required level of profit, Π. Conjecture a value for r_p: the lowest interest rate that is offered by any lender. Since r_p is the lowest interest rate available, it must attract all the borrowers. Therefore the lowest-α lender who is active in market r_p obtains an average quality of:

$$\gamma_0(\alpha) = \frac{G(\omega_g(\alpha))}{G(\omega_g(\alpha)) + B(1) - B(\omega_b(\alpha))} \tag{11}$$

Find the first lender α_0 as the lender who can charge the lowest interest rate and still make profit Π:

$$\alpha_0 = \arg\min_{\alpha} \frac{K(\alpha)+1}{\gamma_0(\alpha)} - 1. \tag{12}$$

The right hand side of (A.1) is the interest rate that lender α needs to charge in order to make profits Π, if they are first in line. From this, we find the pooling interest rate simply as:

$$r_p = \frac{K(\alpha_0)+1}{\gamma_0(\alpha_0)} - 1 \tag{13}$$

Proposition 1 states that Region I is characterized by a density $w(\alpha)$ describing the measure of lenders offering the same r_p but choosing different precisions in a range $\alpha \in [\alpha_0, \alpha_1]$. To see how that can work, first we describe the evolution of the applicant pool under this scenario for an arbitrary $w(\alpha)$. Denote by $g(\omega; r_p, 1, \alpha)$, $b(\omega; r_p, 1, \alpha)$ the densities of good and bad borrowers in the applicant pool of a selective lender with precision α offering r_p interest rate. These are the densities corresponding to the measures $G(\cdot; r_p, 1, \alpha)$, $B(\cdot; r_p, 1, \alpha)$ in Definition 4. Then, for each lender with precision α the mass of acceptable borrowers is

$$T(r_p, 1, \alpha) \equiv \int_0^{\omega_g(\alpha)} g(\omega; r_p, 1, \alpha)\, d\omega + \int_0^{\omega_b(\alpha)} b(\omega; r_p, 1, \alpha)\, d\omega.$$

When lender α lends, he serves $\frac{w(\alpha)}{D(r_p)}$ borrowers, pro-rated among the $T(r_p, 1, \alpha)$ acceptable ones. Therefore, for every ω that lender α finds acceptable, the number of borrowers who remain unserved goes down by a fraction equal to:

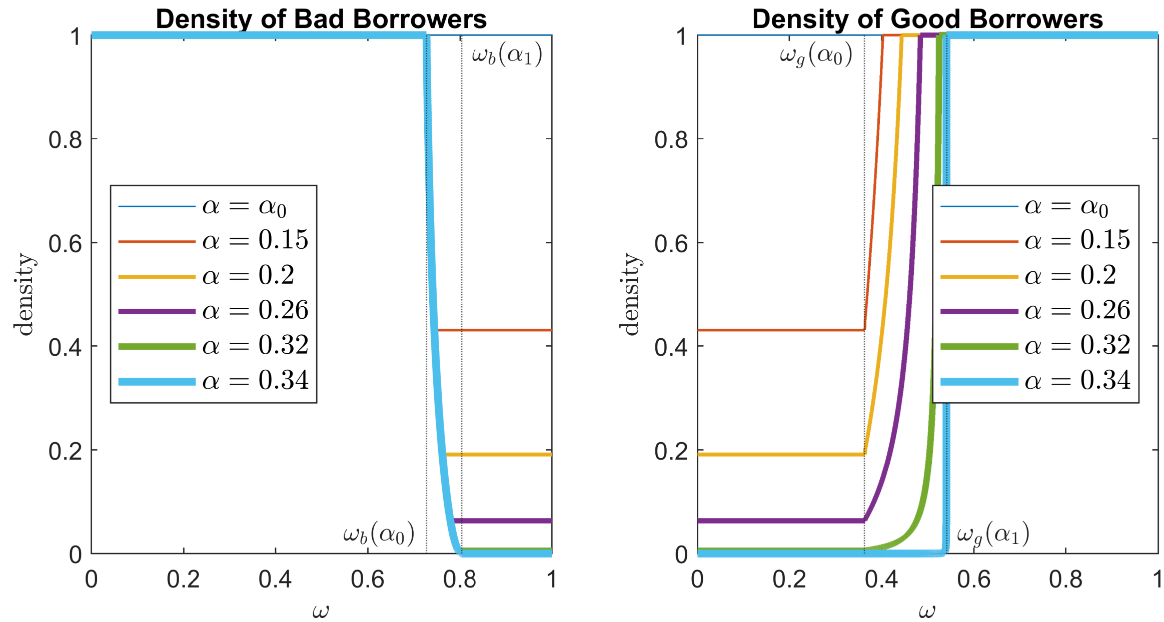

Figure 2: How the remaining density of acceptable borrowers changes after the turn of lender α?

$$\theta(r_p, 1, \alpha) \equiv \frac{w(\alpha)}{D(r_p)T(r_p, 1, \alpha)} \tag{14}$$

Therefore, using (9)-(10), the applicant pools of good and bad borrowers for lender α must satisfy the following differential equations:

$$\frac{\partial g(\omega; r_p, 1, \alpha)}{\partial \alpha} = -\theta(r_p, 1, \alpha)\mathbb{I}(\omega \leq \omega_g(\alpha))\, g(\omega; r_p, 1, \alpha) \tag{15}$$

$$\frac{\partial b(\omega; r_p, 1, \alpha)}{\partial \alpha} = -\theta(r_p, 1, \alpha)\mathbb{I}(\omega \geq \omega_b(\alpha))\, b(\omega; r_p, 1, \alpha) \tag{16}$$

Equations (A.1)-(13), together with (A.24) and (A.25) and initial condition $g(\omega; r_p, 1, \alpha) = g(\omega)$ and $b(\omega; r_p, 1, \alpha) = b(\omega)$ fully define the functions $g(\omega; r_p, 1, \alpha)$ and $b(\omega; r_p, 1, \alpha)$.

To see the intuition, Figure (2) shows how the density of borrowers in a lender's application pool evolves as the queue of lenders advances from $\alpha = \alpha_0$ towards lenders with higher precision. Thicker curves correspond to the density of remaining borrowers available for higher precision lenders.

The left panel shows densities of bad borrowers. The thinnest (light blue) horizontal line at 1 shows that the lender with the lowest precision, α_0 is facing the full distribution of borrowers, which,in this example is assumed to be uniform, $g(\omega) = b(\omega) = 1$ for all ω. This lender accepts the most opaque borrowers with $\omega \geq \omega_b(\alpha_0)$ by mistake and takes a slice proportional to their capital $w(\alpha_0)$. Each group with progressively higher precision $\alpha > \alpha_0$

makes less mistakes, accepting bad borrowers with $\omega \geq \omega_b(\alpha) > \omega_b(\alpha_0)$ which explaining the pattern on the panel. Note that as we proceed to group of lenders with higher precision, there must be an α at which all the good borrowers with $\omega \leq \omega_g(\alpha)$ are depleted. Let us call this precision α_1.

On the right panel are the good borrowers. The lender with the lowest precision α_0 accepts the most transparent good borrowers $\omega \leq \omega_g(\alpha_0)$ and takes a slice of them proportional to $w(\alpha_0)$. The next lender with slightly higher precision, $\alpha = \alpha_0 + d\alpha$ takes a slice of the remaining good borrowers with $\omega \leq \omega_g(\alpha_0 + d\alpha)$. As $\omega_g(\alpha_0 + d\alpha) > \omega_g(\alpha_0)$, this lender accept all good borrowers the previous one accepted depleting the pool with $\omega \leq \omega_g(\alpha_0)$ further. However, he also accept the good borrowers with $\omega \in [\omega_g(\alpha_0), \omega_g(\alpha_0 + d\alpha)]$. We will refer to this group of good borrowers as the own slice of lender $\alpha_0 + d\alpha$ as he just have sufficient precision to recognize them as good. By the point we reach lender $\alpha = 0.15$, the density of available transparent borrowers are severely depleted. The flat part extends to $\omega_g(\alpha_0)$. These are the borrowers all previous lenders accepted. In the range $\omega \in [\omega_g(\alpha_0), \omega_g(\alpha = 0.15)]$ there are increasingly more remaining borrowers as higher precision is needed to accept them. Note that as we proceed to lenders with higher precision, at some choice of α all the good borrowers with $\omega \leq \omega_g(\alpha)$ are depleted. In fact, as we show in the Appendix, it happens at α_1.

Now we turn how to determine lenders' information choice and entry in equilibrium, that is, the equilibrium density $w(\alpha)$. Recall first that as lenders are ex-ante identical, all active lenders has to make the same net profit Π. This fact gives a convenient way to determine α_1, the end of Region I, by

$$r_p = K(\alpha_1) \tag{17}$$

even without solving the differential equation system (A.24)-(A.25). The reason is that for lender $\alpha_1 + d\alpha$ all the bad borrowers he would take by mistake, $\omega \geq \omega_b(\alpha_1)$ are already fully served by definition. As he is always left with his own slice of good borrowers $\omega \in [\omega_g(\alpha_1), \omega_g(\alpha_1 + d\alpha)]$, he will serve good borrowers only. So (17) must hold.[3]

Finally, the construction of Region I is completed by determining the density $w(\alpha)$ over the support $\alpha \in [\alpha_0, \alpha_1]$ which ensures that each active lenders offering interest rate r_p makes the same profit. That is, for lenders in this range,

$$(1 + r_p)\,\gamma(r_p, 1, \alpha) - 1 = K(\alpha) \tag{18}$$

where

$$\gamma(r_p, 1, \alpha) = \frac{\int_0^{\omega_g(\alpha)} g(\omega; r_p, 1, \alpha)\, d\omega}{T(r_p, 1, \alpha)} \tag{19}$$

is the fraction of good borrowers in the accepted borrowers of lender α. Note that equal profit condition (18) immediately implies that the quality of a lender's α loan portfolio, $\gamma(r_p, 1, \alpha)$ has to be monotonically increasing, starting from $\gamma(r_p, 1, \alpha_0) = \gamma_0(\alpha_0)$ and reaching 1 at $\gamma(r_p, 1, \alpha_1) = 1$. In particular, lenders have to enter with the mass $w(\alpha)$ which, through (14) and (A.24)-(A.25), increases the portfolio quality $\gamma(r_p, 1, \alpha + d\alpha)$ for the next lender

[3] If $r_p > K(\alpha_1)$ for all α, $\alpha_1 = 1$ and region II and III does not exists.

choosing $\alpha + d\alpha$ just enough to compensate for the higher cost of precision $C(\alpha + d\alpha)$. We present the formal construction of such $w(\alpha)$ (by discretizing the space of α and then taking the continuous limit) in the Appendix.

Before we move on to Region II, let us emphasize a key properties of our structure. The construction of the pooling region is premised on the fact that lenders in the same market are ordered by increasing α, a property that is built into equation (8) in our definition of equilibrium. We now verify that, if lenders could choose in what order to lend, as in the definition of equilibrium from Kurlat (2016), they would all choose this ordering. Suppose a given lender α lends in a market where other lenders are also active. Starting from any ordering, that lender is given the option of moving further back in the queue, letting other lenders with types in some set A in front of him . Let γ be the average quality lender α gets with his original position and γ' the average quality he gets if he moves back.

Lemma 1 (Endogenous Ordering). *If $\tilde{\alpha} < \alpha$ for all $\tilde{\alpha} \in A$, then $\gamma' > \gamma$. Conversely, if $\tilde{\alpha} > \alpha$ for all $\tilde{\alpha} \in A$, then $\gamma' < \gamma$.*

Lemma 1 says that any lender α prefers to be after less-skilled lenders and before more-skilled lenders. The fact that they prefer to come before more-skilled lenders is standard: more skilled lenders pick out good borrowers, leaving behind an adversely selected pool. What is perhaps more surprising is the lenders are happy to come after less-skilled colleagues. After all, less-skilled does not mean completely unskilled. Would they not also leave a somewhat adversely selected sample? The reason this is not undesirable has to do with the way the information is nested. A less-skilled lender lends to a subset of the good borrowers and a superset of the bad borrowers that are acceptable to a more skilled lender. Therefore, conditional on being acceptable to the more-skilled lender, the pool they leave behind is positively selected.

In fact, this property is also crucial for the endogenous determination of the distribution of lenders $w(\alpha)$. The equilibrium construction relies on the fact that more lenders choose a particular precision α', the more it improves selection for all lenders choosing a higher skill $\alpha'' > \alpha'$. Intuitively, less skilled investors cleanse the pool for more skilled investors as less skilled are serving some of the hardest-to-recognize bad borrowers. Hence, it is possible to find the $w(\alpha)$ at a given skill α in a way that the next group $\alpha + d\alpha$ is just willing to pay the higher cost for the corresponding improved loan quality $\gamma(r_p, 1, \alpha + d\alpha)$.

Region II In Region II we find $r(\alpha)$ by the indifference condition

$$r(\alpha) = K(\alpha). \tag{20}$$

Since in Region II lenders only lend to good borrowers, the interest rate has to be exactly enough to compensate for information costs. Then we find the density $w(\alpha)$ by condition

$$w(\alpha) = D(r(\alpha))\, g(\omega_g(\alpha))\,(1-\beta) \tag{21}$$

which equates supply and demand for each interest rate in this region.

Region III Finally, we find Region III, if it exists. This happens when for some α the implied interest rate $r(\alpha)$ by (20) is relatively high compared to the quality of remaining borrowers. Then, it might be the case that a lender can obtain profit Π by not investing in precision at all, instead lending non-selectively to all remaining borrowers at the given interest rate $r(\alpha)$.

To be more precise, the average quality that a non-selective lender would get in market $r(\alpha)$ is:

$$\gamma^{NS}(\alpha;\alpha_0,\alpha_1) = \frac{G(1) - G(\omega_g(\alpha))}{G(1) - G(\omega_g(\alpha)) + L(\alpha_0,\alpha_1)} \tag{22}$$

Here the quantity $G(1) - G(\omega_g(\alpha))$ is the total mass of good borrowers who have not been served yet because they are too opaque for the lenders with lower skill than α. While

$$L(\alpha_0,\alpha_1) \equiv \int_0^1 b(\omega; r_p, 1, \alpha_1)\, d\omega$$

is the total mass of bad borrowers who were not served by lenders $\alpha \in [\alpha_0, \alpha_1]$ in Region I at interest rate r_p. We will refer to $L(\alpha_0,\alpha_1)$ as the leftover bad borrowers.

For each $r \in [r_p, r(1)]$, compute:

$$\Pi^{NS}(r) = \gamma^{NS}(r;\alpha_0,\alpha_1)(1+r) - 1 \tag{23}$$

where $\gamma^{NS}(r)$ is given by evaluating (A.28) at $\alpha = K^{-1}(r)$. Let r_{NS} be defined by the minimum value of r within the interval $[0, r(1)]$ such that $\Pi^{NS}(r) \geq \Pi$, if such a value exists.[4]. Lenders who enter with $\alpha = 0$ choose $r(0) = r_{NS}$ and $z(0) = 0$, that is they lend non-selectively in market r_{NS}. Let $\alpha_2 \equiv K^{-1}(r_{NS})$ define the the boundary between Region II and Region III. If r_{NS} exists, then W has a mass point at $\alpha = 0$, with mass

$$w^{NS} \equiv \frac{1}{D(r_{NS})}\left[G([0,1]; r_{NS}, 0) + B([0,1]; r_{NS}, 0)\right].$$

That is, the non-selective entry w^{NS} is sufficient to clear all the remaining demand of both bad and good borrowers. This also implies no entry from lenders with precision $\alpha > \alpha_2$ that would require a $r > r_{NS}$ in order to earn Π.

The construction above, for any conjectured Π, fully defines the measure W of lenders who enter at each precision α. Hence, the condition $\bar{W} = W([0,1])$ defines a fixed point problem. In the Appendix, we prove that the right hand side is monotonically decreasing in Π in a range $\Pi \in [0, \Pi^{max}]$, hence the fixed point problem always have a solution in this range.

The equilibrium is therefore:

1. The measure W defined by the construction above

2. Choice of markets and selectiveness

$$r(\alpha) = \begin{cases} r_{NS} & \text{if } \alpha = 0 \\ r_p & \text{if } \alpha \in [\alpha_0, \alpha_1] \\ K(\alpha) & \text{if } \alpha > \alpha_1 \end{cases}$$

[4]This includes as a special case $r_{NS} \leq r_p$, in which case the only market is non-selective

$$z(\alpha) = \begin{cases} 0 & \text{if } \alpha = 0 \\ 1 & \text{otherwise} \end{cases}$$

3. Measures G and B constructed as in the definition of equilibrium implying the continuous loan portfolio quality function for selective lenders

$$\gamma(\alpha) = \begin{cases} \gamma_0(\alpha) & if \quad \alpha \in [0, \alpha_0] \\ \gamma(r_p, 1, \alpha) & if \quad \alpha \in [\alpha_0, \alpha_1] \\ 1 & if \quad \alpha \in [\alpha_1, 1] \end{cases}. \tag{24}$$

and $\gamma^{NS}(\alpha_2, \alpha_0, \alpha_1)$ for non-selective lenders in Region III.

2.4 Equilibrium Properties

In the previous section, we characterized the equilibrium when lenders are ex-ante identical but can choose their skill for a cost implying a nested information structure. We have shown that the resulting equilibrium features a hockey stick pattern as depicted in Figure 1. In this part, we discuss the economic intuition behind the properties of the different segments and how these segments support each other.

As Figure 1 shows, a group of low opacity good borrowers, that is, the ones which can be recognized as good at low cost, can borrow at a uniformly low rate, r_p. This is Region I. Because these good borrowers are served by low precision lenders with $\alpha \in [\alpha_0, \alpha_1]$, the most opaque bad borrowers will be served at this segment too. These are the borrowers whom low precision lenders cannot distinguish from the good ones. Hence,this market is characterized by a moderate, but positive default rate.

Intuitively, this market resembles traditional lending. There is an advertised low rate at which any borrowers can apply. Lenders invest some in due diligence, but prefer to reject those borrowers who are too costly to identify as good. Lenders also do not wish to avoid default at all cost.

Lenders with skill $\alpha \in [\alpha_1, \alpha_2]$ lend to good borrowers only with opacity $\omega \in [\omega_1, \omega_2]$. This is Region II. Each type of lender charges a different interest rate $r(\alpha)$ (or, equivalently, in the ω space, $r(\omega)$) in this region. $r(\alpha)$ has to be increasing in order to compensate for the cost of the higher skill. In fact,

$$\frac{\partial r(\alpha)}{\partial \alpha} = \frac{\partial K(\alpha)}{\partial \alpha}.$$

Then, the mass of lenders $w(\alpha)$ offering that interest rate is determined by the market clearing condition (A.26). That is, the mass of lenders with skill α entering has to be such that their capital is sufficient to serve the lenders whom they can just recognize as good.

Intuitively, this market resembles high-tech lending. The lenders present in this market invest a lot in their screening technology to serve the good borrowers who are hard to assess as good. This is expensive, therefore lenders ask for a high interest rate as compensation.

Note that the existence of Region II critically builds on the existence of Region I: the fact that less-skilled lenders cleanse the pool from the hardest-to-recognize bad borrowers in Region I (that is, the fact that $\gamma(r_p, 1, \alpha_1) \equiv 1$), makes it possible for high-skilled lenders in Region II to lend to good borrowers only.

As the interest rate is increasing in Region II, there might be a high enough $r(\alpha)$ to tempt some lenders to lend indiscriminately to all the remaining borrowers. The advantage is that such non-selective lenders can save the cost of precision $C(\alpha)$. The disadvantage is that they obtain a loan portfolio contaminated by the leftover bad borrowers, $L(\alpha_0, \alpha_1)$. If such α_2 exists, then there is a Region III where non-selective lenders offer the interest rate r_{NS} to any borrower who takes it. The mass of lenders who enter non-selectively with no precision $\alpha = 0$ is just enough to clear all demand from hard-to-recognize bad and good borrowers who were not served at any lower rate.

Intuitively, depending on the context, this region resembles a market with loan-sharks, high-rate credit cards, or low-documentation mortgages. Lenders are not skilled in due diligence. Instead they ask for a high interest rate to compensate for adverse selection. This market exists to profit form the hard-to-recognize good borrowers who cannot obtain loan anywhere else. The interest rate is high, but no applicants are rejected.

