Equilibrium Spillover of Big Data

Maryam Farboodi* Peter Kondor† Pablo Kurlat‡

March 20, 2025§

Abstract

We study a model of credit markets with adverse selection where ex-ante identical lenders invest in a screening technology to reduce their type I and type II error in identifying good borrowers. A rich market structure emerges in equilibrium with continuous heterogeneity in lender screening and non-assortative matching between lenders and borrowers. Furthermore, the equilibrium features a hockey stick interest rate schedule—a segmented market structure with variable degrees of fragmentation across different level of borrower opacity. We demonstrate that this market structure is robust to changes in the screening technology as well as lender entry. We then use the model to 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 benefit the underserved population and counterintuitively, increases the inequality in financial access.

*MIT Sloan, NBER and CEPR; farboodi@mit.edu

†LSE and CEPR; p.kondor@lse.ac.uk

‡USC and NBER; pkurlat@gmail.com

§We thank Fernando Alvarez, David Autor, Ricardo Caballero, John Horton, Simon Johnson, Lars Hansen, Philipp Kircher, Leonid Kogan, Ezra Oberfield, Jonathan Parker, Tong Liu, Esteban Rossi-Hansberg, Elisa Rubbo, Antoinette Shoar and Robert Shimer for their thoughts and comments.

1 Introduction

We develop an equilibrium model of credit markets with adverse selection and two-sided heterogeneity in lender and borrower types. Borrowers are heterogeneous in how creditworthy they are and their opacity—how difficult it is for lenders to recognize them as so. Lenders on the other hand 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 recognizing the different types of error that they will make. Borrowers choose interest rates and their quantity demanded recognizing that in order to obtain credit, they might have to subsidize other borrowers who they are pooled with, or give rents to specialized lenders.

In 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, a *hockey stick interest rate schedule* emerges and borrowers of the same creditworthiness might be able to borrow only at different interest rates depending on how opaque they are. Moreover, the equilibrium matching between borrowers and lenders is non-assortative.

We further show that this market structure is robust to entry of new lenders. We then use our framework to investigate the impact of improvement in big data technologies as well as adoption of big data policies on the financial sector. We find that adoption of AI technology by financial institutions benefits borrowers who face high interest 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.

We have deliberately kept the model stylized to focus on the information asymmetry in the credit market. Specifically, we use a general information structure that allow us to explore lenders' type I and type II errors about borrowers in a systematic manner. Furthermore, we adapt a nested information structure for lenders that we believe resembles the realistic differences between screening technologies used in the financial sector.

We demonstrate our results by examining a market populated by borrowers and lenders, both with (potentially) heterogeneous type distributions. Borrowers have a two-dimensional type. They are either good, creditworthy borrowers who pay the lenders back, or bad, not creditworthy borrowers who default and do not pay back. Moreover, they have different levels of opacity—i.e. the degree of information asymmetry vis-a-vis lenders. Importantly, the only borrower characteristic that impacts lender payoff is whether the borrower is good or bad—he pays the lender back or not. Conditional on knowing this, opacity of borrower is irrelevant for lender's credit extnesion decision.

However, lenders make both type I and type II errors in identifying creditworthy borrowers, depending on borrower's opacity. Lenders are endowed with a common baseline screening technology that determines the relative severity of their type I and type II errors. Furthermore, they endogenously choose their screening precision at a cost. The screening precision corresponds to the technological sophistication of the lenders and determines the total error rate that they make: A naive lender makes many errors in identifying borrower creditworthiness, while a sophisticated one make few. However, of facing the original pool of borrowers, the relative type I and type II error rates of both lenders are the same.

The defining feature of the information structure of lenders with different degrees of screening precision is its *nestedness.* Nestedness of information structure means that if a low precision, less technology-savvy lender correctly identifies the creditworthiness of a given borrower, a high precision, technology-savvy lender certainly does that as well. Alternatively, if a more tech-savvy lender misidentifies the creditworthiness of a borrower, a less tech-savvy lender will make the same mistake. In terms of borrower opacity, less opaque borrowers are those whose creditworthiness is correctly identified by more lenders. In other words, a high degree of borrower opacity requires a high level of lender precision to identify correctly whether the borrower is good or bad.