Note that similarly to Region II, market conditions in Region III crucially depend on what is happening in Region I. The interest rate in Region III is higher whenever there are more leftover bad borrowers, $L(\alpha_0, \alpha_1)$. This quantity is endogenously determined by the mass of lenders entering with different skills in Region I.

Note also that in the presence of Region III, our economy features non-assortative matching between lenders and borrowers. Throughout most of the market structure, more precise lenders lend to harder-to-recognize borrowers at higher interest rates. However, this pattern breaks down in Region III: There, lenders with the least precise screening technology lend to the hardest-to-recognize borrowers at the highest interest rate. As we will see, the competition and spill-overs across regions between lenders with highest and lowest precision has important consequences for impact of adoption of AI in the financial sector on financial inclusion.

Finally, recall that despite of the heterogeneity between the structure of the credit market across regions, the net return on lending is the same everywhere, Π. The heterogeneity across regions comes from the different margins along which lenders can obtain the same profit. In Region I, the increasing cost of skill is compensated by increasing loan quality. In Region II, it is compensated by increasing interest rates. While in Region III, lenders do not invest in more skill, but the adverse selection implied by the remaining leftover bad borrowers have to be compensated by the high interest rates.

2.5 Comparative Statics

Next, we would like to understand how does our economy affected by a change in the main parameters of the model.

First, we explore the effect of more intense overall competition, that is, higher $\bar{W}$, to our equilibrium.

Proposition 2. *A higher mass of lenders, $\bar{W}$, implies*

1. *lower level of profit, Π, for each lender,*
2. *a strictly lower interest rate $r(alpha)$ for each precision in each region*

3. *strictly lower thresholds* α_0, α_1

4. *a smaller mass of bad left-over borrowers* $L(\alpha_0, \alpha_1)$*.*

5. *a strictly better selection,* $\gamma(alpha)$ *in Region 1 for each precision*

The negative relationship between the intensity of competition and profits is intuitive. A larger mass of lenders implies less profit for each. As we mentioned, we exploit this strictly monotonic relationship in our equilibrium construction. This also implies that there is an equivalent way to set up and interpret our model. Instead of taking $\bar{W}$ as primitive and derive Π, we could entertain a large mass of potential entrants, think of Π as their exogenous cost of capital, and treat $K(\alpha)$ as the total cost of entry. Then, a free entry condition would give $\bar{W}$ as the equilibrium mass of active lenders. With this interpretation, each statement in Proposition 2 can be interpreted as the effect of lower cost of capital Π.

The second statement shows that more competition benefits each good borrower in the form of lower rates and more credit, whichever region they borrow at.

The third and fourth statements imply that bad borrowers are also strictly better off. More intense competition implies both the least and most skilled type lending at the lowest interest rate r_p have lower precision. Intuitively, the curves on Figure 2 are pushed to the left. Therefore, by the end of Region 1, at $\alpha = \alpha_1$ there are less bad borrowers are left-over to be cleared at higher interest rates. This makes it possible for interest rate in Region 3 be smaller. To sum up, bad borrowers who has been served at the lowest rate, r_p in region one, are still served in region one, but a lower rate. Some bad borrowers, who were served in Region 3, now are served in Region 1. Bad borrowers who remained to be served in Region 3, are served at a lower rate, r_{NS}.

The last statement characterizes the change in the relative mass of good to bad borrowers, who are served by lenders with a given $\alpha \in [\alpha_0, \alpha_1]$. For any given precision, this ratio, that is, lenders' portfolio quality improves. This is in line with the positive selection we emphasized by Lemma 1. When competition is more intense, each lender type in the interior of Region 1 enjoys a higher degree of positive selection coming from less skilled entrants. Intuitively, less skilled entrants, by accepting some bad very opaque bad borrowers, cleanse the pool for them.

Proposition 2 shows that a change in Π (or, equivalently an opposite change in $\bar{W}$) affects all equilibrium objects. To gain some intuition of the effect of changing informational technology $C(\alpha)$, now we turn to the general effect of a change in the adjusted cost $K(\cdot)$ on equilibrium. For this, we define a particular family of adjusted cost functions parameterized by the scalar κ.

Definition 3 (Selection-Preserving Cost Functions)**.** *Take a baseline equilibrium, with* $K(\alpha)$ *and implied* $\gamma(\alpha, 1, r_p)$ *loan quality in Region I. Any* κ *defines a continuous selection preserving adjusted cost function*

$$K_{SP}(\alpha; \kappa) \equiv \gamma(\alpha) \left(1 + D^{-1}\left(\frac{1}{\kappa} D\left(\frac{1 + K(\alpha)}{\gamma(\alpha)} - 1\right)\right)\right) - 1$$

where $\gamma(\alpha)$ *is defined in* (24)*.*

Note that the choice of $\kappa = 1$ gives back the original technology $K_{SP}(\cdot; 1) = K(\cdot)$. For any fixed α, increasing the parameter κ monotonically increases adjusted cost. However, as opposed to a change of Π only, a change in κ does not imply a parallel shift of the adjusted cost function. Instead, the extant of the change varies across α depending on the parameters and the equilibrium object $\gamma(\alpha, 1, r_p)$ in the baseline economy. Still, as the following Proposition states, an increase in κ implies an *overall technological improvement* in a well defined sense.

Proposition 3 (Overall Technological Improvement)**.** *Take a baseline equilibrium, with $K(\alpha)$ and implied $\gamma(\alpha)$. Replace $K(\alpha)$ with a corresponding $K_{SP}(\alpha; \kappa)$ for some κ for which $K_{SP}(\alpha; \kappa)$ is monotonically increasing in α. Parametrizing the equilibrium objects with κ gives the following comparative statics. For any κ, $\alpha_i(\kappa) = \alpha_i(1)$ for $i = 0, 1$, $r_p(\kappa) = D^{-1}\left(\frac{1}{\kappa} D(r_p)\right)$ and therefore, $\frac{dr_p}{d\kappa} < 0$, and $\frac{dr_{NS}}{d\kappa} > 0$ and $\frac{dr(\alpha;\kappa)}{d\kappa} > 0$ for all $\alpha \in [\alpha_1(\kappa), \alpha_2(\kappa)]$ and $\frac{d\alpha_2}{d\kappa} > 0$.*

The mass of left-over bad borrowers, $L(\alpha_0, \alpha_1)$ is unaffected by a change in κ.

The Proposition states that for any baseline economy, we can define a family of adjusted cost functions which decrease point-by-point with the parameter κ. Decreasing the adjusted cost by decreasing κ keeps the thresholds for Region I unchanged, decreases interest rates monotonically across all regions, and increases the threshold between Region II and Region III. Note that these cost functions are selection preserving in the sense that a change in κ does not affect the mass of left-over bad borrowers by the end of Region I.

In economic terms, the overall technological improvement in our competitive economy benefits all lenders. Those good and bad borrowers who have been served in Region I by traditional lenders are still served by them. Each obtain more credit at lower rates. Those good borrowers who were served in Region II by high-tech lenders continue to do so. Additionally, there will be some opaque good borrowers who are now served by these high-tech lenders instead of the non-selective ones in Region III. At the same time, Region III borrowers also obtain more credit at lower rates.

In Section 4 we discuss other variants of technological improvement which affect certain segments of the cost functions relatively more. We will use the family of Selection-Preserving Cost Functions as a benchmark for that analysis. As we will explain, when compared to this family, the technology improves relatively more (less) for low precision lenders, it tends to push down (up) $L(\alpha_0, \alpha_1)$. This is a downward (upward) force on the interest rate in Region III as this implies less adverse selection for non-selective lenders.

3 New Entrants: The Short Run Impact

We next investigate the impact of new lenders entering the market. In this Section, we are interested in the short-run impact. In particular, assume an incumbent equilibrium has formed at the beginning of period $t = 1$ and let $w(\alpha)$ denote the incumbent wealth distribution. We consider unexpected entry of a positive measure of new lenders at the end of period $t = 1$, before borrowing and lending takes place. These new entrants are endowed with the same basic technology β as incumbents, but they have a potentially different cost of entry

$$K^E(\alpha) = \Pi^E + C^E(\alpha)$$

for a chosen precision α. Let W^E denote the aggregate wealth of active entrants, and $w^E(\alpha)$ their endogenous wealth distribution. In order to study the short run impact of entry, we assume the incumbent lenders cannot change their precision, α, in response to the unexpected entry of the new lenders. However, they can change the interest rate they advertise. For simplicity, we assume that all incumbent cost is sunk, hence they stay active even if their lending activity provides less profit than Π given the new entrants.

Our new Entry Equilibrium follows closely the definition of the Incumbent Equilibrium under Case 2. The critical difference is that new entrants understand that there is a mass of $w(\alpha)$ incumbents who are present at the economy and whom new entrants has to compete against. Still, new entrants solve the analogous problem to incumbents given by

$$\tilde{\Pi}^E(\alpha) = \max_{r,z} \gamma^E(r, z; \alpha)(1 + r) - 1 \tag{25}$$

and

$$0 = \max_{\alpha} \tilde{\Pi}^E(\alpha) - K^E(\alpha). \tag{26}$$

However, the probability of a new entrant with precision α serves a good borrower, $\gamma^E(r, z; \alpha)$ is determined by the evolution of the measures of good and bad borrowers

$$G^E\left(\Omega^G; r, z, \alpha\right) = G^E\left(\Omega^G\right) - \int_{A(r,z,\alpha)} \frac{G^E\left(\Omega^G \cap I^G(\alpha, z); r, z, \alpha\right)}{G^E(I^G(\alpha, z); r, z, \alpha) + B^E(I^B(\alpha, z); r, z, \alpha)} \frac{1}{D(r(\alpha))} d\left(W(\alpha) + W^E(\alpha)\right) \tag{27}$$

$$B^E\left(\Omega^B; r, z, \alpha\right) = B^E\left(\Omega^B\right) - \int_{A(r,z,\alpha)} \frac{B^E\left(\Omega^B \cap I^B(\alpha, z); r, z, \alpha\right)}{G^E(I^G(\alpha, z); r, z, \alpha) + B^E(I^B(\alpha, z); r, z, \alpha)} \frac{1}{D(r(\alpha))} d\left(W(\alpha) + W^E(\alpha)\right) \tag{28}$$

which take into account the present of incumbents. Then, the definition of the equilibrium is as follows.

Definition 4 (**Entry Equilibrium**). *For any given W measure of incumbents with various screening precision, the Entry Equilibrium consists of*

1. *A measure W^E over new entrants screening precision such that $W^E([0, 1]) = W$,*
2. *choice-of-market function $r^E(\alpha)$ and a (binary) choice-of selectiveness function $z^E(\alpha)$ for each lender α in the support of W^E,*
3. *measures of good and bad borrowers available to lender α with selectivity z^E in market r^E: $G^E(\cdot; r, z, \alpha)$ and $B^E(\cdot; r, z, \alpha)$*

such that

1. *Given α, $r^E(\alpha)$ and $z^E(\alpha)$ solve the new entrants' problem* (25)*, with γ^E defined by the appropriately modified versions of* (5)*,* (6) *and* (7)*,*
2. *Every α in the support of W^E solves the entrants problem* (26)*,*
3. *The measures $G^E(\cdot; r, z, \alpha)$ and $B^E(\cdot; r, z, \alpha)$ satisfy* (27) *and* (28) *respectively.*

3.1 Equilibrium Construction and Properties

As we show in this section, the structure of the equilibrium remains similar. The Entry Equilibrium still features the hockey stick schedule, however, it can be "broken" as described by the following definition.

Definition 5 (The Broken Hockey Stick Interest Rate Schedule). *A broken hockey stick interest rate schedule is a version of a hockey stick interest rate schedule with the following modifications.*

The interest rate schedule can discretely jump at points ω_1 *and* ω_2*:* $r(\omega_1) \geq r_p$ *and* $r(\omega_2) \leq r_{NS}$.

Region II is divided into

1. *Region IIa: where good borrowers with* $\omega \in [\omega_1, \omega_2']$ *borrows at a single interest rate served by lenders with a moderately high degree of precision* $\alpha \in [\alpha_1, \alpha_2']$*, as in Region II in Proposition 1 and*

2. *Region IIb: where incumbent lenders with the highest degree of precision* $\alpha \in [\alpha_2', \alpha_2]$ *compete with a zero measure of non-selective lenders to serve good borrowers in each market characterized by an increasing interest rate schedule* $\tilde{r}(\omega)$ *for* $\omega \in [\omega_2', \omega_2]]$.

The following proposition states the main result of this section.

Proposition 4 (Entry Equilibrium). *Consider an incumbent equilibrium where lenders make profit* Π*. A measure* W^E *ex-ante identical new lenders enter with cost of capital* Π^E *and cost of precision* $C(\alpha)$.

The unique entry equilibrium is heterogeneous in lender precision and every incumbent lender makes $0 \leq \Pi' \leq \Pi$ *profits.*

The equilibrium features a hockey-stick interest rate schedule described in Proposition 1 with the following modifications.

1. *The interest rate schedule can discretely jump at points* α_1 *and* α_2*:* $r(\omega_g(\alpha_1)) \geq r_p$ *and* $r(\omega_g(\alpha_2)) \leq r_{NS}$.

2. *Region II is divided into*

   - *Region IIa: where good borrowers with* $\omega \in [\omega_1, \omega_2']$ *borrows at a single interest rate served by lenders with a moderately high degree of precision* $\alpha \in [\alpha_1, \alpha_2']$*, as in Region II in Proposition 1 and*

   - *Region IIb: where incumbent lenders with the highest degree of precision* $\alpha \in [\alpha_2', \alpha_2]$ *compete with a zero measure of non-selective lenders to serve good borrowers in each market characterized by an increasing interest rate schedule* $\tilde{r}(\omega)$ *for* $\omega \in [\omega_2', \omega_2]]$.

When any of these modifications apply, we refer to the interest rate schedule as having the form of a broken hockey stick.

The construction follows the steps of the construction of an Incumbent Equilibrium in Case 2. We describe these steps in the Appendix in detail. Here we give only a draft and highlight the main forces and properties. Then, in the next section we illustrate our main insights with a few economically relevant examples.

The main difference compared to the construction of an Incumbent Equilibrium is that in every step we have to check whether new entrants wish to enter in a given region. This decision is mostly driven by their comparative advantage for the given level of skill α. In particular, when the adjusted cost of entry at a given precision α, $K^E(\alpha)$ is small relative to $K(\alpha)$ for incumbents, new entrants tend to choose the given α and enter in the corresponding region. When entrants decide to do so, they affect the equilibrium along two channels. First, they enter because they can offer a lower interest rate creating losses for incumbents and gains for borrowers served in the given market. Second, new entrants change the pool of borrowers for all lenders with higher skill or offering a higher interest rate. This potentially creates spill-overs over the economy: a main focus of our analysis.

Region I Just as in Section 2.3, we start by finding the lowest skill-level, α_0^E and the interest rate r_p^E at which new entrants prefer to enter and which satisfies the zero profit condition. If that r_p^E is smaller than r_p in the incumbent equilibrium, there is entry in Region I.

Then, we find the distribution W^E by discretizing the space and taking the continuous limit. In particular, we find a sequence of values α_n and corresponding masses w_n^E in a way that all new entrants make net profit Π^E. In this case, it is possible that this process stops at an $\alpha_n < \alpha_1$. That is, new lenders do not enter everywhere in the original pooling region. This is typically the case when $K^E(\alpha)$ increases steeply eroding new entrants comparative advantage approaching the end of Region I. As the incumbents are still lending, as we will see in the next section, this might lead to a discrete jump in the interest rate schedule at α_1 from r_p^E to r_p. In contrast, when $K^E(\alpha)$ remains low compared to $K(\alpha)$, more specifically when

$$\alpha_1^E \equiv (K^E)^{-1}(r_p^E) > K^{-1}(r_p) = \alpha_1$$

, then new entrants enter everywhere, and the region extends to the right until α_1^E.

Region II As we noted in Definition 5 this region might feature a new segment.

Region IIa is similar to Region II of the incumbent equilibrium. That is, for the endogenous thresholds $\alpha \in [\alpha_1^E, \alpha_2^E]$ the interest rate follows

$$r^E(\alpha) = \min\left(K^E(\alpha), K(\alpha)\right).$$

The expression illustrates that if for any skill-level new entrants have a comparative advantage, they enter and push down interest rates accordingly.

Region IIb arises in an Entry Equilibrium when non-selective lenders find it profitable to compete with incumbent high-tech selective lenders with precision $\alpha \in [\alpha_2^E, \alpha_2]$. While only zero measure of them enter at a given market, their threat of entry is sufficient to push the interest rate down to a level $\tilde{r}(\alpha) < \min\left(K^E(\alpha), K(\alpha)\right)$. At that interest rate non-selective lenders, serving a mixture of hard-to-recognize bad and good borrowers make the same profit as in Region III below. However, the incumbent lenders suffer even largest losses than in region IIa.

In the next section, we provide some economically intuitive examples when this case arises.

Region III The outcome in Region III depends on which group has the comparative advantage to lend to good borrowers just above opacity $\omega = \beta + \alpha_2(1-\beta)$. These are the least opaque borrowers who were served in Region III in the Incumbent Equilibrium. Namely, we have to compare three interest rates.

$$D^{-1}\left(\frac{w^{NS}}{(L^E(\alpha_0^E, \alpha_1^E) + [G(1) - G(\omega_g(\alpha_2))])}\right) = r'$$
$$\frac{K^E(0)}{\gamma^{NS,E}(\alpha_2)} - 1 = r''$$
$$K^E(\alpha_2) = r'''$$

where $L^E(\alpha_0^E, \alpha_1^E)$ and $\gamma^{NS,E}(\alpha_2)$ are the left-over bad borrowers and the probability a non-selective lender entering at interest rate $r^E(\alpha_2)$ serves a good borrower. Both these objects are defined analogously to their counterpart in the Incumbent Equilibrium.

If $r' = \min(r', r'', r'')$ then non-selective incumbents have the comparative advantage over new entrants, and there will not be new entrants in this region. Whether the interest rate goes up or down critically depends on the whether there are more or less left-over bad borrowers after entry in Region I.

If $r'' = \min(r', r'', r'')$, than non-selective entrants have the comparative advantage, leading to a smaller interest rate in this region.

Perhaps it is useful to note that Region IIb we described in the previous part arises if in any of these cases $\min(r', r'', r'') < r_{NS}$. That is, when the interest rate what non-selective incumbents or entrants can offer is smaller than the non-selective interest rate in the Incumbent Equilibrium. This is the case, when non-selectives can compete with high-skilled incumbents.

Finally, if $r''' = \min(r', r'', r'')$ then the cost advantage of new entrants is sufficiently large that high-skilled entrants serve some of the good borrowers who previously were served by non-selective incumbents. It implies that Region II extends to the right.

In the next part, we provide more economic intuition for these various cases by some applications and examples.

4 Examples and Applications

In this part, we go through a number of examples and applications to shed more light on the economic intuition behind our results. To warm up, we start with two instructive examples where we specify new entrants' cost function $C^E(\alpha)$ in a way to limit the moving parts effecting the results. Then, we move on to applications to explain the potential spill-overs in our economy from entry in some region to the outcome of other regions.

4.1 Illustrative Example I: New Entrants, Same Technology

In this part, we study the benchmark entry equilibrium where entrants are endowed with the same cost function as incumbents, $C^E(\alpha) = C(\alpha)$.

Corollary 1 (Entry Equilibrium with Homogeneous Technology). *Consider an incumbent equilibrium with mass of lenders' $\bar{W}$, implied level of profit Π, and cost function $C(\alpha)$. Consider new entrants with total wealth $\bar{W}^E$, and identical information technology $C^E(\alpha) = C(\alpha)$.*

The unique Entry Equilibrium is equivalent to an Incumbent Equilibrium with a single group of lenders with total mass $\bar{W} + \bar{W}^E$ and information technology $C(\alpha)$.

The equilibrium has the same, hockey stick structure as stated in Proposition 1.

This result is intuitive. It simply says that when the new entrants have no technological advantage relative to incumbents, they spread out across the full spectrum of incumbent lender precision distribution. At each precision α all lenders, incumbents and new entrants, have paid the same cost, lend to the same portfolio of borrowers at the same interest rate and face the same default rate. As such, they will all be as well of as each other.

The proof proceeds by showing that and increase in incumbent aggregate wealth leads to a pointwise increase at wealth at every precision α which is chosen in the original incumbent equilibrium $\alpha \in \{0 \cup [\alpha_0, \alpha_2]\}$, and no increase in wealth outside this range, $\alpha \notin \{0 \cup [\alpha_0, \alpha_2]\}$. This benchmark result illustrates that without any technological improvement, an increase in supply of lender capital benefits all borrowers. Every incumbent lender makes less profit as the supply of capital has increased.

4.2 Illustrative Example II: Selection-Preserving Technology

In this example, we study how new entrants change the equilibrium in the special case when the cost function of new entrants are in the family we defined in Definition 3. As the following Lemma demonstrates under this treatment Entry Equilibrium and Incumbent Equilibrium can be compared easily. It is also apparent under what conditions the hockey stick interest schedule becomes broken.

Lemma 2. *Take a baseline equilibrium, with $K(\alpha)$ and implied $\gamma(\alpha)$. Consider new entrants with a corresponding selection preserving cost function, $K^E(\alpha) = K_{SP}(\alpha; \kappa)$ with parameter $\kappa < 1$. This implies an entrant equilibrium where $\alpha_0^E = \alpha_0$, $\alpha_1^E = \alpha_1$, $D(r_p^E) = \frac{1}{\kappa} D(r_p)$ and $r^{NS,E} < r_{NS}$ and $r(\alpha) > r^E(\alpha)$ for all $\alpha \in \left[\alpha_1^E, \alpha_2^E\right]$. Furthermore, there exists a critical $\bar{\Pi}^E$ (strictly smaller than Π) such that*

1. *if*

$$\Pi^E > \bar{\Pi}^E$$

   $\alpha_2^E > \alpha_2$ and there exists no Region IIb ($\omega_2' = \omega_2$).

2. *Otherwise, there exists a Region IIb where good firms with opacity*

$$\omega \in \left[\omega_g(\alpha_2^E), \omega_g(\alpha_2)\right]$$

   obtain loans at an interest rate, $\tilde{r}(\alpha)$, where $\tilde{r}(\alpha) < K^E(\alpha) < r(\alpha)$.

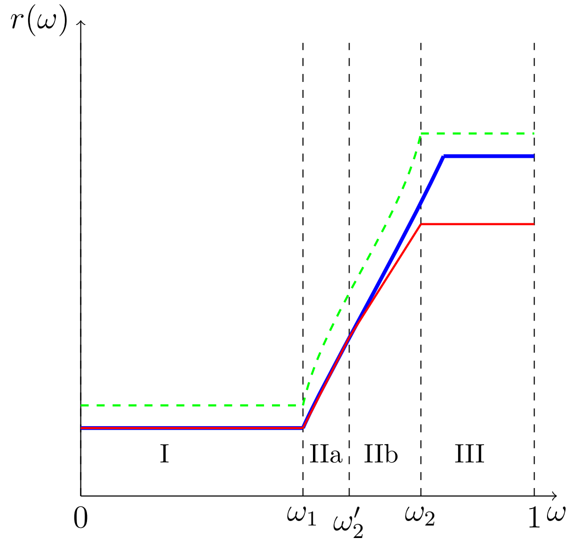

Figure 3: New Entry with Selection Preserving Technology

As we discussed in Section **??** when new entrants' cost function is $K^E_{SP}(\alpha;\kappa)$, a smaller parameter κ implies a larger comparative advantage in all regions. For any $\kappa < 1$ the interest rate in Region I is pushed down, the increasing interest rate schedule in Region IIa is smaller, and the non-selective interest rate in Region III decreases too. At the same time, the thresholds determining the limits across regions remain the same, except perhaps between Region II and III.

The last part of the Lemma describes when should we expect a region IIb to arise. That is, when will the hockey stick break. As it is apparent this is the case when the comparative advantage stems mostly from smaller cost of capital, Π^E. This favors non-selective entry. In fact, in that case, a large group of new entrants choose to not to learn at all, $\alpha = 0$, and threaten the highest skilled incumbents to enter and serve some of the most opaque good borrowers instead of them. To avoid this, those incumbents are forced to decrease the interest rate even below the level new entrants would offer. Hence, the highest skilled incumbents make the largest losses.

Figure 3 illustrates the results in Lemma 2. The dashed green curve is the interest rate schedule for good borrowers in the Incumbent Equilibrium. The blue curve is the same object in the Entry Equilibrium when the new entrants cost function is selection preserving and their cost of capital is large, $\Pi^E > \bar{\Pi}^E$. The red curve is when the opposite inequality holds (the red curve is on the top of the blue curve everywhere on the left from ω_2'). Note that region IIb is present only in the last case.

4.3 Applications: Big Data Innovation & Policy in Credit Markets

There is an active policy debate concerning how adoption of big data technologies impacts the credit markets and how it should be regulated. In this section, we use our model to study the consequences of growth in big data technologies and adoption of policies related to consumer data on the credit market equilibrium. We will interpret a reduction in the cost of screening precision as an improvement in data processing technology or improved access to consumer data. With this interpretation, a "directed" change in the screening cost affects the cost of vetting different borrowers and the rates their are offered deferentially.

In what follows we consider a few different realistic changes to lenders' screening cost, and study how the market structure changes with the entry of new lenders with the improved screening technology. This analysis allows us to identify the borrowers who benefit or are harmed by entry of new lenders, and in particular whether there is any *spillover* in equilibrium. To be precise, Definition 6 defines the equilibrium spillover of big data in our framework.

Definition 6 (Equilibrium Spillover of Big Data). *There is spillover of big data in equilibrium if there is a change in market conditions in markets where no new lender with improved screening technology enters.*

We will focus on three applications. In order to abstract away from the direct impact of increase in the supply of capital in the credit market, we restrict attention to when Π^E is high in all of these applications. As we have noted earlier in the paper, this corresponds to limited capital of new entrants.

The first application is *Open Banking*. Open Banking refers to mandatory data sharing among financial institutions, if requested by their clients. In the terms of reducing screening cost of borrowers, Open Banking makes the data of already-served borrowers more broadly available and reduces the cost of screening them for creditors, in a directed fashion. Note that in our model, the best served borrowers are served by low α lenders at low rates—many of them at the (lowest) pooling interest rate. Thus, we model Open Banking as a directed change in reducing cost of lower α. To be more specific, we assume there is $\hat{\alpha}$ such that $C^E(\alpha) < C(\alpha)$ for $\alpha < \hat{\alpha}$.

Interestingly, we find that adopting Open Banking has expected and unexpected implications for credit market conditions, which depends on the detail of implementation and which lenders primarily enter the market. Figure 4 illustrates two different possible outcomes that can happen if Open Banking is adopted in a credit market. The left panel corresponds to a directed cost reduction for $C(\alpha)$ for very low screening technology levels. We interpret this as a limited adoption of Open Banking. Alternatively, the right panel corresponds to a directed cost reduction for intermediate screening technology levels, which we interpret as a broader adoption of Open Banking.

Let us start from the more expected impact of Open Banking. As cost of low and/or intermediate screening technology decreases, borrowers who are served in the markets which are (partially) served by directly impacted lenders will benefit, irrespective of whether they are served by the new entrants or existing incumbents. In particular, assume $\hat{\alpha} < \alpha(\omega_1)$. There will be new lenders who choose screening technology $\alpha < \hat{\alpha}$ and enter the pooling market. Supply of capital will increase in this market and the pooling interest rate r_p falls. The left hand side of both panels in Figure 4 depicts the decrease in the prevailing interest rate in the pooling market, from baseline green to blue, as a result of adoption of Open Banking.

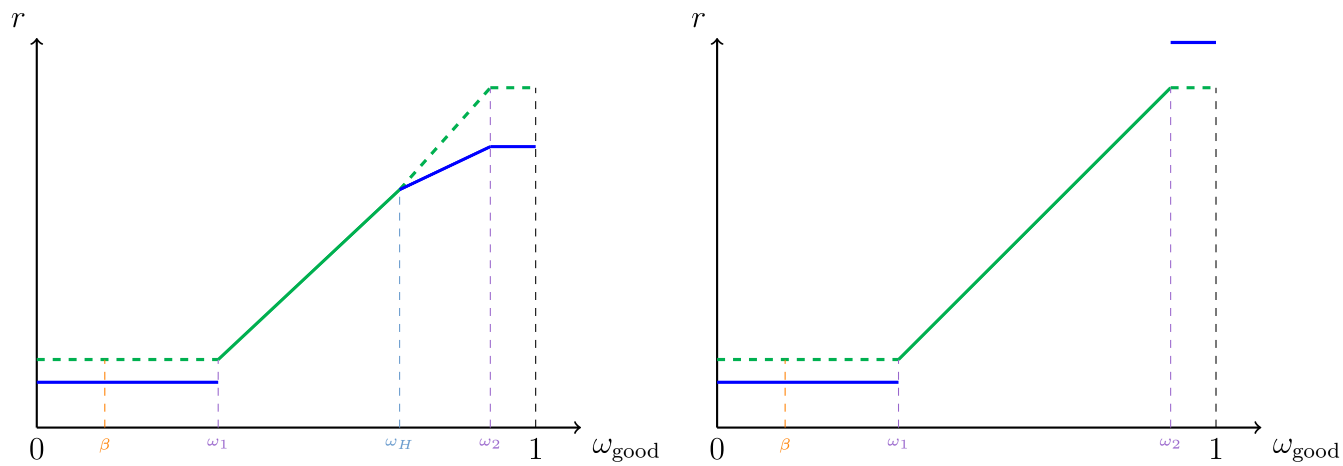

Figure 4: Adoption of Open Banking policy: The left panel corresponds to a directed cost reduction for $C(\alpha)$ for very low screening technology level. The right panel corresponds to a directed cost reduction for $C(\alpha)$ for intermediate screening technology level.

The more interesting impact of adoption of Open Banking in the credit market is through a spillover to market segments that are served at higher interest rates. Counter-intuitively, the interest rates in these segments can increase or decrease, depending on the exact implementation and scale of adoption of Open Banking. Lets call the borrowers who are served on the non-selective segment of the credit market, at the highest interest rates, the *financially excluded*, and consider why Open Banking can impact them positively or negatively.

The left panel of Figure 4 depicts the case where the financially excluded benefit from adoption of Open Banking, as displayed in the right end of the panel. This happens when $\hat{\alpha}$ is quite low, which we interpret as a limited adoption of Open Banking, when only the data of most served borrowers is shared even more broadly. In other words, $C^E(\alpha)$ reduces only for very low levels of α. This implies that all the new entrants have a relatively low level of screening expertise, thus, they disproportionately absorb the demand by bad opaque borrowers in the pooling segment. This in turn implies that the quality of the pool of remaining borrowers to be served at the highest interest rate improves. Recall that the lender-borrower matching is non-assortative in that region and only lenders with $\alpha = 0$ serve that market—incumbents or new entrants. Thus, as quality of the pool improves the interest rate falls, and there can be some extra spillover to the left, to the borrowers who were served at high interest rates in the right end of the cash-in-the-market region. This segment is referred to as Region IIb in Definition 5 and as explained in Section 4.1. Thus, there is *positive equilibrium spillover of big data*.

Alternatively, the right panel of Figure 4 depicts the case where the financially excluded are harmed by adoption of Open Banking and now face an even higher interest rate, as displayed in the right end of the panel. This happens when $\hat{\alpha}$ is in an intermediate range. We interpret this as a widespread adoption of Open Banking, as the data of a wider range of borrowers is shared among the institutions. This implies that the new entrants in the pooling market have an intermediate level of screening expertise, thus, they disproportionately absorb the demand by good relatively transparent borrowers in the pooling segment. This in turn

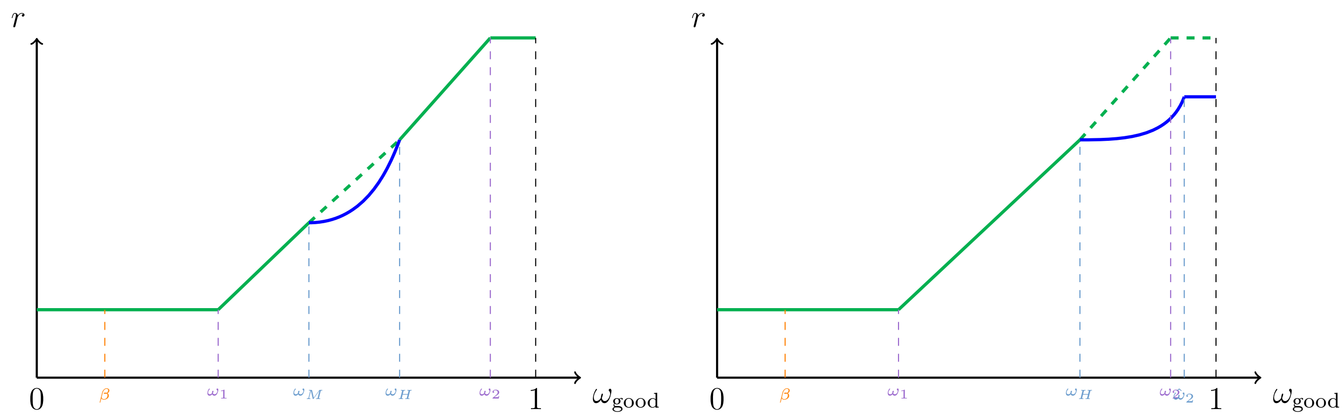

Figure 5: The left panel corresponds to a change in market structure if existing screening technologies become more widely available. The right panel corresponds to innovation in Artificial Intelligence (AI) and Machine Learning (ML) technology in the credit market.

implies that the quality of the pool of remaining borrowers to be served at the highest interest rate worsens. Thus, the lenders with the most basic level of expertise who lend in this market are only willing to provide credit at a higher interest rate and the credit conditions for these borrowers worsen. Thus, there is *negative equilibrium spillover of big data*.

The second application corresponds to cheaper availability of existing technologies in the credit market. We interpret this adoption as a decrease in the screening cost for intermediate, high levels of α, but not the highest levels of screening technology. The left panel of Figure 5 shows an example of this technological growth. In this example, $C^E(\alpha)$ falls only for the range $\alpha \in (\alpha(\omega_L), \alpha(\omega_H))$, and there is entry by new lenders who have access to this improved, cheaper screening technology only in this range. As the supply of capital has increased in these markets, interest rates fall and borrowers served in these markets benefit. However, there is no spillover to markets served at higher or lower interest rates.

Lastly, we consider innovation in Artificial Intelligence (AI) and Machine Learning (ML) technology in the credit market. As AI enables lenders to screen the less-traditional borrowers better and less traditional borrowers are highly opaque, we interpret this innovation as a directed cost reduction for high levels of screening technology. In particular, assume there is $\hat{\alpha}$ such that $C^E(\alpha) < C(\alpha)$ for $\alpha > \bar{\alpha}$. The right panel of Figure 5 depicts the impact of this innovation in the credit market structure. New entrants will enter at the high end of the market—right end of CIM market and possibly to the right of it, serve the most opaque borrowers and make a lot of profits. Importantly, there is no spillover in markets within the upward sloping range of interest rate. However, there is an interesting spillover to the financially excluded, as the lowest technology lenders who are serving the most opaque borrowers now face more fierce competition from very high tech lenders and are forced to offer lower interest rates. Here, there is *positive equilibrium spillover of big data*.

5 Model Variants

In this section, we illustrate the role of our main assumptions by investigating how the equilibrium changes when we replace them. In particular, we discuss two variants of our model.

First, we replace ex-ante identical lenders who optimally choose their technology with exogenous distribution of lender's types, $W(\cdot)$. We demonstrate that some features of the equilibrium are affected, for instance, not all bad borrowers are served. At the same time, the main structure, the hockey stick interest rate schedule in particular, is robust to this change.

Second, we replace the nested information structure to the more standard, conditionally iid structure where a lender with a higher precision might make mistakes about the same or different borrowers compared to a lender with a lower precision. We show that nestedness is critical for the structure of our equilibrium. In fact, there is a single region under this alternative information structure where all good and some bad borrowers are served.

5.1 Exogenous Distribution of Screening Technology

The next Proposition provides the comparison between the incumbent equilibrium with exogenous and with endogenous information.

Proposition 5 (Incumbent Equilibrium: Exogenous Distribution of Lenders). *Assume the wealth distribution of lenders $w(\alpha)$ is exogenous. An equilibrium exists, is unique, and features a hockey-stick interest rate schedule, defined by the endogenous thresholds on lender screening precision: α_0, α_1, α_2 satisfying $0 \leq \alpha_0 \leq \alpha_1 \leq \alpha_2 \leq 1$. The implied three regions differ from the ones defined in Proposition 1 in the following features:*

1. **Region I:** *Apart from every easy-to-recognize good borrower with $\omega \leq \omega_g(\alpha_1)$, and every hard-to-recognize bad borrower, (b, ω) with $\omega > \omega_b(\alpha_1)$, some easy-to-recognize bad borrowers with $\omega < \omega_b(\alpha_1)$ also borrow at r_p.*

2. **Region II:** *No change.*

3. **Region III:** *Only some easy-to-recognize bad borrowers, (b, ω) with $\omega < \omega_b(\alpha_2)$ borrow at r_{NS}. The rest of this group of borrowers are not served.*

   Lenders with the low technology level, $\alpha < \alpha_0$ who are non-selective along with very-high technology lenders $\alpha > \alpha_2$ (who are selective) lend at r_{NS}

Clearly, by and large, the fragmented equilibrium structure with the hockey-stick interest rate schedule is robust to taking the skill-distribution of lenders as exogenous. Still, there are some notable differences. From an economic point of view, perhaps the most significant difference is in Region 3. Recall that this region exists, because the interest rate is sufficiently high that even lending without any selection can be profitable. With endogenous information, this fact incentivizes a sufficient mass of uninformed lenders to enter and sweep up all remaining borrowers, bad and good. Thus, eventually all borrowers are served.

With exogenous information, unskilled lenders might not exist. Instead, lenders with insufficient skill to lend at the lowest rate, the ones with $\alpha < \alpha_0$, lend at the rate r_{NS}. Conditional on α_0 their mass is exogenous. They will pick up some left-over bad borrowers and some opaque good borrowers according to their skill and mass of capital. The remaining good borrowers will be picked up by the highly skilled lenders, $\alpha > alpha_2$. The remaining bad borrowers, unlike with endogenous information, are excluded.

Finally, the next proposition formalizes the equivalence between the economies with exogenous and endogenous information.

Proposition 6 (Duality of Distribution of Lenders' Wealth & Cost of Acquiring Precision). *For an exogenous lender precision distribution $w(\alpha)$, there exists an increasing cost function $C(\alpha)$ such that $w(\alpha)$ is an equilibrium if lenders choose α endogenously. Moreover, $C(\alpha)$ is unique on support of $w(\alpha)$.*

5.2 Non-nested Information Structure

In this section we provide a benchmark to highlight the features of the credit market structure that are unique to the nested information structure. This benchmark is identical to our main model in every respect, except that the following information structure replaces the nested information structure of Definition 1.

Definition 1-iid (IID Information structure). *When lender α meets borrower (τ, ω), screening technology β with precision α generates signal:*

$$x^{iid}(\tau; \alpha, \beta) = \begin{cases} \tau & \textit{if} \begin{cases} \tau = g & \textit{w. iid p. } 1-(1-\beta)(1-\alpha) \\ \tau = b & \textit{w. iid p. } 1-\beta(1-\alpha) \end{cases} \\ \\ -\tau & \textit{otherwise} \end{cases}$$

The IID information structure of Definition 1-iid shares two of the main properties of the nested information structure of Definition 1. First, the total error rate of lender with precision α is $1-\alpha$. Second, β determines the fraction of errors that are Type I and type II for every lender, $\frac{\text{type I error rate}}{\text{type II error rate}} = \frac{\beta}{1-\beta}$. More precisely,

$$\Pr(\text{false positive error}) = \beta(1-\alpha)$$
$$\Pr(\text{false negative error}) = (1-\beta)(1-\alpha)$$

The distinction between the two information structures is their *nestedness*. In particular, the IID information structure is non-nested. As such, unlike the nested information structure of Definition 1, the mistakes made by lenders of different technological precision with the IID information structure are perfectly uncorrelated. In other words, *opacity* of a borrower has no relevance for the percentage of lenders who identify the borrower's type correctly.