There is a continuum of markets, each defined by an interest rate. After lenders choose their technology precision, lenders and borrowers choose which market to go to, and the lenders decide to which borrowers on that market to lend. When considering a borrower, the lender receives a signal that depends on the two-dimensional borrower type, as well as the baseline and the precision of the lender screening technology. The decision of the lender whether to extend credit to the borrower or not has to measurable with respect to the lender signal. Alternatively, borrowers have a demand for credit that is downward sloping in the interest rate that they can borrow at. In equilibrium, the interest rate is determined in each market to equalize the demand and supply of credit in equilibrium.

We first establish a benchmark result for exogenous distribution of lender expertise. We show that the market structure is segmented with variable degrees of fragmentation across different level of borrower opacity. The interest rates follow a hockey stick interest rate schedule, as depicted in Figure 2. The least opaque good and most opaque bad borrowers borrow at a low common pooling interest rate from a heterogeneous group of lenders with intermediate levels of precision. This segment of the credit market features the most favorable interest rate to borrowers as moderately skilled lenders compete to relatively transparent good borrowers.

The next segment of the credit market is a fragmented segment where hard-to-recognize, opaque good borrowers raise credit from high precision lenders who are uniquely positioned to identify them. In this segment, borrowers are perfectly separated and face an interest rate schedule which is increasing in their degree of opacity. Lender specialization leads to low ex-post competition in different markets in this segment, which in turn means lenders engage in rent seeking.

In the final segment of the equilibrium structure, the least technology sophisticated lenders compete with the most sophisticated ones to serve the highly opaque good borrowers who are willing to pay a high interest rate to raise credit. This segment of the credit market resembles the financially excluded population.

We use this benchmark equilibrium to demonstrate two important properties of this market structure. First, lenders with a better technology have a higher quality portfolio. In other words, this is a credit market in which better lenders engage in creme skimming. Second, the matching between lender and borrowers is non-assortative: The least and most tech-savvy lenders both lend to the most opaque good borrowers at high interest rates.

We next show our first result that features endogenous choice of expertise. When lenders are ex-ante identical and can choose their screening precision at a cost, the credit market has a unique equilibrium in which ex-ante homogeneous lenders choose heterogeneous levels of

screening precision, and the hockey stick interest rate schedule we just describes persists. We also establish a duality result between the exogenous wealth distribution of lender precision and cost of acquiring precision.

Our second main result uses this duality argument to characterize the market structure when a group of lenders enter an equilibrium populated by incumbent lenders. Importantly, we show that the equilibrium market structure with the hockey stick interest rate schedule is robust to unanticipated entry of new lenders.

A central feature of our framework is the robustness of the equilibrium market structure to new entry and exogenous changes in the economic environment. Furthermore, despite being highly stylized, the hockey stick equilibrium structure of the model matches 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 literature, 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.

This evidence implies that asymmetric information is one of most fundamental frictions in credit markets. We have formulated the lender cost function in our framework to precisely attenuate this friction, for instance through better processing and more availability of consumer data. As such, we believe that our framework is well suited to study the impact of the recent big data technological growth, as well as policies that specifically target consumer data availability in the financial sector.

Our final main results use our model to explore the consequences of directed reductions in the cost function, that resemble specific big data technological improvements or policies, on the credit market. We focus on two particular exercises: Adoption of AI in the financial sector and a mandatory data sharing policy. We model adoption of AI in the financial sector as a directed reduction in cost of acquiring an advanced screening technology, and show that such a technological growth positively impact financial inclusion in the credit market.

Alternatively, we argue that mandatory data sharing policies mainly reduces the cost of moderate levels of screening. We show that not only such policies do not have any spillover to the under-served population, but also counterintuitively, they even amplify the inequality in financial access.

1.1 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 nestesness 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.

The rest of the paper is organized as follows. Section 2 lays out the model. Section 3 provides the characterization of the credit market equilibrium. Section 4 demonstrates that the equilibrium market structure is robust with respect to entry of new lenders as well as changes in the incumbents screening technology. Section 5 investigates the impact of big data technological growth and related policies on the credit market equilibrium of the main model. Section 6 provides a benchmark with non-nested information structure. Section 7

concludes.