It turns out that this difference has profound implications on the market structure. Propositions 1-iid and 2-iid provide parallel results to Propositions 5 and 1 with IID information structure.

Proposition 1-iid (Incumbent Equilibrium: Exogenous Wealth Distribution). *Assume the wealth distribution of lenders $w(\alpha)$ with precision α is exogenous and lenders have iid information structure.*

For a given exogenous profit level $\Pi \geq 0$ for the active lender with the lowest level of expertise, a unique incumbent equilibrium exists. The equilibrium features a strictly increasing interest rate schedule.

Proposition 2-iid (Incumbent Equilibrium: Endogenous Wealth distribution). *Assume the information structure of lenders is iid. Consider a measure W of ex-ante identical lenders with cost $C(\alpha)$ who choose their precision and populate the credit market.*

Lender choose heterogeneous levels of precision and make the same profit $\Pi(W)$ in the unique incumbent equilibrium. The equilibrium features a strictly increasing interest rate schedule.

The above two results highlight the two important properties of the equilibrium market structure with nested information that is absent when lenders have IID information. First, recall that with a nested-information structure, the degree of fragmentation varies throughout the market. The best borrowers are served at a low, integrated interest rate with minimal fragmentation, i.e. in the pooling credit market segment. At higher levels of interest rate the interest rate schedule becomes fragmented and different borrowers are served at different interest rates, i.e. the separating credit market segment. With an IID information structure, the degree of fragmentation does not vary. All borrowers are served at separated interest rates and no pooling segment emerges.

Second, the nested information structure leads to a credit market which features non-assortative matching between lenders and borrowers. The lenders with the lowest precision serve the market segment which features the highest interest rate. This phenomena is absent with an IID information structure: The lenders who charge the highest interest rates are those who have the highest degree of precision, acquired at the highest cost when lender distribution is endogenously determined.

6 Conclusion

We develop an equilibrium model of credit markets with adverse selection and two-sided heterogeneity in lender and borrower types. Borrower type is two-dimensional, they are heterogeneous in both creditworthiness and opacity. Lenders have nested information structures and choose the precision of their screening technology to reduce the type I and II error rates that they make about borrowers' creditworthiness. Lenders also set interest rates to be compensated for different types of error that they will make. Borrowers choose interest rates and their quantity demanded to maximize their payoff.

In equilibrium, ex-ante homogeneous lenders choose heterogeneous levels of screening precision. The market structure is segmented with variable degrees of fragmentation across different level of borrower opacity and a hockey stick interest rate schedule. We then show that this market structure is robust to entry of new lenders and use our framework to investigate the impact of changes in big data technologies and policies on the financial sector. We find that adoption of AI technology benefits borrower who face high rates and improves financial inclusion. However, a mandatory data sharing policy not only does not have any spillover to the under-served population, but also exacerbates the inequality in financial access.

References

Babina, Tania, Saleem Bahaj, Greg Buchak, Filippo De Marco, Angus Foulis, Will Gornall, Francesco Mazzola, and Tong Yu, "Customer data access and fintech entry: Early evidence from open banking," *Journal of Financial Economics*, 2025, *169*, 103950.

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

Board, Simon, Moritz Meyer ter Vehn, and Tomasz Sadzik, "Recruiting talent," Technical Report, UCLA working paper 2017.

Chang, Briana, "Adverse selection and liquidity distortion," *The Review of Economic Studies*, 2018, *85* (1), 275–306.

Chari, VV, Ali Shourideh, and Ariel Zetlin-Jones, "Reputation and persistence of adverse selection in secondary loan markets," *American Economic Review*, 2014, *104* (12), 4027–4070.

Daley, Brendan and Brett Green, "Waiting for news in the dynamic market for lemons," *Econometrica*, 2012, *80* (4), 1433—-1504.

Eisfeldt, Andrea L, "Endogenous liquidity in asset markets," *The Journal of Finance*, 2004, *59* (1), 1–30.

Farboodi, Maryam and Péter Kondor, "Heterogeneous global booms and busts," *American Economic Review*, 2022, *112* (7), 2178–2212.

_ and _ , "Cleansing by tight credit: Rational cycles and endogenous lending standards," *Journal of Financial Economics*, 2023, *150* (1), 46–67.

Gale, Douglas, "A Walrasian theory of markets with adverse selection," *The Review of Economic Studies*, 1992, *59* (2), 229–255.

Goldstein, Itay, Chong Huang, and Liyan Yang, "Open banking under maturity transformation," *Available at SSRN*, 2022.

Guerrieri, Veronica and Robert Shimer, "Dynamic adverse selection: A theory of illiquidity, fire sales, and flight to quality," *American Economic Review*, 2014, *104* (7), 1875–1908.

_ , _ , and Randall Wright, "Adverse selection in competitive search equilibrium," *Econometrica*, 2010, *78* (6), 1823–1862.

He, Zhiguo, Jing Huang, and Jidong Zhou, "Open banking: Credit market competition when borrowers own the data," *Journal of financial economics*, 2023, *147* (2), 449–474.

Kaya, Ayça and Kyungmin Kim, "Trading dynamics with private buyer signals in the market for lemons," *The Review of Economic Studies*, 2018, *85* (4), 2318–2352.

Kurlat, Pablo, "Asset markets with heterogeneous information," *Econometrica*, 2016, *84* (1), 33–85.

Lauermann, Stephan and Asher Wolinsky, "Search with adverse selection," *Econometrica*, 2016, *84* (1), 243–315.

Li, Qi and Robert Shimer, "Liquidity and Price Dispersion in Markets with Information Heterogeneity," in "2019 Meeting Papers" number 1491 Society for Economic Dynamics 2019.

Lockwood, Ben, "Information externalities in the labour market and the duration of unemployment," *The Review of Economic Studies*, 1991, *58* (4), 733–753.

Strahan, Philip E, "Borrower risk and the price and nonprice terms of bank loans," *FRB of New York staff report*, 1999, (90).

Tirole, Jean, "Overcoming adverse selection: How public intervention can restore market functioning," *American economic review*, 2012, *102* (1), 29–59.

Appendix

A Proofs

A.1 Equilibrium construction: The Proofs of Proposition 1 and 2

Here, we give the remaining details of the equilibrium construction of our baseline case. The comparative statics results in Proposition 2 are part of the proof.

Steps of equilibrium construction Following the description in the main text, we summarize and detail the construction of the three equilibrium regions in the following steps.

1. Fix Π

2. Find the marginal lender:

$$\alpha_0 = \arg\min_{\alpha} \frac{\Pi + C(\alpha) + 1}{\gamma_0(\alpha)} - 1$$

   where γ_0 is defined in Equation (11).

3. Find the pooling interest rate:

$$r_p = \frac{\Pi + C(\alpha_0) + 1}{\gamma_0(\alpha_0)} - 1$$

4. Find α_1, the highest α lender in the r_p market. The indifference condition is:

$$r_p - C(\alpha_1) = \Pi.$$

   Therefore

$$\alpha_1 = \begin{cases} 1 & \text{if } (1 + r_p) - 1 - C(\alpha_1) > \Pi \text{ for all } \alpha \\ C^{-1}(\Pi - r_p) & \text{otherwise} \end{cases}$$

   Find W by discretizing, iterating and then taking the limit. Note that by assumption,

$$(1 + r_p)\,\gamma_0(\alpha') - 1 - C(\alpha') \leq \Pi \quad \text{for all } \alpha' > \alpha_0$$

   (i.e. no one can make Π with the original distribution).

   Furthermore, if W has a mass point w at α, then for all $\alpha' > \alpha$, the quantity $\gamma^S(r, \alpha')$ is increasing in w, and reaches $\gamma^S(r, \alpha') = 1$ for some finite w.

   Fix Δ. For each n, suppose W_Δ has a mass point w_n at α_n. Find w_n such that:

$$\max_{\alpha' \geq \min\{\alpha_n + \Delta, \alpha_1\}} (1 + r_p)\,\gamma(r_p, \alpha') - 1 - C(\alpha') = \Pi.$$

Call the argmax, α_{n+1}.[5] Continue until you reach $\alpha_{n+1} \geq \alpha_1$. Denote by W_Δ the resulting measure over the interval $[\alpha_0, \alpha_1]$. For any subset $A \subseteq [\alpha_0, \alpha_1]$, let $W(A) = \lim_{\Delta \to 0} W_\Delta(A)$.

5. For $\alpha > \alpha_1$, find $r(\alpha)$ by the indifference condition

$$r(\alpha) - C(\alpha) = \Pi,$$

and the density $w(\alpha)$ by the condition

$$w(\alpha) = D(r(\alpha))\, g(\beta + \alpha(1-\beta))(1-\beta).$$

The last region is the non-selective region, if it exists.

To find the first point of non-selective entry, for each $r \in [r_p, r(1)]$, compute

$$\Pi^{NS}(r) = \gamma^{NS}(r, 0)(1+r) - 1.$$

Let r_{NS} be defined by the minimum value of r within the interval $[0, r(1)]$ such that $\Pi^{NS}(r) \geq \Pi$, if such a value exists. This includes as a special case $r_{NS} \leq r_p$, in which case the only market is non-selective

If r_{NS} exists, then the non-selective region exists and lenders with no expertise serve borrowers there. Thus, W has a mass point at $\alpha = 0$, with mass

$$\frac{1}{D(r_{NS})}\left[G([0,1]; r_{NS}, 0) + B([0,1]; r_{NS}, 0)\right],$$

i.e. enough to satisfy all the demand at this point. Furthermore, $r(0) = r_{NS}$, i.e. unskilled lenders go to market r_{NS}, and $s(0) = 1$, i.e. they choose to be non-selective.

For any $A \subset [C^{-1}(r_{NS} - \Pi), 1]$, $W(A) = 0$, i.e. there is no entry for values of α that would require $r > r_{NS}$ in order to earn Π.

6. Compute the total mass of entrants W

7. As we prove below, Steps 1-6 $W(\Pi)$. The last step is to invert this function to find the level of Π that is consistent with the exogenous total wealth of lenders W.

The equilibrium is therefore:

1. The measure W defined by the construction above

2. Choice of markets and selectiveness

$$r(\alpha) = \begin{cases} r_{NS} & \text{if } \alpha = 0 \\ r_p & \text{if } \alpha \in [\alpha_0, \alpha_1] \\ \Pi + C(\alpha) & \text{if } \alpha > \alpha_1 \end{cases}$$

$$z(\alpha) = \begin{cases} 0 & \text{if } \alpha = 0 \\ 1 & \text{otherwise} \end{cases}$$

3. Measures G and B constructed as in the definition of equilibrium

[5] There are two possibilities: a corner solution with $\alpha_{n+1} = \alpha_n + \Delta$, or an interior solution the leaves a gap between $\alpha_n + \Delta$ and α_{n+1}.

Verifying optimality The above construction is an equilibrium as

1. Given their choice of α, all lenders are optimizing over r and s

   - Optimization over s is immediate because of the way we define the non-selective point.
   - Optimization over r:
     - For $\alpha > \alpha_1$; at higher r they would have to be non-selective and at lower r they get the same γ but a worse price
     - For $\alpha \in [\alpha_0, \alpha_1]$: at higher r they would have to be non-selective and at lower r they get weakly worse γ(same γ only for α_0) but a worse price
     - For $\alpha = 0$: by construction it's the only market where they can make at least Π, so they are optimizing

2. Lenders are optimizing over α

   - By construction, they are indifferent

3. The accounting conditions hold by construction

Showing Uniqueness and Proving Comparative Statics The following steps prove uniqueness and Proposition 2.

Proposition A.1. *Taking Π as given, the only equilibrium is the one that follows the construction above.*

Proof. We start by showing that the lowest-α active lender α_0 and lowest-active interest rate $r(\alpha_0)$ must be as in our equilibrium construction.

Lemma A.1. *Let α_0 be the lowest α in the support of W. Then*

$$\alpha_0 = \arg\min_{\alpha} \frac{\Pi + C(\alpha) + 1}{\gamma_0(\alpha)} - 1 \tag{A.1}$$

and

$$r(\alpha_0) = \frac{\Pi + C(\alpha) + 1}{\gamma_0(\alpha)} - 1 \tag{A.2}$$

Proof. Define, as in our construction

$$\gamma_0(\alpha) = \frac{G([0, \beta + \alpha(1-\beta)])}{G([0, \beta + \alpha(1-\beta)]) + B([1 - \beta + \alpha\beta, 1])}$$

In any equilibrium, it must be that

1. α_0 makes profits Π

$$\Pi = (1 + r(\alpha_0))\,\gamma_0(\alpha_0) - 1 - C(\alpha_0)$$

2. No lender can make profits higher than Π at interest rate $r(\alpha_0)$ when faced with the original pool

$$(1+r(\alpha_0))\gamma_0(\alpha)(r(\alpha_0),\alpha) - 1 - C(\alpha) \leq \Pi \qquad \text{for all } \alpha$$

This immediately implies (A.1) and (A.2) □

Now Let $\tilde{A}(r) = \{\tilde{\alpha} : r(\tilde{\alpha}) \leq r\}$ be the set of lenders that in equilibrium choose interest rate r or less. The remainder of good borrowers in market r, defined for an arbitrary subset $\Omega^G \subseteq [0,1]$, is:

$$R\left(\Omega^G; r\right) = G\left(\Omega^G\right) - \int_{\tilde{A}(r)} \left[G^S\left(\Omega^G; r(\alpha), \alpha\right) s(\alpha) + G^{NS}\left(\Omega^G; r(\alpha), \alpha\right)[1 - s(\alpha)]\right] dW(\alpha)$$

Lemma A.2. *There can be no interest rate r such that a lender α in the support of W chooses $r(\alpha) = r$, $s(\alpha) = 1$ and has a positive-measure remainder of acceptable good assets, i.e. $R([0, \beta + \alpha(1-\beta)], r) > 0$*

Proof. Assume the contrary for market r. Let α be the highest type lender that chooses market r. Lender α's profits are $(1+r)\gamma^S(r,\alpha) - 1 - C(\alpha)$. Since $R([0, \beta + \alpha(1-\beta)], r) > 0$ then for small enough ϵwe have $G([0, \beta + \alpha(1-\beta)], r+\epsilon, \alpha) > 0$ (i.e. the adjacent markets have acceptable good assets available). If lender α chooses market $r+\epsilon$, profits are:

$$(1+r+\epsilon)\gamma^S(r+\epsilon,\alpha) - 1 - C(\alpha) \tag{A.3}$$

where

$$\gamma^S(r+\epsilon,\alpha) = \frac{G^S([0,1]; r+\epsilon, \alpha)}{G^S([0,1]; r+\epsilon, \alpha) + B^S([0,1]; r+\epsilon, \alpha)}$$

The fact that $G([0, \beta + \alpha(1-\beta)], r+\epsilon, \alpha) > 0$ implies that $\gamma^S(r+\epsilon,\alpha) > 0$. The derivative with respect to ϵ is:

$$\begin{aligned}
\frac{\partial \gamma^S(r+\epsilon,\alpha)}{\partial \epsilon} &= \frac{\frac{\partial G^S}{\partial \epsilon}\left[G^S + B^S\right] - \left(\frac{\partial G^S}{\partial \epsilon} + \frac{\partial B^S}{\partial \epsilon}\right) G^S}{\left[G^S + B^S\right]^2} \\
&= \frac{\frac{\partial G^S}{\partial \epsilon} B^S - \frac{\partial B^S}{\partial \epsilon} G^S}{\left[G^S + B^S\right]^2} \\
&= \frac{\gamma^S(r+\epsilon,\alpha) B^S - \left(1 - \gamma^S(r+\epsilon,\alpha)\right) G^S}{\left[G^S + B^S\right]^2} \frac{w\left(r^{-1}(r+\epsilon)\right)}{d(r+\epsilon)}
\end{aligned}$$

Evaluating it at $\epsilon = 0$, this is

$$\begin{aligned}
\left.\frac{\partial \gamma^S(r+\epsilon,\alpha)}{\partial \epsilon}\right|_{\epsilon=0} &= \frac{\gamma^S(r,\alpha) B^S - \left(1 - \gamma^S(r,\alpha)\right) G^S}{\left[G^S + B^S\right]^2} \frac{w(\alpha)}{d(r)} \\
&= \frac{\frac{G^S}{G^S+B^S} B^S - \frac{B^S}{G^S+B^S} G^S}{\left[G^S + B^S\right]^2} \frac{w(\alpha)}{d(r)} \\
&= 0
\end{aligned}$$

Therefore, expression (A.3) is strictly increasing at $\epsilon = 0$, so choosing r cannot have been optimal for lender α. □

We now use Lemmas (A.1) and (A.2) to show that the equilibrium must be as in our equilibrium construction

Lemma A.3. *In any equilibrium with profit level* Π*, the measure of lenders who choose* $\alpha \in [0, \alpha_1]$, $r(\alpha) = r(\alpha_0)$ *and* $s(\alpha) = 1$ *must coincide with the one defined by the construction above.*

Proof. Assume the contrary. By Lemma A.1, no one chooses $\alpha < \alpha_0$ and $s = 1$, so the measures coincide on $[0, \alpha_0)$.

Now suppose that for some $\tilde{\alpha}$, the measure over $[\alpha_0, \tilde{\alpha})$ is strictly lower than the one defined by the construction above. Since, by Lemma REF, profits for $\alpha' > \tilde{\alpha}$ are increasing in the mass of lenders who choose $\alpha < \alpha'$, this implies that lenders cannot make profits Π by choosing $\alpha' > \tilde{\alpha}$ and $r(\alpha_0)$, so no lenders do so. But this implies that $R([0, \beta + \tilde{\alpha}(1-\beta)], r) > 0$, which contradicts Lemma A.2.

Conversely, suppose that for some $\tilde{\alpha}$,the measure over $[\alpha_0, \tilde{\alpha})$ is strictly greater than the one defined by the construction above. By Lemma REF, lenders can make profits higher than Π by choosing $\alpha > \tilde{\alpha}$ and $r(\alpha_0)$, which contradicts that there is an equilibrium with profits Π. □

Lemma A.4. *In any equilibrium with profit level* Π*, the measure of lenders who choose* $\alpha \in \left[\alpha_1, \alpha^{NS}\right]$ *and their choice of* r *and* s *must coincide with the one defined by the construction above.*

Proof. Conditional on α, the choice of r as in the construction above is the only one consistent with profits Π, so it remains to show that the measure coincides. By the same reasoning as in Lemma A.3, a lower measure would mean a positive remainder, contradicting Lemma A.2. Conversely, a greater measure would mean a negative remainder, which contradicts the definition of equilibrium. □

Lemma A.5. *The measure of entrants who choose* $\alpha = 0$*,* $s = 0$ *and* r_{NS} *must coincide with the one defined by the construction above.*

Proof. Given that the choices of selective lenders must coincide with construction ADD REFERENCE, lenders can make profits exactly Π only in the market r_{NS}. If the mass of non-selective entrants at r_{NS} was lower, then there would be a positive remainder of good assets, so lenders could make higher profits at market $r_{NS} + \epsilon$. If the mass of non-selective entrants at r_{NS} was higher, it would lead to a negative measure of remaining assets, which contradicts the definition of equilibrium. □

Together, Lemmas A.3, A.4 and A.5 imply the result. □

Proposition A.2. *Let* $W(\Pi)$ *be the total wealth of entering lenders as a function of their equilibrium profit level.* $W(\Pi)$ *is strictly decreasing*

Proof. We start by showing how the values of α_0, α_1 and $\gamma(\alpha)$ depend on Π.

Lemma A.6. α_0 *is increasing in* Π

Proof. Let

$$R(\alpha, \Pi) \equiv \frac{\Pi + C(\alpha) + 1}{\gamma_0(\alpha)} - 1$$

be the interest rate that will make a lender who chooses α make profits Π, conditional on facing the entire pool of borrowers.