2 Model

We model a credit market where the main friction is the information asymmetry between borrowers and lenders. There are two dates, $t = 1, 2$ and there is no discounting. There are two types of agents, lenders and borrowers. Both lenders and borrowers have heterogeneous type distributions. Furthermore, each lenders and borrowers have a two-dimensional type.

There is a continuum of heterogeneous borrowers, each with a project that requires investment at $t = 1$ and pays back at $t = 2$. Borrowers have no capital. Let $d(r)$ denote the demand of a borrower who faces interest rate r. We assume $d(r)$ is downward sloping.[1]

Each borrower has a two-dimensional type, (τ, ω). The critical property of borrower type is that the dimension that controls borrower performance is different from the dimension that controls information asymmetry between lender and borrowers. In particular, $\tau \in \{G, B\}$ controls borrower performance vis-a-vis the lender. It determines whether the borrower is creditworthy ($\tau = G$), i.e., he will pay the lender back; or not ($\tau = B$). Alternatively, the second dimension of borrower type, $\omega \in [0, 1]$, is the opacity of the borrower. $\omega = 0$ (1) is the least (most) opaque. Each lender receives a signal $s \in \{G, B\}$ from each borrower (τ, ω). More lenders receive a signal s that does not match the true τ for borrowers with higher ω. Put differently, τ of the more opaque borrowers is more difficult to identify correctly for more lenders, thus, they face more information asymmetry. The continuum of borrowers is distributed with CDF/PDF $G(\omega)/g(\omega)$ and $B(\omega)/b(\omega)$ on $\omega \in [0, 1]$.

There is a mass of lenders. Each lender has one unit of capital. All lenders are endowed with a common exogenous *basic screening technology*, $\beta \in [0, 1]$. Furthermore, each lender has a one-dimensional type $\alpha \in [0, 1]$ which denotes the *precision* of his screening technology. Screening precision can be exogenous, or endogenously chosen at cost $C(\alpha)$ by lenders at $t = 1$.

At the beginning of date $t = 2$, a continuum of markets are open. Each market is defined by an interest rate. Lenders and borrowers choose which market(s) to participate in. In each market, lenders receive (heterogeneous) signals from borrowers. Each lender lends to a portfolio of borrower that has to be measurable with respect to the signal lender receives from him.[2] Borrowers invest what they can raise from lenders. At the end of period $t = 2$, borrowers' projects pay off and good borrowers pay back the lenders.

The following definition introduces the information structure of the lenders.

Definition 1 (Nested Information structure). *When lender α meets borrower (τ, ω),*

[1] Here, we take $d(r)$ as an exogenous demand curve. Farboodi and Kondor (2022) provides a micro-foundation for a downward sloping demand curve which satisfy the conditions necessary for our market structure.

[2] Please see Kurlat (2016) and Farboodi and Kondor (2022) for detail of the market structure.

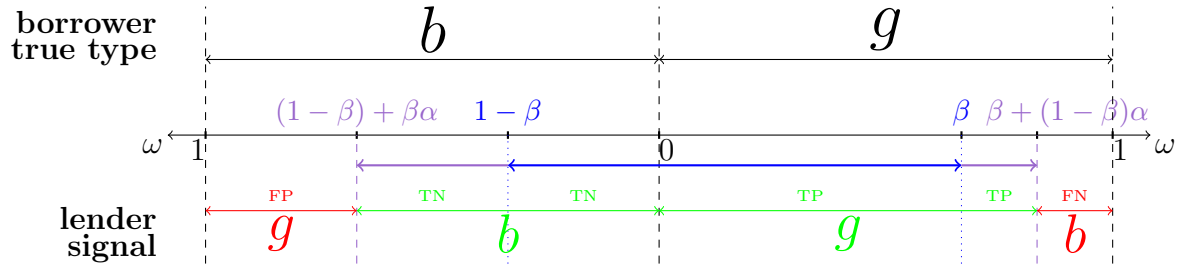


Figure 1: Signal structure of lender α from borrower (τ, ω), as specified in (1).

screening technology β with precision α generates signal:

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

Figure 1 illustrates the information structure proposed in Definition 1. This information structure has implications for individual lenders' type I and type II error rates, as well as how the correlation of errors among lenders. We next explain each of these implications, under the null hypothesis that "the borrower is good."

Type II errors: False Negative $\omega_g(\beta,\alpha) = \beta + \alpha(1-\beta)$ is the the highest-ω (most-opaque) good borrower who is identified as good by lender α. Put differently, lender α perceives all good borrowers with $\omega < \omega_g$ correctly as good while he perceives more opaque good borrowers, with $\omega < \omega_g$, as bad ones incorrectly. Alternatively, for a (g,ω) borrower, $\alpha(g,\omega) = \frac{\omega-\beta}{1-\beta} > 0$ is the lowest precision lender who identifies that borrower as a good one. If $\alpha(g,\omega) \leq 0$, borrower (g,ω) is perceived as a good borrower by all lenders, which is the case for $\omega \leq \beta$. For instance, in the extreme case of $\beta = 1$ (type I errors only), every lender recognizes every good borrower as good.

Type I errors: False Positive Among the bad borrowers, $\omega_b(\beta,\alpha) = 1 - \beta + \alpha\beta$ is the the highest-ω (most-opaque) bad borrower who is identified as bad by lender α. Lender α perceives all bad borrowers with $\omega < \omega_b$ correctly as bad and all bad borrowers while he perceives bad borrowers with $\omega > \omega_b$ mistakenly as good one. Alternatively, for a (b,ω) borrower, $\alpha(b,\omega) = \frac{\omega-(1-\beta)}{\beta} > 0$ is the lowest precision lender who identifies that borrower as a bad one. If $\alpha(b,\omega) \leq 0$, borrower (b,ω) is perceived as a bad borrower by all lenders, which is the case for $\omega \leq 1-\beta$. For instance, in the extreme case of $\beta = 0$ (type II errors only), every lender recognizes every bad borrower as bad.

The above implications demonstrate the first two properties of the information structure in Definition 1. First, lender type, α, determines the total error rate that they make. In

particular, lender with precision α has a lender specific total error rate of $1-\alpha$: Higher expertise lenders, i.e. lenders with high α, make fewer error.

Second, every lender makes both type I and type II errors in lending to borrowers. Technology parameter β controls the relative severity of type I and type II error for every lenders. In particular, for every lender, $\frac{\text{type I error rate}}{\text{type II error rate}} = \frac{\beta}{1-\beta}$. One interpretation of β is that it corresponds to aggregate business cycle conditions. For instance, in good times lenders are more liniment towards type I (false positive) mistakes as long as they make sufficient profits on the good investment opportunities. This corresponds to large values of β. Alternatively, in bad times lenders are cautious and prefer to avoid making any bad investment, at the cost of losing out on good investment opportunities. This implies higher propensity of type II errors and corresponds to a small β. A different interpretation is the information that lenders are allowed to condition on. For instance, an ongoing policy discussion is whether to allow insurers to condition the rates on past health outcomes. Allowing the use of adverse past outcomes corresponds to a reduction in parameter β: It is less likely that insurers mistakenly provide cheap insurance to risky borrowers, but it is more likely that they make mistakes about now-healthy individuals due to prior adverse conditions.

Correlation of errors across lenders Consider two lenders $\alpha' < \alpha$, i.e. lender α' is *less precise* than lender α. The errors that lenders α and α' make are correlated. In particular, error that lender α makes are a subset of errors that lender α' makes. To be more precise, for any borrower of opacity ω, if the less precise lender α identifies borrower τ correctly, the more precise lender α also identifies that borrower correctly. Alternatively, if lender α does not identify borrower τ correctly, lender α' does not identify τ correctly either.

This correlation underlies the third and last property of the information structure in Definition 1, *nestedness*: The information structure of lenders of difference levels of expertise are nested. This property is what differentiates the information structure in Definition 1 from the commonly used iid information structure. Section 6 highlights the the differences between our equilibrium structure with the equilibrium that relies on the iid, non-nested information structure.

The next definition uses our proposed information structure to formalize the notion of equilibrium in this framework.