By definition, α_0 solves

$$\alpha_0(\Pi) = \arg\min_{\alpha} R(\alpha, \Pi)$$

with FOC and SOC:

$$R_1(\alpha_0(\Pi), \Pi) = 0 \tag{A.4}$$

$$R_{11}(\alpha_0(\Pi), \Pi) > 0 \tag{A.5}$$

By the implicit function theorem:

$$R_{11}(\cdot)\,\alpha_0'(\Pi) + R_{12}(\cdot) = 0$$

$$\alpha_0'(\Pi) = -\frac{R_{12}(\cdot)}{R_{11}(\cdot)}$$

Compute $R_{12}(\cdot)$:

$$R_{12}(\cdot) = -\gamma_0(\alpha)^{-1}\,\gamma_0'(\alpha) < 0$$

which, together with the SOC (A.5), implies

$$\alpha_0'(\Pi) > 0$$

□

Lemma A.7. *α_1 is increasing in Π*

Proof. Since lender α_1 lends at $r_p = R(\alpha_0(\Pi), \Pi)$ and faces $\gamma(\alpha_1) = 1$, α_1 satisfies:

$$R(\alpha_0(\Pi), \Pi) - C(\alpha_1(\Pi)) = \Pi$$

Take derivatives on both sides:

$$R_1(\alpha_0(\Pi), \Pi)\,\alpha_0'(\Pi) + R_2(\cdot) - C'(\cdot)\,\alpha_1'(\Pi) = 1$$

Use the FOC (A.4) and rearrange:

$$\alpha_1'(\Pi) = \frac{R_2(\cdot) - 1}{C'(\cdot)} = \frac{\frac{1}{\gamma_0(\alpha)} - 1}{C'(\cdot)} > 0$$

□

Lemma A.8. *For any $\alpha \in [\alpha_0, \alpha_1]$, $\gamma(\alpha; \Pi)$ is decreasing in Π*

Proof. In an equilibrium with profits Π $\gamma(\alpha;\Pi)$ must satisfy:

$$\gamma(\alpha;\Pi)(1+R(\alpha_0(\Pi),\Pi))-C(\alpha)-1=\Pi$$

Rearrange:

$$\gamma(\alpha;\Pi)=\frac{1+\Pi+C(\alpha)}{(1+R(\alpha_0(\Pi),\Pi))}$$

Compute the derivative w.r.t. Π, use the FOC (A.4) and rearrange:

$$\begin{aligned}
\frac{\partial\gamma(\alpha;\Pi)}{\partial\Pi} &= \frac{(1+R(\alpha_0(\Pi),\Pi))-[R_1(\cdot)\alpha_0'(\Pi)+R_2(\cdot)][1+\Pi+C(\alpha)]}{(1+R(\alpha_0(\Pi),\Pi))^2}\\
&= \frac{(1+R(\alpha_0(\Pi),\Pi))-\frac{1}{\gamma_0(\alpha)}[1+\Pi+C(\alpha)]}{(1+R(\alpha_0(\Pi),\Pi))^2}\\
&= \frac{(1+R(\alpha_0(\Pi),\Pi))-\frac{1}{\gamma_0(\alpha)}[\gamma(\alpha;\Pi)(1+R(\alpha_0(\Pi),\Pi))]}{(1+R(\alpha_0(\Pi),\Pi))^2}\\
&= \frac{1-\frac{\gamma(\alpha;\Pi)}{\gamma_0(\alpha)}}{(1+R(\alpha_0(\Pi),\Pi))}<0
\end{aligned}$$

□

Lemma A.9. *Let* $B^R(\Omega^B)=B(\Omega^B)-\int_{\tilde{A}(r(\alpha_0))}B^S(\Omega^B;r(\alpha_0),\alpha)\,dW(\alpha)$ *be the measure (with density* b^R*) of bad borrowers who remain after market* $r(\alpha_0)$*.* $b^R(\omega)$ *is increasing in* Π *for all* ω*.*

Proof. Consider two equilibria, with profits Π and $\hat{\Pi}>\Pi$ respectively.

Define $\omega_0\equiv\omega_b(\alpha_0)=1-\beta+\alpha_0\beta$ and $\hat{\omega}_0\equiv\omega_b(\hat{\alpha}_0)=1-\beta+\hat{\alpha}_0\beta$. By Lemma A.6, $\hat{\omega}_0>\omega_0$. We know that borrowers with $\omega<\omega_0$ cannot borrow in either equilibrium, so $b^R(\omega)=\hat{b}^R(\omega)=b(\omega)$. Borrowers $\omega\in[\omega_0,\hat{\omega}_0]$ borrow with positive probability in the Π equilibrium but cannot borrow in the $\hat{\Pi}$ equilibrium, so $b^R(\omega)<\hat{b}^R(\omega)$ in this range.

Similarly, define $\omega_1\equiv\omega_b(\alpha_1)=1-\beta+\alpha_1\beta$ and $\hat{\omega}_1\equiv\omega_b(\hat{\alpha}_1)=1-\beta+\hat{\alpha}_1\beta$. By Lemma A.6, $\hat{\omega}_1>\omega_1$. We know that borrowers with $\omega>\hat{\omega}_1$ borrow with probability 1 in both equilibria, so $b^R(\omega)=\hat{b}^R(\omega)=0$. Borrowers $\omega\in[\omega_1,\hat{\omega}_1]$ borrow with probability 1 in the Π equilibrium but probability less than 1 in the $\hat{\Pi}$ equilibrium, so so $b^R(\omega)<\hat{b}^R(\omega)$ in this range.

Now assume the statement is not true. Since $b^R(\omega)\le\hat{b}^R(\omega)$ for all $\omega\le\hat{\omega}_0$ and $\omega\ge\omega_1$, there must be values $\omega\in(\hat{\omega}_0,\omega_1)$ such that $b^R(\omega)>\hat{b}^R(\omega)$. By continuity, this implies that there are at least two values $\omega\in(\hat{\omega}_0,\omega_1)$ such that $b^R(\omega)=\hat{b}^R(\omega)$. Let ω^* be the highest of these values, and define $\alpha^*\equiv\frac{\omega^*-(1-\beta)}{\beta}$. Let $B(\cdot;r_p,\alpha^*)$ and $\hat{B}(\cdot;\hat{r}_p,\alpha^*)$ be the measures of borrowers faced by lender α^* in the two equilibria. Since any lender that accepts bad borrower ω^* also accepts all bad borrowers $\omega>\omega^*$, the measures B and $\hat{B}$ coincide on the set $[\omega^*,1]$. By Lemma A.8, $\gamma(\alpha^*)>\hat{\gamma}(\alpha^*)$, and since $\gamma(\alpha^*)=\frac{G([0,\beta+\alpha^*(1-\beta)];r_p\alpha^*)}{G([0,\beta+\alpha^*(1-\beta)];r_p\alpha^*)+B([\omega^*,1];r_p,\alpha^*)}$, this implies that

$$G([0,\beta+\alpha^*(1-\beta)];r_p,\alpha^*)>\hat{G}([0,\beta+\alpha^*(1-\beta)];\hat{r}_p,\alpha^*)$$

Now consider a discrete approximation where lenders choose between α^* and $\alpha^* + \Delta$ for a fixed Δ. The mass of entrants at α^* must be such that lenders are willing to also enter at $\alpha^* + \Delta$. In the Π equilibrium, this means that:

$$\begin{aligned}
\Pi &= \gamma\left(\alpha^* + \Delta\right)\left(1 + r_p\right) - 1 - C\left(\alpha^* + \Delta\right) \\
\Rightarrow \gamma\left(\alpha^* + \Delta\right) &= \frac{\Pi + 1 + C\left(\alpha^* + \Delta\right)}{1 + r_p} \\
&\approx \gamma\left(\alpha^*\right) + \frac{C'\left(\alpha\right)}{1 + r_p}\Delta \\
&= \frac{G}{G + B} + \frac{C'\left(\alpha\right)}{1 + r_p}\Delta
\end{aligned} \tag{A.6}$$

where

$$\begin{aligned}
G &= G\left(\left[0, \beta + \alpha^*\left(1 - \beta\right)\right]; r_p, \alpha^*\right) \\
B &= B\left(\left[\omega^*, 1\right]; r_p, \alpha^*\right)
\end{aligned}$$

Suppose that the entrants into market α^* are enough to lend to a fraction θ of acceptable borrowers in that market. Given the construction above, this implies that:

$$\gamma\left(\alpha^* + \Delta\right) \approx \frac{\left(1 - \theta\right)G + \left(1 - \beta\right)g\Delta}{\left(1 - \theta\right)G + \left(1 - \beta\right)g\Delta + \left(1 - \theta\right)B - \beta b\Delta} \tag{A.7}$$

where:

$$\begin{aligned}
g &= g\left(\beta + \alpha^*\left(1 - \beta\right)\right) \\
b &= b^R\left(\omega^*\right)
\end{aligned}$$

Equate A.6 and A.7 and solve for $1 - \theta$ to obtain:

$$\begin{aligned}
\frac{G}{G + B} + \frac{C'\left(\alpha\right)}{1 + r_p}\Delta &= \frac{\left(1 - \theta\right)G + \left(1 - \beta\right)g\Delta}{\left(1 - \theta\right)G + \left(1 - \beta\right)g\Delta + \left(1 - \theta\right)B - \beta b\Delta} \\
\Rightarrow 1 - \theta &= \left(1 + r_p\right)\frac{\left(1 - \beta\right)gB + \beta bG}{\left(G + B\right)^2 C'(\alpha)} + \frac{\left[\beta b - \left(1 - \beta\right)g\right]}{G + B}\Delta
\end{aligned} \tag{A.8}$$

Now take the derivative with respect to Δ in A.6 and A.7, evaluate them at and equate them:

$$\begin{aligned}
\frac{C'\left(\alpha\right)}{1 + r_p} &= \frac{\left(1 - \beta\right)gB + \beta bG}{\left(G + B\right)^2} \\
\Rightarrow \left(1 + r_p\right)\frac{\left(1 - \beta\right)gB + \beta bG}{\left(G + B\right)^2 C'(\alpha)} &= 1
\end{aligned}$$

Replace this in (A.8):

$$\theta = \frac{\left(1 - \beta\right)g - \beta b}{G + B}\Delta$$

Since g is the original density, it is the same in both the Π and the Π' equilibrium; by assumption b is the same in both equilibria, and we proved above that this implies B is the same in both equilibria, and that $G > \hat{G}$. This implies that:

$$\hat{\theta} > \theta$$

This implies that at any point ω^* where the densities $b^R(\omega^*)$ and $\hat{b}^R(\omega^*)$ are equal, the density $\hat{b}^R(\omega^*)$ must have a lower (more negative) slope than $b^R(\omega^*)$. But since $\hat{b}^R(\omega) > b^R(\omega)$ for $\omega > \omega_1$, this cannot be true of the last time the densities intersect, which represents a contradiction. □

Lemma A.10. *For every borrower, the interest rate at which they borrow is increasing in* Π

Proof. Start with good borrowers. For $\omega \leq \beta + \alpha_1(1-\beta)$, they borrow at r_p. By (A.2) and the FOC (A.4), this is increasing in Π. For $\omega \geq \beta + \alpha_1(1-\beta)$, indifference for lenders implies that they borrow at $r = \Pi + C\left(\frac{\omega-\beta}{1-\beta}\right) + 1 - 1$, which is also increasing in Π . For bad borrowers, consider first the case where there is no non-selective region. In that case, bad borrowers borrow only in the r_p market, where the interest rate is increasing in Π. Instead, in the case where there is a nonselective region, Lemma A.9 implies that $\gamma^{NS}(r, \alpha)$ and therefore $\Pi^{NS}(r)$ is decreasing in Π. This in turn implies that r_{NS} is increasing in Π, so the interest rate increases both for bad borrowers who borrow at r_p and those who borrow at r_{NS} □

In the case where the equilibrium has a non-selective region, all borrowers end up borrowing but, by Lemma A.10, at higher interest rates when Π is higher. Therefore, since $d(r)$ is a decreasing function, the total wealth needed to lend to them is lower when Π is higher. In the case where there is no non-selective region, all good borrowers end up borrowing and, by Lemma A.9 fewer borrowers end up borrowing when Π is higher, so the same conclusion applies. □

A.2 Proof of Proposition 4

Proof. Suppose we have constructed an equilibrium with endogenous entry giving the measure W, $\alpha_0, \alpha_1, \alpha_2, r(\alpha)$ where incumbents have a cost function $C(\alpha)$ and cost of capital Π. Here we consider the construction of a new equilibrium where a group of entrants with cost function $C^E(\alpha)$ and cost of capital Π^E enter at various markets determining the equilibrium measure W^E. (Note that the equilibrium construction is the same regardless whether the incumbents' measure W is exogenously given or constructed as above.) We proceed in steps.

1. Let $\gamma^E\left(\alpha; r, \tilde{W}^E\right)$ the probability that an atomistic selective lender with precision α would get a good borrower on market r, under the assumption that all lenders from the incumbent group enter with measure W given by the incumbent equilibrium while new entrants enter with measure $\tilde{W}^E$. The formal expressions are analogous to (5)-(8)

with the exception that we fix $z=1$, and the updating rule (7-8) modifies to

$$G\left(\Omega^G;r,1,\alpha\right)=G\left(\Omega^G\right)-\int_{A(r,z,\alpha)}\Pr{}_G\left(\Omega^G;r\left(\alpha\right),z\left(\alpha\right),\alpha\right)\frac{1}{d\left(r\left(\alpha\right)\right)}D\left(W\left(\alpha\right)+\tilde{W}^E\left(\alpha\right)\right) \tag{A.9}$$

and

$$B\left(\Omega^B;r,1,\alpha\right)=B\left(\Omega^B\right)-\int_{A(r,z,\alpha)}\Pr{}_B\left(\Omega^B;r\left(\alpha\right),z\left(\alpha\right),\alpha\right)\frac{1}{d\left(r\left(\alpha\right)\right)}D\left(W\left(\alpha\right)+\tilde{W}^E\left(\alpha\right)\right) \tag{A.10}$$

2. Find the (potential) marginal entrant in the pooling region and the pooing interest rate by defining

$$\alpha_0^E\left(r\right)\equiv\arg\min_{\alpha\leq\alpha_1}\frac{\Pi^E+C^E\left(\alpha\right)+1}{\gamma^E\left(\alpha;r,0\right)}-1$$

and solving

$$r_p^E=\frac{\Pi^E+C^E\left(\alpha_0^E\left(r_p^E\right)\right)+1}{\gamma^E\left(\alpha_0^E\left(r_p^E\right);r_p^E,0\right)}-1. \tag{A.11}$$

If $r_p^E\geq r_p$ there is no entry in the pooling region: $W^E\left([0,\alpha_1]\right)=0$. In this case, $\alpha_1^E=\alpha_1$, and the construction continues from step 7.

3. Find the implied cash-in-the market interest rate, $\hat{r}\left(\alpha\right)$, by

$$w\left(\alpha\right)=D\left(\hat{r}\left(\alpha\right)\right)g\left(\beta+\alpha\left(1-\beta\right)\right)\left(1-\beta\right)$$

for all α (note that for $\alpha\geq\alpha_1$ $\hat{r}\left(\alpha\right)=C\left(\alpha\right)+\Pi$ by definition.) let $\alpha_1'^E$ the smallest element of the set $\left\{\alpha:\hat{r}\left(\alpha'\right)\geq r_p^E \text{ for all } \alpha'\in[\alpha,\alpha_1]\ \right\}$. This is the end of the new pooling region if new entrants are not active everywhere along the region.

4. find $\alpha_1''^E$, by the indifference condition:

$$r_p^E-C^E\left(\alpha_1''^E\right)=\Pi^E. \tag{A.12}$$

This is the end of the new pooling region if new entrants are active at that point. Let $\alpha_1^E=\max\left(\alpha_1'^E,\alpha_1''^E\right)$

5. Find the measure $W^E\left((0,\alpha_1^E]\right)$ by discretizing, iterating and then taking the limit. In particular, for any fixed Δ we are specifying steps to build up a discrete measure W_Δ, with finite number of masses $\{w^0,w^1...w^n,...\}$ at the corresponding mass points $\{\alpha^0,\alpha^1...\alpha^n,...\}$ where $\alpha^0=\alpha_0^E$ Taking the limit $\Delta\to 0$ will give $W^E\left((0,\alpha_1^E]\right)$.

   (a) Fix a Δ.

   (b) step 1:

      i. Let $\alpha^0=\alpha_0^E$. By assumption,

$$\left(1+r_p^E\right)\gamma^E\left(\alpha;r_p^E,0\right)-1-C^E\left(\alpha'\right)\leq\Pi^E\quad\text{for all } \alpha'>\alpha^0$$

      (i.e. no one can make Π^E with the incumbent distribution)

ii. However, for any $\alpha' > \alpha^0$, $\gamma\left(\alpha'; r_p^E, W_\Delta^0\right)$ is increasing in w, and reaches $\gamma^E\left(\alpha'; r_p^E, W_\Delta^0\right) = 1$ for some finite w where W_Δ^0 is defined as a measure with w mass at α^0 point and 0 mass everywhere else.

iii. Find w^0 such that:

$$\max_{\alpha' \geq \min\left\{\alpha_0^E + \Delta, \alpha_1\right\}} \left(1 + r_p^E\right) \gamma^E\left(\alpha'; r_p^E, W_\Delta^0\right) - 1 - C^E\left(\alpha'\right) = \Pi^E$$

Call the argmax α^1. [There are two possibilities: a corner solution with $\alpha^1 = \alpha^0 + \Delta$. or an interior solution the leaves a gap between $\alpha^0 + \Delta$ and α_1^E

(c) step $n > 1$:

i. By assumption,

$$\left(1 + r_p^E\right) \gamma^E\left(\alpha; r_p^E, W_\Delta^{n-1}\right) - 1 - C^E\left(\alpha'\right) \leq \Pi^E \quad \text{for all } \alpha' > \alpha^{n-1}$$

(i.e. no one can make Π^E with the incumbent distribution)

ii. However, for any $\alpha' > \alpha^{n-1}$, $\gamma\left(\alpha'; r_p^E, W_\Delta^{n-1}\right)$ is increasing in w, and reaches $\gamma\left(\alpha'; r_p^E, W_\Delta^{n-1}\right) = 1$ for some finite w where W_Δ^{n-1} is defined as a measure $\{w^0, w^1 ..., w^{n-2}, w\}$ at the corresponding mass points $\{\alpha^0, \alpha^1 .., \alpha^{n-2}, .\alpha^{n-1}\}$ 0 mass everywhere else.

iii. Find w^{n-1} such that:

$$\max_{\alpha' \geq \min\left\{\alpha_0^E + \Delta, \alpha_1\right\}} \left(1 + r_p^E\right) \gamma^E\left(\alpha'; r_p^E, W_\Delta^{n-1}\right) - 1 - C^E\left(\alpha'\right) = \Pi^E$$

Call the argmax α^n.

(d) stop when $\alpha^n + \Delta > \alpha_1^E$.

(e) For any subset $A \subseteq \left[\alpha_0^E, \alpha_1^E\right]$, let

$$W^E(A) = \lim_{\Delta \to 0} W_\Delta^E(A) \tag{A.13}$$

6. If $\alpha_1^E < \alpha_1$ check whether in the range $\alpha \in \left[\alpha_1^E, \alpha_1\right]$ $\hat{r}(\alpha)$ is non-monotonic. If yes, we need ironing (To Pablo: how, exactly?). With an abuse of notation, let $\hat{r}(\alpha)$ be the ironed version.

7. Find the point of non-selective entry, if it exists. Let $B^{NS,E} = B\left([0,1]; r_P^E, 1, \alpha_1^E\right)$ be all bad borrowers who did not borrow at the pooling market. Let $\gamma^{NS,E}(\alpha) = \frac{[G(1) - G(\beta + \alpha(1-\beta))]}{[G(1) - G(\beta + \alpha(1-\beta))] + B^{NS,E}}$ the fraction of good applicants a non-selective entrant gets if enters at market $r_{CIM}(\alpha)$.First,we determine the group of lenders who will lend to good firms with opacity just above $\beta + \alpha_2(1-\beta)$.

$$d^{-1}\left(\frac{w^{NS}}{\left(B^{NS,E} + \left[G(1) - G\left(\beta + \alpha_2(1-\beta)\right)\right]\right)}\right) = r'$$

$$\frac{\left(1 + \Pi^E\right)}{\gamma^{NS,E}(\alpha_2)} - 1 = r''$$

$$C^E(\alpha_2) + \Pi^E = r'''$$

(a) if $r' = \min(r', r'', r'')$ then non-selective incumbents serve these good firms (along with bad ones). there are no non-selective entrants

i. if also $r' > r_{NS}$, then $r^{NS,E} = r'$, $\alpha_2^E = \alpha_2$ (where there is a jump).

ii. if $r' < r_{NS}$, then α_2^E is smaller than α_2, and incumbent non-selectives enter at $\tilde{r}(\alpha) < r_{CIM}(\alpha)$ in the range $\alpha \in \left[\alpha_2^E, \alpha_2\right]$ and $r^{NS,E} = \tilde{r}(\alpha_2)$. We describe how to determine $\tilde{r}(\alpha)$ and α_2^E in the next step.

(b) if $r'' = \min(r', r'', r'')$ the non-selective entrants serve these good firms. (along with bad ones) and similarly

i. if also $r'' > r_{NS}$, then $r^{NS,E} = r'$, $\alpha_2^E = \alpha_2$ (where there is a jump).

ii. if $r'' < r_{NS}$, then α_2^E is smaller than α_2, and nonselective entrants enter at $\tilde{r}(\alpha) < r_{CIM}(\alpha)$ in the range $\alpha \in \left[\alpha_2^E, \alpha_2\right]$ and $r^{NS,E} = \tilde{r}(\alpha_2)$. We describe how to determine $\tilde{r}(\alpha)$ and α_2^E in the next step..

(c) if $r''' = \min(r', r'', r'')$ then skilled entrants serve these good firms. It implies that the CIM region will extend to the right, $\alpha_2^E > \alpha_2$. We have to find α_2^E as follows. Let $\alpha_2^{E\prime}$ and $\alpha_2^{E\prime\prime}$ solve (these are the intercepts of the CIM interest rate curve of new entrants with the rate nonselective entrants and incumbents were willing to offer to the same group)

$$\frac{\left(1+\Pi^E\right)}{\gamma^{NS,E}\left(\alpha_2^{E\prime\prime}\right)} - 1 = C^E\left(\alpha_2^{E\prime\prime}\right) + \Pi^E\left(\alpha_2^{E\prime\prime}\right)$$

$$d^{-1}\left(\frac{w^{NS}}{\left(B^{NS,E} + \left[G(1) - G\left(\beta + \alpha_2^{E\prime}(1-\beta)\right)\right]\right)}\right) = C^E\left(\alpha_2^{E\prime}\right) + \Pi^E\left(\alpha_2^{E\prime}\right)$$

(if any of these equations do not have a solution in the unit interval, pick $\alpha_2^{E\prime\prime} = 1, \alpha_2^{E\prime} = 1$ respectively. If there is more than one, pick the smaller.) Let $\alpha_2^E = \min\left(\alpha_2^{E\prime\prime}, \alpha_2^{E\prime}\right)$, and $r^{NS,E} = C^E\left(\alpha_2^E\right) + \Pi^E\left(\alpha_2^E\right)$. If, $\alpha_2^E = \alpha_2^{E\prime\prime}$, then nonselective entrants, otherwise, nonselective incumbents clear good firms with opacity higher than $\beta + \alpha_2^E(1-\beta)$, while skilled incumbents clear good firms at the

$$r_{CIM}(\alpha) = C^E\left(\alpha_2^{E\prime}\right) + \Pi^E\left(\alpha_2^{E\prime}\right)$$

in the range $\alpha \in \left[\alpha_2, \alpha_2^E\right]$. (there might be a jump at α_2 upward, if $r''' > r_{NS}$).

8. Determining $\tilde{r}(\alpha)$ and α_2^E

(a) In the case of 8bii, we have non-selective entry. Then α_2^E is the solution of

$$\gamma^{NS,E}(\alpha)\left(1 + \min\left(C^E(\alpha) + \Pi^E(\alpha), C(\alpha) + \Pi(\alpha)\right)\right) = \left(1+\Pi^E\right).$$

and $\tilde{r}(\alpha)$ is given by the equivalent definitions of

$$\gamma^{NS,E}(\alpha)\left(1+\tilde{r}(\alpha)\right) = \left(1+\Pi^E\right) \tag{A.14}$$

or

$$\gamma^{NS,E}(\alpha)\left(1+\tilde{r}(\alpha)\right) = \gamma^{NS,E}\left(\alpha_2^E\right)\left(1 + r_{CIM}^E\left(\alpha_2^E\right)\right) \tag{A.15}$$

the mass of non-selectives entering along with incumbents in markets $r \in \left[r^E_{CIM}\left(\alpha^E_2\right), r^{NS,E}\right]$ is given by

$$w^{NS,E}(\alpha) = -\phi'(\alpha)\left(B^{NS,E} + \left[G(1) - G(\beta + \alpha(1-\beta))\right]\right) D(\tilde{r}(\alpha)) \quad \text{(A.16)}$$

which are only new entrants in case B and only incumbents in case A and

$$\phi(\alpha) = \frac{w(\alpha)}{D(\tilde{r}(\alpha))\, g(\beta + \alpha(1-\beta))(1-\beta)} \quad \text{(A.17)}$$

with

$$\phi\left(\alpha^E_2\right) = 1.$$

In this case, it must be that the the total required capital is larger than W^{NS}, that is

$$\int_{\alpha^E_2}^{\alpha_2} -\phi'(\alpha)\left(B^{NS,E} + \left[G(1) - G(\beta + \alpha(1-\beta))\right]\right) D(\tilde{r}(\alpha))\, d\alpha + \\ + D(\tilde{r}(\alpha_2))\,\phi(\alpha_2)\left(B^{NS,E} + \left[G(1) - G(\beta + \alpha_2(1-\beta))\right]\right) > W^{NS}.$$

In this case, there is an atom at $\tilde{r}(\alpha_2)$, $w^{NS,E}$ given by the difference of the two sides of this inequality.
In the case of 8cii, we do not have non-selective entry. Instead, we have to figure out how the incumbent non-selectives enter in the new equilibrium. For this,we have to conjecture an α^E_2, which by (A.15) gives an $r(\alpha)$, which by (A.17) gives a $\phi(\alpha)$ for which we can check whether

$$\int_{\alpha^E_2}^{\alpha_2} -\phi'(\alpha)\left(B^{NS,E} + \left[G(1) - G(\beta + \alpha(1-\beta))\right]\right) D(\tilde{r}(\alpha))\, d\alpha + \\ + D(\tilde{r}(\alpha_2))\,\phi(\alpha_2)\left(B^{NS,E} + \left[G(1) - G(\beta + \alpha_2(1-\beta))\right]\right) = W^{NS}.$$

As we explain now, under these expressions all the non-selective lenders, make the same profit when enter at $r \in \left[r^E_{CIM}\left(\alpha^E_2\right), \tilde{r}(\alpha_2)\right]$ and all the good with $\alpha > \alpha^E_2$ are cleared at weakly increasing interest rate in α. Suppose that $B(\alpha)$ is the total number of bad borrowers who are applying to market α. Then, we define $\phi(\alpha)$ is

$$\phi(\alpha) = \frac{B(\alpha)}{B^{NS,E}} \quad \text{(A.18)}$$

the fraction compared to $B^{NS,E}$. This implies that the measure of total borrowers non-selectives face must be

$$\underbrace{B(\alpha)}_{\text{bad borrowers}} + \underbrace{\phi(\alpha)\left[G(1) - G(\beta + \alpha(1-\beta))\right]}_{\text{good borrowers that did not't borrow from non-selective lenders before } \alpha} \\ = \phi(\alpha)\left(B^{NS,E} + \left[G(1) - G(\beta + \alpha(1-\beta))\right]\right)$$

This explains why (A.16) is a market clearing condition. Also, the profit of non-selectives in market α is given by

$$\frac{\phi(\alpha)[G(1)-G(\beta+\alpha(1-\beta))]}{\phi(\alpha)[G(1)-G(\beta+\alpha(1-\beta))]+\phi(\alpha)B^{NS,E}}(1+\tilde{r}(\alpha))-1-C^E(0)$$
$$=\frac{[G(1)-G(\beta+\alpha(1-\beta))]}{[G(1)-G(\beta+\alpha(1-\beta))]+B^{NS,E}}(1+\tilde{r}(\alpha))-1-C^E(0)$$
$$=\gamma^{NS,E}(\alpha)(1+\tilde{r}(\alpha))-1-C^E(0).$$

This expression explains how the definition of $\tilde{r}(\alpha)$ implies the same non-selective profit for each α in expressions (A.14)-(A.14). Also, (A.17) must hold, because the incumbent with precision α, and with one unit of capital has to serve the good types who cannot be served by a slightly lower α

$$D(\tilde{r}(\alpha))\phi(\alpha)g(\beta+\alpha(1-\beta))(1-\beta)=w(\alpha).$$

Intuitively, the left hand side is the demand for capital from good applicants with $\omega=\beta+\alpha(1-\beta)$, while the right hand side is the supply of incumbents with exactly that precision.
Finally, (A.18) is given by the following arguments. $\phi\left(\alpha_2^E\right)=1$ by definition. Then

$$\begin{aligned}\phi(\alpha+\epsilon) &= \\ &= \phi(\alpha)-\frac{w^{NS}(\alpha)\epsilon}{[(B^{NS,E}+[G(1)-G(\beta+\alpha(1-\beta))]])]d(\tilde{r}(\alpha))} \\ &= \phi(\alpha)-\frac{\left[\phi(\alpha)D(\tilde{r}(\alpha))-\frac{w(\alpha)}{g(\beta+\alpha(1-\beta))(1-\beta)}\right]}{D(\tilde{r}(\alpha))} \\ &= \frac{w(\alpha)}{D(\tilde{r}(\alpha))g(\beta+\alpha(1-\beta))(1-\beta)}\end{aligned}$$

which also implies ,

$$\frac{\phi(\alpha+\epsilon)-\phi(\alpha)}{\epsilon}=-\frac{w^{NS}(\alpha)}{(B^{NS,E}+[G(1)-G(\beta+\alpha(1-\beta))])D(\tilde{r}(\alpha))}$$

or, in the limit $\epsilon\to 0$, (**??**).

9. The CIM region is $\alpha\in[\alpha_1^E,\alpha_2^E]$ and in this range

$$r_{CIM}^E(\alpha)=\left\{\begin{array}{ccc}\min\left(\hat{r}(\alpha),C^E(\alpha)+\Pi^E\right) & if & \alpha<\alpha_2 \\ C^E(\alpha)+\Pi^E & otherwise & \end{array}\right\}$$

Entry in the CIM region with precision α is

$$w^E(\alpha)\equiv\min\left(0,D\left(r_{CIM}^E(\alpha)\right)g(\beta+\alpha(1-\beta))(1-\beta)-w(\alpha)\right).$$

.

□

A.3 Proof of Proposition 5

We prove this result first by constructing the equilibrium allocation and then adding the verification that it indeed satisfies the definition of equilibrium, and that it is unique.

We start from Region I. Start from a conjectured value for r_p: the lowest interest rate that is offered by any lender, and a conjectured level of profits Π for the least-informed lender $\alpha = 0$. Since r_p is the lowest interest rate available, it must attract all the borrowers. Therefore the lowest-α lender who is active in market r_p and is selective obtains an average quality of 11. We'll find lender α_0 as the lender who makes profits of Π by lending in this market, so α_0 solves:

$$\gamma_0\left(\alpha_0\right)\left(1+r_p\right)-1=\Pi \tag{A.19}$$

Since $\gamma_0\left(\alpha\right)$ is increasing in α, equation (A.19) defines a one-to-one negative relationship between r_p and α_0: the higher the interest rate, the less skilled the first lender needs to be to achieve profits Π. It is convenient to invert this relationship: conjecture a value of α_0 and use (A.19) to define an interest rate $r_p\left(\alpha_0\right)$.

Lenders of different skill α will pool in market r_p. Since they pick borrowers sequentially in order of increasing α, the composition of the pool they face changes with α. Denote by $g\left(\cdot;\alpha,\alpha_0\right)$ and $b\left(\cdot;\alpha,\alpha_0\right)$ the pdfs over good and bad borrowers respectively that remain when it's lender α's turn to lend in market r_p (taking α_0 as given). Then compute:

$$G\left(\alpha;\alpha_0\right)=\int_0^{\beta+\alpha(1-\beta)} g\left(\omega;\alpha,\alpha_0\right)d\omega \tag{A.20}$$

$$B\left(\alpha;\alpha_0\right)=\int_{1-\beta+\alpha\beta}^{1} b\left(\omega;\alpha,\alpha_0\right)d\omega \tag{A.21}$$

$$T\left(\alpha;\alpha_0\right)=G\left(\alpha;\alpha_0\right)+B\left(\alpha;\alpha_0\right) \tag{A.22}$$

$$\gamma\left(\alpha;\alpha_0\right)=\frac{G\left(\alpha;\alpha_0\right)}{G\left(\alpha;\alpha_0\right)+B\left(\alpha;\alpha_0\right)} \tag{A.23}$$

Expressions (A.20)-(A.23) represent, respectively the total mass of acceptable good borrowers, the total mass of acceptable bad borrowers, the total mass of acceptable borrowers and the average quality received by lender α in market r_p.

When lender α lends he serves $\frac{w(\alpha)}{D(r_p)}$ borrowers, pro-rated among the $T\left(\alpha;\alpha_0\right)$ acceptable ones. Therefore, for every ω that lender α finds acceptable, the number of borrowers who remain unserved goes down by a fraction equal to:

$$\theta\left(\alpha;\alpha_0\right)=\frac{w\left(\alpha\right)}{D\left(r_p\right)T\left(\alpha;\alpha_0\right)}$$

Therefore we have that the densities $g\left(\cdot;\alpha,\alpha_0\right)$ and $b\left(\cdot;\alpha,\alpha_0\right)$ must satisfy the following differential equations:

$$\frac{\partial g\left(\omega;\alpha,\alpha_0\right)}{\partial\alpha}=-\theta\left(\alpha;\alpha_0\right)\mathbb{I}\left(\omega\leq\beta+\alpha\left(1-\beta\right)\right)g\left(\omega;\alpha,\alpha_0\right) \tag{A.24}$$

$$\frac{\partial b\left(\omega;\alpha,\alpha_0\right)}{\partial\alpha}=-\theta\left(\alpha;\alpha_0\right)\mathbb{I}\left(\omega\geq 1-\beta+\alpha\beta\right)b\left(\omega;\alpha,\alpha_0\right) \tag{A.25}$$

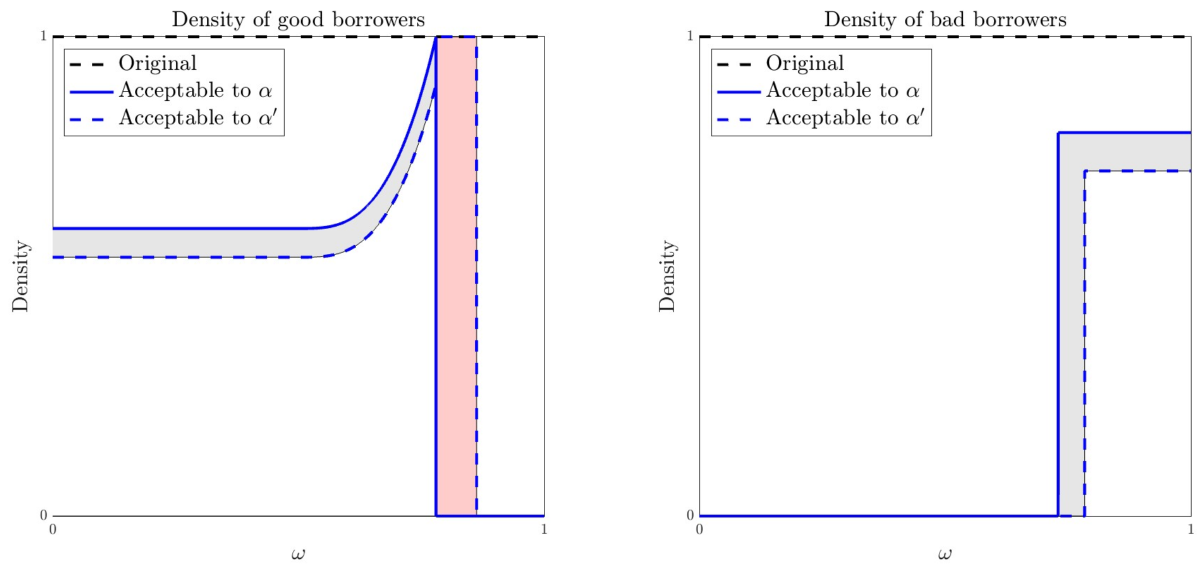

Figure A.1: How the density of acceptable borrowers changes from lender α to lender α'

Figure (A.1) shows a discretized example of how the density of acceptable borrowers evolves as the queue of lenders advances from lender α to lender α'. On the left panel are the good borrowers. The density of acceptable borrowers faced by α is shown in solid blue. It includes no borrowers to the right of $\omega = \beta + \alpha (1 - \beta)$; even though they are present in market r_p when it's lender $\alpha's$ turn, they are too opaque to be acceptable. To the left of this point, the density is lower than the original density due to the lenders that came before α. After α lends, the density falls proportionately, so the area shaded in gray is not available to lender α'. At the same time, borrowers with $\omega \in (\beta + \alpha (1 - \beta), \beta + \alpha' (1 - \beta)]$, who were not acceptable to lender α but are acceptable to lender α' enter the acceptable pool (this is the area shaded in red). The right panel shows the bad borrowers. Here the acceptable borrowers for α are those to the right of $\omega = 1 - \beta + \alpha\beta$ who have not borrowed yet. This whole density falls proportionately, and in addition the whole region $\omega = (1 - \beta + \alpha\beta, 1 - \beta + \alpha'\beta]$ goes away as those borrowers are unacceptable to lender α'. These two areas are shaded gray.

Equations (A.19), together with (A.24) and (A.25) and initial condition $g(\omega; \alpha_0, \alpha_0) = g(\omega)$ and $b(\omega; \alpha_0, \alpha_0) = b(\omega)$ fully define the functions $g(\cdot; \alpha, \alpha_0)$ and $b(\cdot; \alpha, \alpha_0)$ and therefore also define $G(\alpha; \alpha_0)$, $B(\alpha; \alpha_0)$, $T(\alpha; \alpha_0)$ and $\gamma(\alpha; \alpha_0)$. We show below that the function $T(\alpha; \alpha_0)$ is increasing in α_0: the later in α-space that lending begins, the higher the number of remaining borrowers faced by a given lender α. Define $\alpha_1(\alpha_0)$ as the value of α that solves:

$$T(\alpha; \alpha_0) = 0$$

(if a solution exists). By definition, lender α_1 finds a zero measure of acceptable borrowers in market r_p: all good borrowers with $\omega \in [0, \beta + \alpha_1 (1 - \beta))$ and all bad borrowers with $\omega \in (1 - \beta + \alpha_1\beta, 1]$ have been served by lenders with $\alpha \in [\alpha_0, \alpha_1)$. We show in the Appendix that $\alpha_1(\alpha_0)$ is an increasing function: a higher starting point for the pooling region implies a higher ending point as well.

Point α_1 is the boundary of Region I and Region II. The fact that $T(\alpha_1;\alpha_0)=0$ means that from α_1 onwards, lenders only lend to borrowers were unacceptable to the previous lenders. In terms of Figure 2, this means their pool of lenders consists only of vertical slices like the area shaded in red. The interest rate at which they lend is a cash-in-the-market rate: just high enough to equate supply and demand. The supply of loans by a neighborhood of lenders $[\alpha,\alpha+d\alpha]$ is $w(\alpha)\,d\alpha$, while the demand, coming from good borrowers with $\omega\in[\beta+\alpha(1-\beta),\beta+(\alpha+d\alpha)(1-\beta)]$, is $D(r(\alpha))\,g(\beta+\alpha(1-\beta))(1-\beta)\,d\alpha$. Equating supply and demand we have:

$$w(\alpha)=D(r(\alpha))\,g(\beta+\alpha(1-\beta))(1-\beta) \tag{A.26}$$

Therefore, lender α lends in the market defined by the interest rate:

$$r(\alpha)=D^{-1}\left(\frac{w(\alpha)}{g(\beta+\alpha(1-\beta))(1-\beta)}\right) \tag{A.27}$$

For simplicity, we assume that (A.27) defines an increasing function. Otherwise, it's easy to extend the equilibrium construction with ironing. In this cash-in-the-market segment, each good borrower is served in lowest-rate market where there are lenders who can identify them as good borrowers. Conversely, each lender chooses the highest-interest market where they can detect good borrowers.

The final possibility is that there is a Region III. This happens when for some α we have $r(\alpha)$ sufficiently high and the pool of remaining borrowers sufficiently good that some lenders can obtain profits Π by lending non-selectively. The average quality that a non-selective lender would get in market $r(\alpha)$ is:

$$\gamma^{NS}(\alpha)=\frac{G(1)-G(\beta+\alpha(1-\beta))}{G(1)-G(\beta+\alpha(1-\beta))+L(\alpha_0,\alpha_1)} \tag{A.28}$$

Here the quantity $G(1)-G(\beta+\alpha(1-\beta))$ is the total mass of good borrowers with $\omega>\beta+\alpha(1-\beta)$, none of whom have been served yet because they are too opaque for the lenders before α. While

$$L(\alpha_0,\alpha_1)\equiv\int_0^1 b(\omega;\alpha_1,\alpha_0)\,d\omega$$

is the total mass of bad borrowers who were not served by lenders $\alpha\in[\alpha_0,\alpha_1]$ in the pooling market r_p. We will refer to $L(\alpha_0,\alpha_1)$ as the leftover bad borrowers. If, for any $\alpha\in(\alpha_1,1)$ we have that

$$\gamma^{NS}(\alpha)(1+r(\alpha))-1>\Pi \tag{A.29}$$

then the low-α lenders with $\alpha\in[0,\alpha_0)$ will find it profitable to enter. Define α_2 as the solution (with equality) to (A.29), if such a solution exists. In this case, α_2 defines the boundary between Region II and Region III. In Region III, non-selective lenders will want to lend a total of $\int_0^{\alpha_0} w(\alpha)\,d\alpha$, pro-rated among all remaining borrowers, so that a fraction

$$\theta^{NS}=\frac{\int_0^{\alpha_0} w(\alpha)\,d\alpha}{D(r(\alpha_2))\left[G(1)-G(\beta+\alpha_2(1-\beta))+\int_0^1 b(\omega;\alpha_1,\alpha_0)\,d\omega\right]} \tag{A.30}$$

will be served by non-selective lenders. The remaining bad borrowers will be left unserved, and the remaining good borrowers will be served by the lenders with $\alpha > \alpha_2$. Therefore, we must have that:

$$\frac{\int_{\alpha_2}^{1} w(\alpha)\, d\alpha}{D(r(\alpha_2))\,[G(1) - G(\beta + \alpha_2(1-\beta))]} = 1 - \theta^{NS} \tag{A.31}$$

Equation (A.31) says that the total wealth of the highest-α lenders is exactly enough to satisfy the demand of the fraction $1 - \theta^{NS}$ of opaque good borrowers who were not served by non-selective lenders.

To complete the construction of equilibrium, we must find the correct values of α_0 and Π. Given Π, we find α_0 with a smooth pasting condition. Use (A.20)-(A.22) to compute:

$$\frac{\partial T(\alpha;\alpha_0)}{\alpha} = -\frac{w(\alpha)}{D(r_p(\alpha_0))} + (1-\beta)\, g(\beta + \alpha(1-\beta)) - \beta b(1 - \beta + \alpha\beta; \alpha; \alpha_0) \tag{A.32}$$

Evaluating this expression at α_1, we have

$$\left|\frac{\partial T(\alpha;\alpha_0)}{\alpha}\right|_{\alpha=\alpha_1} = -\frac{w(\alpha_1)}{D(r_p(\alpha_0))} + (1-\beta)\, g(\beta + \alpha_1(1-\beta)) \tag{A.33}$$

(the last term in (A.32) vanishes because $T(\alpha_1;\alpha_0) = 0$ implies $b(1-\beta+\alpha_1\beta;\alpha_1;\alpha_0)$: if there are no remaining acceptable borrowers, it must mean that the density of acceptable bad borrowers is also zero). If it were the case that $\left|\frac{\partial T(\alpha;\alpha_0)}{\alpha}\right|_{\alpha=\alpha_1} < 0$, then (A.27) and (A.33) imply a downward discontinuity in the interest rate schedule, with the cash-in-the-market rate below the pooling rate. This cannot be part of an equilibrium, because the higher-α lenders would prefer to lend in the pooling market. Hence in equilibrium we must have:

$$\left|\frac{\partial T(\alpha;\alpha_0)}{\alpha}\right|_{\alpha=\alpha_1} = 0 \tag{A.34}$$

Equation (A.34) guarantees continuity in the transition between Region I and Region II. We find the equilibrium by finding the value of α_0 that ensures that the last lender in the pooling market lends to no more than exactly a vertical slice like the red shaded ares in Figure (2).

Finally, we find the correct value of Π by ensuring that the market-clearing condition (A.31) holds. There are two possibilities: either $\Pi > 0$, (A.30) holds with equality and all lenders lend their entire wealth; or $\Pi = 0$ and (A.30) is slack: low-α lenders are indifferent between lending or not, and just enough of them lend so that all good borrowers end up being served. In case equation(A.29) does not't't have a solution, Region III does not exist and we must have $\Pi = 0$. This completes the construction of the equilibrium allocation.

Given this construction, a hockey-stick stick equilibrium where lender distribution is exogenous is defined by the following set of equations:

$$\begin{aligned}
G(\alpha;\alpha_0) &= \int_0^{\beta+\alpha(1-\beta)} g\left(\omega;\alpha\right) d\omega \\
B(\alpha;\alpha_0) &= \int_{1-\beta+\alpha\beta}^{1} b\left(\omega;\alpha\right) d\omega \\
\frac{\partial g(\omega;\alpha)}{\partial\alpha} &= -\theta\left(\alpha\right)\mathbb{I}\left(\omega \leq \beta+\alpha\left(1-\beta\right)\right) g\left(\omega;\alpha\right) \\
\frac{\partial b(\omega;\alpha)}{\partial\alpha} &= -\theta\left(\alpha\right)\mathbb{I}\left(\omega \geq 1-\beta+\alpha\beta\right) b\left(\omega;\alpha\right) \\
\theta(\alpha) &= \frac{w\left(\alpha\right)}{D\left(r_p\right)T\left(\alpha\right)} \\
T(\alpha;\alpha_0) &= B(\alpha;\alpha_0)+G(\alpha;\alpha_0) \\
T_1(\alpha_1;\alpha_0) &= -\frac{w\left(\alpha_1\right)}{D\left(r_p\right)} + \left(1-\beta\right) g\left(\beta+\alpha_1\left(1-\beta\right)\right) \\
\gamma(\alpha) &= \frac{G(\alpha)}{T(\alpha)} \\
\gamma_0(\alpha) &= \frac{G\left(\left[0,\beta+\alpha\left(1-\beta\right)\right]\right)}{G\left(\left[0,\beta+\alpha\left(1-\beta\right)\right]\right)+B\left(\left[1-\beta+\alpha\beta,1\right]\right)} \\
\gamma^{NS}(\alpha) &= \frac{\int_{\beta+\alpha(1-\beta)}^{1} g(\omega)d\omega}{\int_{\beta+\alpha(1-\beta)}^{1} g(\omega)d\omega + \int_{1-\beta+\alpha_0\beta}^{1-\beta+\alpha_1\beta} b(\omega,\alpha_1)d\omega} \\
w(\alpha) &= D(r_{CIM}(\alpha))g(\beta+\alpha(1-\beta))(1-\beta) \qquad\qquad \alpha\in[\alpha_1,\alpha_2]
\end{aligned}$$

Furthermore,

1. α_0 lender, i.e. the lowest entrant with positive level of expertise: makes profit Π.
2. $\alpha = 0$ lenders, i.e. those who lend in the non-selective region: make profit Π.

Thus, the equilibrium is defined by the following system of 7 equations- 7 unknowns.

$$\begin{aligned}
&\gamma_0(\alpha_0)(1+r_p) = 1 \\
&T(\alpha_1;\alpha_0) = 0 \\
&T_1(\alpha_1;\alpha_0) = 0 \\
&\gamma^{NS}(\alpha_2)\left(1+r_{CIM}(\alpha_2)\right) = 1 \\
&\omega_i = \beta+\alpha_i(1-\beta) \qquad i=0,1,2
\end{aligned}$$

The first three equations need to be solved together. The next four equations can be solved one-by-one.

We next provide a constructive proof for the equilibrium.

Fix the value of β. We provide a constructive proof for the equilibrium with interest rate schedule in Proposition 5 where every investor makes profit at least $\Pi \geq 0$. If $\Pi = 0$ then the lowest skilled entrants make zero profits.

Consider a market where lender with precision α participate in. Let $g(\omega;\alpha)/b(\omega;\alpha)$ denote the mass of good/bad borrowers with opacity ω who are present in that market, are acceptable to lender α, and are not cleared by lenders who choose before lender α.

Region I Following Proposition 5, in this region borrowers with opacity $\omega \in [0, \omega_1]$ borrow from lenders with precision $\alpha \in [\alpha_0, \alpha_1]$ at common interest rate r_p, and $\omega_1 = \beta + (1-\beta)\alpha_1$. We next characterize α_0, α_1, and r_p.

Mass of good, bad, and total borrowers acceptable to lender with precision α are given by Equations (**??**), (**??**) and (**??**), respectively. Furthermore, Equation (**??**) defines the quality of borrowers for lender with precision α and Equation (14) defines the rate of depletion of borrowers after lender α lends to his chosen portfolio. For a given interest rate r_p, in Region I the pdfs evolve according to Equations (**??**) and (**??**), where r_p determines $\theta(\alpha)$.

Take the derivative with respect to α to compute $G'(\alpha)$, $B'(\alpha)$ and $T'(\alpha)$:

$$\begin{aligned}
G'(\alpha) &= \int_0^{\beta+\alpha(1-\beta)} \frac{\partial g(\omega;\alpha)}{\partial \alpha} d\omega + (1-\beta)\, g(\beta+\alpha(1-\beta);\alpha) \\
&= -\int_0^{\beta+\alpha(1-\beta)} \theta(\alpha)\, g(\omega;\alpha)\, d\omega + (1-\beta)\, g(\beta+\alpha(1-\beta);\alpha) \\
&= -\frac{w(\alpha)}{D(r_p)\, T(\alpha)} G(\alpha) + (1-\beta)\, g(\beta+\alpha(1-\beta);\alpha) \\
&= -\frac{w(\alpha)}{D(r_p)} \gamma(\alpha) + (1-\beta)\, g(\beta+\alpha(1-\beta);\alpha) \\
&= -\frac{w(\alpha)}{D(r_p)} \gamma(\alpha) + (1-\beta)\, g(\beta+\alpha(1-\beta))
\end{aligned}$$

Where $g(\beta+\alpha(1-\beta))$ to denote the original density. Alternatively

$$\begin{aligned}
B'(\alpha) &= \int_{1-\beta+\alpha\beta}^{1} \frac{\partial b(\omega;\alpha)}{\partial \alpha} d\omega - \beta b(1-\beta+\alpha\beta;\alpha) \\
&= -\int_{1-\beta+\alpha\beta}^{1} \theta(\alpha)\, b(\omega;\alpha)\, d\omega - \beta b(1-\beta+\alpha\beta;\alpha) \\
&= -\frac{w(\alpha)}{D(r_p)\, T(\alpha)} B(\alpha) - \beta b(1-\beta+\alpha\beta;\alpha) \\
&= -\frac{w(\alpha)}{D(r_p)} (1-\gamma(\alpha)) - \beta b(1-\beta+\alpha\beta;\alpha)
\end{aligned}$$

Notice that $b(1-\beta+\alpha\beta;\alpha)$ in the last step is **not** the original density. Finally, add up to get $T'(\alpha)$:

$$T'(\alpha) = -\frac{w(\alpha)}{D(r_p)} \gamma(\alpha) + (1-\beta)\, g(\beta+\alpha(1-\beta)) - \frac{w(\alpha)}{D(r_p)} (1-\gamma(\alpha)) - \beta b(1-\beta+\alpha\beta;\alpha)$$

$$T'(\alpha) = -\frac{w(\alpha)}{D(r_p)} + (1-\beta)\, g(\beta+\alpha(1-\beta)) - \beta b(1-\beta+\alpha\beta;\alpha) \tag{A.35}$$

Note that that $T'(\alpha)$ is **not** guaranteed to be negative. The reason is that an improvement in precision α adds some good borrowers who are identified only by lender α, but the mass of good borrowers who were identified by $\alpha' < \alpha$ decreases, as well as the mass of bad borrowers who are now identified as bad as not funded by α.

We now determine the range of lender precision who participate in **Region I** market, $[\alpha_0, \alpha_1]$. We first characterize $\alpha_1(\alpha_0))$ for each α_0, and then show that only a unique pair $(\alpha_0, \alpha_1(\alpha_0)$ can arise in equilibrium.

Characterization of $\alpha_1(\alpha_0)$ As α_0 is the lender with the lowest strictly positive precision level who lends in this market, interest rate r_p is determined such that it makes lender α_0 indifferent between entering or staying out. Recall that we are assuming lenders make profit Π by staying out. The indifference condition of α_0 lender determines the pooling interest rate in Equation (13).

Recall that $\gamma_0(\alpha_0)$, the quality of borrowers acceptable to lenders α_0 is given by Equation (11). Since γ_0 is increasing in α_0, then r_p is decreasing in α_0.

Define $T(\alpha, \alpha_0)$ as is the solution to the differential equation (A.35) with initial condition α_0. $T(\alpha, \alpha_0)$ is the answer to the question: "if the lowest entrant to the pooling region is α_0, what is the mass of acceptable borrowers to lender α in the pooling region?"

For a given α_0, let $\alpha_1(\alpha_0)$ denote the lowest solution to

$$T(\alpha, \alpha_0) = 0$$

if there is a solution. α_1 is the lowest precision in the pooling region who reaches zero-acceptable-supply if the first entrant to the pooling region is α_0, $T(\alpha_1(\alpha_0), \alpha_0) = 0$.

Existence of (α_0, α_1) We show three statements: 1) if α_0^* is the lowest precision of any lender in **Region I** (pooling region), then there is a unique α_1 for which $T(\alpha_1, \alpha_0^*) = 0$, 2) there is a unique α_0 for which an α_1 that satisfies equilibrium conditions exist, 3) a pair (α_0, α_1) exists.

We start by proving a series of intermediate claims.

Claim 1. *$T(\alpha, \alpha_0)$ is increasing in α_0, $\frac{\partial T(\alpha,\alpha_0)}{\partial \alpha_0} > 0\ \forall \alpha$.*

Proof. Let $\gamma(\alpha; \alpha_0)$ denote the quality of borrowers faced by lender with precision α if the lowest precision in the pooling region is α_0. Consider 3 values of α, $\alpha_0^l < \alpha_0^m \leq \alpha_h$, such that $\gamma(\alpha_0^m; \alpha_0^l) < \gamma(\alpha_h; \alpha_0^l) < 1$, i.e. if the pooling region starts at α_0^l, both α_0^m and α_h are in the pooling region.

The goal is to show that for all such $\alpha_0^m \leq \alpha_h$, $T(\alpha_h, \alpha_0^m) > T(\alpha_h, \alpha_0^l)$, where $T(\alpha_h, \alpha_0)$ denotes the pool of borrowers available to lender α_h if α_0 is the first (lowest precision) lender who lends in **Region I**. Consider moving α_0 (start of pooling region) from α_0^l to α_0^m.

This exercise impacts the $T(.)$ function through two channels: 1) decrease in interest rate r_p, which increases the demand of every borrower in the pooling region, 2) decrease in the set of lenders available to absorb the demand of borrowers, which amplifies the increase in set of remaining borrowers. In the proof below, we address the two channels concurrently. However, they can be separately considered as well.

We use a proof by induction on α_h using Equations (**??**), (**??**), (**??**), (14), (**??**) and (**??**).

Base step The base step is for $\alpha_h = \alpha_0^m$.[6] As $g(\omega;\alpha) \leq g(\omega)$ and $b(\omega;\alpha) \leq b(\omega)$, $\forall\ (\alpha,\omega)$, starting the pooling region at α_0^m to the right of α_0^l increases $G(\alpha_0^m), B(\alpha_0^m), T(\alpha_0^m)$ from Equations (**??**), (**??**), (**??**). Thus, $T(\alpha_0^m;\alpha_0^m) > T(\alpha_0^m;\alpha_0^l)$.

Furthermore, from Equation (13) and $\frac{d\gamma_0(\alpha)}{d\alpha} > 0$, the interest rate r_p decreases when α_0 increases from α_0^l to α_0^m, $r_p^m < r_p^l$. Thus, $D(r_p^m) > D(r_p^l)$. Using Equation (14), these two observations imply that $\theta(\alpha_0^m;\alpha_0^m) < \theta(\alpha_0^m;\alpha_0^l)$. Thus, from Equations (**??**) and (**??**), the rate of change of both $g(\omega;\alpha)$ and $b(\omega;\alpha)$ are smaller at $\alpha = \alpha_0^m$, for every ω, when the start of the pooling region moves to the right, from α_0^l to α_0^m. In other words, α_0^m lending leads to a lower rate of decline in the mass of both good and bad borrowers that he lends to. The starting levels of both good and bad borrowers at every ω, before α_0^m lends, are higher as all the lenders with precision $\alpha \in [\alpha_0^l, \alpha_0^m)$ used to absorb some demand but now they are not lending, and each borrower is demanding more because the interest rate is lower. Higher starting level and lower rate of decline imply that the final level after α_0^m has lent is higher for both bad and good borrowers that he has lent to and unchanged to those who he has not. Using Equations (**??**), (**??**) and (**??**) at $\alpha = \alpha_0^m + d\alpha$, $\lim_{\alpha\to\alpha_0^m} T(\alpha;\alpha_0^m) > \lim_{\alpha\to\alpha_0^m} T(\alpha;\alpha_0^l)$, where α converges to α_0^m from above.

Inductive step Assume $T(\alpha_h;\alpha_0^m) > T(\alpha_h;\alpha_0^l)$ for $\alpha_h > \alpha_0^m$. Show that $T(\alpha_h + d\alpha;\alpha_0^m) > T(\alpha_h + d\alpha;\alpha_0^l)$ when α_0 increases from α_0^l to α_0^m.[7]

The argument is exactly the same as the base step. First, $r_p^m < r_p^l$ and thus $D(r_p^m) > D(r_p^l)$. Second, by be inductive assumption, $T(\alpha_h;\alpha_0^m) > T(\alpha_h;\alpha_0^l)$. Thus, using Equation (14), $\theta(\alpha_h;\alpha_0^m) < \theta(\alpha_h;\alpha_0^l)$. From Equations (**??**) and (**??**), the rate of change of both $g(\omega;\alpha)$ and $b(\omega;\alpha)$ are smaller at $\alpha = \alpha_h$ when α_0 increases from α_0^l to α_0^m, for every ω that α_h lends to; and is zero otherwise. Furthermore,

$$T(\alpha_h + d\alpha) = \lim_{\alpha\to\alpha_h^+} T(\alpha_h) = T(\alpha_h) + \int_0^1 \frac{\partial g(\omega;\alpha)}{\partial\alpha}|_{\alpha=\alpha_h} d\omega + \int_0^1 \frac{\partial b(\omega;\alpha)}{\partial\alpha}|_{\alpha=\alpha_h} d\omega,$$

i.e. the only difference between $T(\alpha_h;\alpha_0^m)$ and $T(\alpha_h + d\alpha;\alpha_0^m)$ comes from the lending of lenders with precision α_h, as a rate of change from base of $T(\alpha_h;\alpha_0^m)$, which is higher as by assumption of inductive step, $T(\alpha_h;\alpha_0^m) > T(\alpha_h;\alpha_0^l)$. Again, higher base and lower rate of decline imply that the final level is higher, i.e. $T(\alpha_0^h + \delta h;\alpha_0^m) > T(\alpha_0^h + \delta_h;\alpha_0^l)$, which completes the proof. □

Intuitively, consider α_0^l and two larger values of α smaller than α_1, $\alpha_0^l < \alpha_0^m < \alpha_h < \alpha_1$. Moving the first lender in the pooling region from α_0^l to α_0^m has two impacts on the pool of borrowers available to lender α_h. Both effects go in the same direction and enlarge the pool.

First, the interest rate in the pooling region is determined by Equation (13). As $\gamma_0(.)$ is increasing in the precision of the first lender in the pooling region, r_p is decreasing in it, thus $r_p(\alpha_0^m) < r_p(\alpha_0^l)$. Furthermore, demand of every borrower is decreasing in the interest rate he faces. As such, if the pooling region starts with a lender with higher precision, the

[6] I am actually extending the base step to be for $\alpha_h \downarrow \alpha_0^m$ to be sure.

[7] Or in limit notation, $\lim_{\alpha\to\alpha_h^+} T(\alpha;\alpha_0^m) > \lim_{\alpha\to\alpha_h^+} T(\alpha;\alpha_0^l)$.

pooling interest rate is lower and the demand of every borrower is higher, which pushes up the acceptable demand by α_h.

Second, if we compare the pooling region that starts from precision α_0 to the one that starts with α_m, there are lenders with precision $[\alpha_0^l, \alpha_0^m)$ who used to lend in the former pooling region and clear out some of the demand but do not lend in the latter. In other words, fewer lenders lend before α_h, which also pushes up acceptable demand by α_h.

Claim 2. $\frac{\partial T(\alpha,\alpha_0)}{\partial \alpha}|_{\alpha=\alpha_1} = 0$.

Proof. To show that $T'(\alpha_1) = 0$, we first show that $T'(\alpha_1) \leq 0$. The reason is that $T(\alpha_0, \alpha_0)$ is positive, and α_1 is defined as the lowest solution to $T(\alpha, \alpha_0) = 0$, thus, at $\alpha = \alpha_1$, $T(\alpha, \alpha_0)$ must approach zero from above, $T'(\alpha_1) \leq 0$. Note that in general, $T'(\alpha)$ can be positive, as we show in Claim 3.

Taking derivative of $T(\alpha)$:

$$T'(\alpha) = -\frac{w(\alpha)}{D(r_p)} + (1-\beta)\, g(\beta + \alpha(1-\beta)) - \beta b(1-\beta+\alpha\beta;\alpha)$$

As $T \to 0$ when $\alpha \to \alpha_1$, $b(1-\beta+\alpha\beta;\alpha) \to 0$ at the same time. Thus the above simplifies to

$$T'(\alpha_1) = -\frac{w(\alpha_1)}{D(r_p)} + (1-\beta)\, g(\beta + \alpha_1(1-\beta)).$$

To prove $T'(\alpha_1) = 0$, we need to show that $T'(\alpha_1) \nless 0$. We prove this by contradiction. Suppose that $T'(\alpha_1) < 0$. Economically, this means that lender α_1 lends to two groups of borrowers: 1) good borrowers with $\omega_1 = \beta + (1-\beta)\alpha_1$, 2) (a vanishing share of) all the good borrowers with $0 \leq \omega < \omega_1$. This would imply We have:

$$\begin{aligned}
& -\frac{w(\alpha_1)}{D(r_p)} + (1-\beta)\, g(\beta + \alpha_1(1-\beta)) < 0 \\
& \frac{w(\alpha_1)}{D(r_p)} > (1-\beta)\, g(\beta + \alpha_1(1-\beta)) \\
& D(r_p) < \frac{w(\alpha_1)}{(1-\beta)\, g(\beta + \alpha_1(1-\beta))} \\
& D(r_p)(1-\beta)\, g(\beta + \alpha_1(1-\beta)) < w(\alpha_1)
\end{aligned}$$

Recall that $w(.)$ is continuous. Thus, the last expression implies that in the first cash-in-the-market pricing region at the switch from pooling at α_1, if the interest rate is pooling interest rate r_p, there well be an excess supply of capital. As such, the equilibrium interest rate in this market will be lower than r_p, a contradiction. It follows that $T'(\alpha_1) = 0$ while $T'(\alpha_) < 0$, thus $T'(\alpha_1) = 0$. □

Claim 3. $\frac{\partial^2 T(\alpha,\alpha_0)}{\partial^2 \alpha} > 0\ \forall \alpha$.

Proof. Compute the second derivative of $T(\alpha)$

$$T''(\alpha) = -\underbrace{\underbrace{\frac{w'(\alpha)}{D(r_p)}}_{(-)}}_{(+)} + (1-\beta)\underbrace{\frac{dg_0(\beta+\alpha(1-\beta))}{d\alpha}}_{=0} \underbrace{-\beta\underbrace{\frac{db(1-\beta+\alpha\beta;\alpha)}{d\alpha}}_{(-)}}_{(+)}.$$

Using Assumption **??**, $\frac{dw(\alpha)}{d\alpha} < 0$ and $g(\omega)$ and $b(\omega)$ are uniform. Thus, $T''(\alpha) > 0$, i.e. $T(\alpha)$ is convex. □

To show that for the equilibrium α_0^*, α_1 is unique, if it exists, Claim 2 shows that the first derivative is zero at α_1, $\frac{\partial T(\alpha,\alpha_0)}{\partial\alpha}|_{\alpha=\alpha_1} = 0$. As Claim 3 shows that $T(\alpha, \alpha_0^*)$ is globally convex in α, α_1 is unique.

Next, in order to show that α_0 is unique, let α_0^* denote the value of α_0 such that $T(\alpha,\alpha_0) = 0$ and $\frac{\partial T(\alpha,\alpha_0)}{\partial\alpha} = 0$ hold for the same α (we have already shown that it can hold for at most one value of α). Since T is convex in α, $T(\alpha,\alpha_0) \geq 0$ for all α. For any $\tilde{\alpha}_0 > \alpha_0^*$, the function $T(\alpha,\tilde{\alpha}_0)$ is strictly higher than a non-negative function, so $T(\alpha,\tilde{\alpha}_0) > 0$ for all α and we do not have an equilibrium. Alternatively, for any $\tilde{\alpha}_0 < \alpha_0^*$, the function $T(\alpha,\tilde{\alpha}_0)$ is strictly lower from the function $T(\alpha,\alpha_0^*)$ thus it can never cross it, and furthermore it is globally convex. It follows that $T(\alpha,\tilde{\alpha}_0)$ have to cross $\alpha = 0$ twice, once at $\alpha_L < \alpha_1$ and once at $\alpha_H > \alpha_1$. Thus, $T'(\alpha_L) < 0$ and $T'(\alpha_H) > 0$ and neither can be an equilibrium. It follows that $\tilde{\alpha}_0 < \alpha_0^*$ cannot be an equilibrium either and α_0^* is the unique equilibrium.

To complete the construction of **Region I** equilibrium, we have to show that a pair (α_0,α_1) exist.

For $\beta = 1$, $\alpha_1 = 1$. For $\beta = 0$, I think $\alpha_0 = 0$. Are these corner cases? I do think $\alpha_1 = 1$ in $\beta = 1$ is a corner where $T(\alpha_1) = T'(\alpha_1) = 0$, but I am not 100% sure. I wonder if for the proof it is sufficient to say that Farboodi and Kondor (2022) show that an equilibrium exists for $\beta = 0, 1$ and equilibrium is continuous in β.

This completes the characterization of **Region I.**

Finally, Claim 2 implies that the interest rate is continuous at α_1, the switch from **Region I** (pooling interest rate) to **Region II** (cash-in-the-market interest rate).

Region II When $\alpha > \alpha_1$, we enter the region where there is a continuum of markets each with a cash-in-the-market pricing equilibrium. Each market is served by lenders of a single precision α, with $\gamma(\alpha) = 1$, who lends only to good borrowers with $\omega = \beta + \alpha(1-\beta)$. The interest rate in each market is determined such that the market clears:

$$dW(\alpha) = w(\alpha)d\alpha = D(r(\alpha))\,g(\beta+\alpha(1-\beta))(1-\beta)d\omega.$$

$$r(\alpha) = D^{-1}\left(\frac{w(\alpha)}{g(\beta+\alpha(1-\beta))(1-\beta)}\right) \tag{A.36}$$

Region III For each $\alpha > \alpha_1$ define

$$\gamma^{NS}(\alpha) = \frac{\int_{\beta+\alpha(1-\beta)}^{1} g(\omega)d\omega}{\int_{\beta+\alpha(1-\beta)}^{1} g(\omega)d\omega + \int_0^{1-\beta+\alpha_1\beta} b(\omega,\alpha_1)d\omega}$$

The numerator is all the good borrowers who are not served (cleared) by lenders with precision $\alpha' < \alpha$ and are willing to borrow at interest rates $r \geq r(\alpha)$. The second term in the denominator is all the bad borrowers who are willing to borrow at such rates, i.e. bad borrowers who are sufficiently transparent $\omega < \omega_1$ who are not cleared by the lenders who lend in the pooling region. As such, $\gamma^{NS}(\alpha)$ is the quality of borrowers that a lender with precision zero will get if he tries to lend in the market where lenders with precision α lend. $\gamma^{NS}(\alpha)$ represents the selection that a no-precision lender receives if he lends in the cash-in-the-market pricing where lender α lends.

Lender $(\beta, 0)$ has to receive profit Π for lending, thus we have

$$\gamma^{NS}(\alpha)\left(1 + r_{NS}(\alpha)\right) = 1 + \Pi$$
$$r_{NS}(\alpha) = \frac{1+\Pi}{\gamma^{NS}(\alpha)}.$$

As the interest rate in the cash-in the-market region is strictly increasing. Let α_2 denote the lowest solution to

$$r_{NS}(\alpha) = r(\alpha)$$
$$\frac{(1+\Pi)\left(\int_{\beta+\alpha(1-\beta)}^{1} g(\omega)d\omega + \int_{0}^{1-\beta+\alpha_1\beta} b(\omega,\alpha_1)d\omega\right)}{\int_{\beta+\alpha(1-\beta)}^{1} g(\omega)d\omega} = D^{-1}\left(\frac{w(\alpha)}{g(\beta+\alpha(1-\beta))(1-\beta)}\right).$$

At $\alpha = \alpha_2$, the equilibrium interest rate schedule switches from **Region II** (cash-in-the-market pricing) to **Region III** (non-selective) pricing.

Any lenders with $\alpha > \alpha_2$ will also lend at the same price as all good borrowers ca be served by $\alpha = 0$ non-selective lenders at $r_N S(\alpha_2)$ and will not accept any higher interest rate. Using the equilibrium definition in Definition 4, lenders with $\alpha = 0$ are serves first in this market. On the other hand, lenders with $\alpha > \alpha_2$, have total wealth of $\int_{\alpha_2}^{1} w(\alpha)d\alpha$ and only lender to good borrowers. Each borrowers demands $D(r_{NS}(\alpha_2))$ in **Region III**. As such, the measure of good borrowers absorbed by the wealth of high precision lenders is given by

$$M_g = \frac{\int_{\alpha_2}^{1} w(\alpha)d\alpha}{D\left(r_{NS}(\alpha_2)\right)}.$$

The remaining good lenders have to be absorbed by the wealth of $\alpha = 0$ lenders who lend to portfolio quality $\gamma^{NS}(\alpha_2)$, which implies

$$w_{NS} = \frac{D\left(r_{NS}(\alpha_2)\right)\left(\int_{\beta+\alpha_2(1-\beta)}^{1} g(\omega)d\omega - M_g\right)}{\gamma^{NS}(\alpha_2)}.$$

A.4 IID signals

Proof of Proposition 1-iid.
We'll solve the equilibrium as a series of functions of α.

For each α, $G(\alpha)$ and $B(\alpha)$ denote the sizes of remaining pools when it's the turn of α and $r(\alpha)$ is interest rate in market that α visits.

Define $z(\alpha) = \frac{G(\alpha)}{G(\alpha)+B(\alpha)}$. The quality of the pool faced by α is:

$$\gamma(\alpha) = \frac{z(\alpha)[\beta + \alpha(1-\beta)]}{z(\alpha)[\beta + \alpha(1-\beta)] + (1 - z(\alpha))\beta(1-\alpha)}.$$

Similar to the nested case, define $\gamma(\alpha, \tilde{\alpha})$ as the pool faced by α in the market $\tilde{\alpha}$

$$\gamma(\alpha, \tilde{\alpha}) = \frac{z(\tilde{\alpha})[\beta + \alpha(1-\beta)]}{z(\tilde{\alpha})[\beta + \alpha(1-\beta)] + (1 - z(\tilde{\alpha}))\beta(1-\alpha)},$$

so, in this notation, $\gamma(\alpha) = \gamma(\alpha, \tilde{\alpha})$.

Let $\underline{\alpha}$ denote the lowest entrant. Thus, $z(\underline{\alpha}) = \frac{G}{G+B}$. Profits for lender α are given by

$$\Pi(\alpha, \tilde{\alpha}) = \gamma(\alpha, \tilde{\alpha})(1 + r(\tilde{\alpha})) - 1,$$

while evolution of G and B are given by

$$G'(\alpha) = -\frac{w(\alpha)}{D(r(\alpha))}\gamma(\alpha)$$
$$B'(\alpha) = -\frac{w(\alpha)}{D(r(\alpha))}[1 - \gamma(\alpha)]$$

The optimality condition for type α is

$$\alpha \in \arg\max_{\tilde{\alpha}} \gamma(\alpha, \tilde{\alpha})(1 + r(\tilde{\alpha})) - 1,$$

with FOC

$$\left.\frac{\partial \gamma(\alpha, \tilde{\alpha})}{\partial \tilde{\alpha}}\right|_{\tilde{\alpha}=\alpha} (1 + r(\alpha)) + r'(\alpha)\gamma(\alpha, \alpha) = 0$$

or, equivalently

$$r'(\alpha) = -\frac{\left.\frac{\partial \gamma(\alpha, \tilde{\alpha})}{\partial \tilde{\alpha}}\right|_{\tilde{\alpha}=\alpha}(1 + r(\alpha))}{\gamma(\alpha, \alpha)} \tag{A.37}$$

There are two terminal conditions. $\underline{\alpha}$ and $r(\underline{\alpha})$ have to be such that the following two conditions hold

1. The marginal lender is indifferent,

$$\gamma(\underline{\alpha}, \underline{\alpha}) = (1 + r(\underline{\alpha})) - 1.$$

2. Every good borrower is served

$$G(1) = 0.$$

□

Proof of Proposition 2-iid.

Construct the equilibrium as follows:

1. Fix Π.

2. Let

$$z_0 \equiv \frac{G}{G+B}$$
$$\gamma_0(\alpha) \equiv \frac{z_0[\beta+\alpha(1-\beta)]}{z_0[\beta+\alpha(1-\beta)]+(1-z_0)\beta(1-\alpha)}$$

Find the lender who can lend cheapest and still make Π

$$\alpha_0 = \arg\min \frac{\Pi+1+C(\alpha)}{\gamma_0(\alpha)}.$$

3. Lender α_0 goes to market

$$r(\alpha_0) = \frac{\Pi+1+C(\alpha)}{\gamma_0(\alpha)} - 1.$$

4. Find W by discretizing.

In order to do that, we will solve a series of systems of 2 equations-2 unknowns.

Fix Δ. For each n, suppose that W has a mass point $w(\alpha_n)$ at α_n. Think of one lender with mass $w(\alpha_n)$. Then when we take the limit this should not matter.

At the next market $n+1$, $\alpha_{n+1} = \alpha_n + \Delta$, the quality of the pool $z(\alpha_{n+1})$ is given by:

$$G(\alpha_n+\Delta) = G(\alpha_n) - \frac{w(\alpha_n)}{D(r(\alpha_n))}\gamma(\alpha_n)$$
$$B(\alpha_n+\Delta) = B(\alpha_n) - \frac{w(\alpha_n)}{D(r(\alpha_n))}[1-\gamma(\alpha_n)]$$

so

$$\begin{aligned} z(\alpha_{n+1}) &= \frac{G(\alpha_n+\Delta)}{G(\alpha_n+\Delta)+B(\alpha_n+\Delta)} \\ &= \frac{G(\alpha_n) - \frac{w(\alpha_n)}{D(r(\alpha_n))}\gamma(\alpha_n)}{G(\alpha_n) - \frac{w(\alpha_n)}{D(r(\alpha_n))}\gamma(\alpha_n) + B(\alpha_n) - \frac{w(\alpha_n)}{D(r(\alpha_n))}[1-\gamma(\alpha)]} \\ &= \frac{G(\alpha_n) - \frac{w(\alpha_n)}{D(r(\alpha_n))}\gamma(\alpha_n)}{G(\alpha_n) + B(\alpha_n) - \frac{w(\alpha_n)}{D(r(\alpha_n))}} \end{aligned}$$

We need $w(\alpha_n)$ and $r(\alpha_{n+1})$ to be such that lenders are indifferent between choosing α_n and α_{n+1} and constitute a 2 eq-2 unknown system of equations. To ensure this, let

$$\gamma(\alpha, \alpha_{n+1}) = \frac{z(\alpha_{n+1})[\beta + \alpha(1-\beta)]}{z(\alpha_{n+1})[\beta + \alpha(1-\beta)] + (1 - z(\alpha_{n+1}))\beta(1-\alpha)},$$

which is the quality that a lender gets if he has skill α and chooses market α_{n+1}). Indifference requires:

$$\gamma(\alpha_{n+1}, \alpha_{n+1})(1 + r(\alpha_{n+1})) - 1 - C(\alpha_{n+1}) = \Pi. \quad \text{(A.38)}$$

Furthermore, lenders who already chose skill α_n are indifferent between visiting market α_n and visiting market α_{n+1}. Indifference requires:

$$\gamma(\alpha_n, \alpha_{n+1})(1 + r(\alpha_{n+1})) - 1 - C(\alpha_n) = \Pi \quad \text{(A.39)}$$

Equations (A.38) and (A.39) are a system of two equations in two unknowns: $w(\alpha_n)$ and $r(\alpha_{n+1})$.

The interpretation is that we need to find $r(\alpha_{n+1})$ that is high enough that it compensates for the higher cost of the skill level α_{n+1}. However, if w_n is too low, α_n lenders will find market α_{n+1} attractive: you can go to it without acquiring extra skill, the selection is almost the same and the interest rate is better. So we need w_n to be high enough to deter α_n types from deviating to market α_{n+1} by worsening the pool. In turn, this will require an even higher rate to keep α_{n+1} types indifferent, so you must keep doing this until you find a fixed point, where conditions in market α_{n+1} leave you indifferent between three options:

- Choosing α_n and going to market α_n
- Choosing α_{n+1} and going to market α_{n+1}
- Choosing α_n and going to market α_{n+1}

Single crossing ensures that this fixed point exists: worsening the pool but raising the rate benefits α_{n+1} types more than α_n types

5. Continue until you reach $\alpha = 1$
6. Denote by W_Δ the resulting measure over the interval $[\alpha_0, 1]$.
7. Define the function $r_\Delta(\alpha)$ over the interval $[\alpha_0, 1]$ as:

$$r_\Delta(\alpha) = r(\alpha_n) \quad \text{for } \alpha \in [\alpha_n, \alpha_{n+1})$$

8. For any subset $A \subseteq [\alpha_0, 1]$, let: $W(A) = \lim_{\Delta \to 0} W_\Delta(A)$
9. For any $\alpha\,[\alpha_0, 1]$, let $r(\alpha) = \lim_{\Delta \to 0} r_\Delta(\alpha)$
10. Compute the total mass of entrants W
11. These steps define a decreasing function $W(\Pi)$. The last step is to invert this function to find the level of Π that is consistent with the exogenous total wealth of lenders W.

□

B Microfoundation for Borrowers Demand

Consider a borrower with type (τ, ω) endowed with a unit of capital and a project. She wants to obtain a loan $\ell(\tau, \omega)$ to invest $i(\tau, \omega)$ in period 1 to consume the proceeds in period 2. Each unit of investment in the morning produces ρ return. The cost of investment has to be covered by the borrower's initial endowment or credit, implying the following budget constraint

$$i(\tau, \omega) = 1 + \ell(\tau, \omega). \tag{B.40}$$

Furthermore, each borrower has to pledge her investment as collateral to obtain credit. Seizing the collateral is the only threat to enforce repayment from the borrowers, thus $(1 + r_t(\tau, \omega))\ell_t(\tau, \omega) \leq i_t(\tau, \omega)$. Using (B.40) this simplifies to

$$\ell_t(\tau, \omega) \leq \frac{1}{r_t(\tau, \omega)}. \tag{B.41}$$

Given the linear technology, all borrowers would like to borrow the maximum

$$D(r) = \frac{1}{r}$$

at the minimal interest rate they can obtain loans as described in Assumption 1.