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

1. *measure $w(\alpha)$ over lender screening precision that adds up to W,*
2. *choice-of-market function $r(\alpha)$ and a (binary) choice-of selectiveness function $s(\alpha)$ for each lender α,*
3. *measures of good and bad borrowers available to lender α in market r: $G(\cdot; r, \alpha)$ and $B(\cdot; r, \alpha)$*

such that

1. *Given α, lenders optimize over what market to go to and how selective to be:*

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

where $\gamma(\cdot)$ is defined below;

2. *Lenders optimize over α, i.e. every α in the support of $w(\alpha)$ satisfies:*

$$\alpha \in \arg\max_{\alpha} \pi(\alpha) = \gamma(r(\alpha), s(\alpha); \alpha)(1 + r(\alpha)) - 1 - C(\alpha);$$

3. *The measures G and B satisfy the accounting condition specified below.*

Let Ω^G and Ω^B be any subsets of good and bad borrowers.

- *If lender α is selective, he will obtain a measure (per unit of wealth) over these of*

$$G^S\left(\Omega^G; r, \alpha\right) = \frac{G\left(\Omega^G \cap [0, \beta + \alpha(1-\beta)]; r, \alpha\right)}{G([0, \beta + \alpha(1-\beta)]; r, \alpha) + B([1 - \beta + \alpha\beta, 1]; r, \alpha)} \tag{2}$$

$$B^S\left(\Omega^B; r, \alpha\right) = \frac{B\left([1 - \beta + \alpha\beta, 1] \cap \Omega^B; r, \alpha\right)}{G([0, \beta + \alpha(1-\beta)]; r, \alpha) + B([1 - \beta + \alpha\beta, 1]; r, \alpha)} \tag{3}$$

  In these expressions, the denominator is the mass of all borrowers the lender accepts, and the numerator in the mass in subsets Ω^G and Ω^B respectively.

  Each of these borrowers demand $d(r)$ units of capital.

- *If the lender is non-selective, he will obtain:*

$$G^{NS}\left(\Omega^G; r, \alpha\right) = \frac{G\left(\Omega^G; r, \alpha\right)}{G([0,1]; r, \alpha) + B([0,1]; r, \alpha)}$$

$$B^{NS}\left(\Omega^B; r, \alpha\right) = \frac{B\left(\Omega^B; r, \alpha\right)}{G([0,1]; r, \alpha) + B([0,1]; r, \alpha)}$$

  Each of these borrowers demand $d(r)$ units of capital.

  Using this, define:

$$\gamma^S(r, \alpha) = \frac{G^S([0,1]; r, \alpha)}{G^S([0,1]; r, \alpha) + B^S([0,1]; r, \alpha)} \tag{4}$$

$$\gamma^{NS}(r, \alpha) = \frac{G^{NS}([0,1]; r, \alpha)}{G^{NS}([0,1]; r, \alpha) + B^{NS}([0,1]; r, \alpha)} \tag{5}$$

  when the denominators are positive, and zero otherwise. Aggregate these two definitions into

$$\gamma(r, s; \alpha) = s\gamma^S(r, \alpha) + (1 - s)\gamma^{NS}(r, \alpha). \tag{6}$$

Finally, these measures have to satisfy two accounting condition. Let $A(r, \alpha) = \{\tilde{\alpha} : r(\tilde{\alpha}) < r\} \cup \{\tilde{\alpha} : r(\tilde{\alpha}) = r, \tilde{\alpha} < \alpha\}$ be the set of lenders that choose a lower-interest-rate market than α or choose the same interest rate but pick before, i.e. all the lenders that pick before α.

The measures over good/bad types faced by lender α in market r, defined for an arbitrary subset Ω^{G/Ω^B}, are

$$G\left(\Omega^G; r, \alpha\right) = G\left(\Omega^G\right) - \int_{A(r,\alpha)} \left[G^S\left(\Omega^G; r\left(\alpha\right), \alpha\right) s\left(\alpha\right) + G^{NS}\left(\Omega^G; r\left(\alpha\right), \alpha\right)\left[1 - s\left(\alpha\right)\right]\right] dW\left(\alpha\right) \tag{7}$$

$$B\left(\Omega^B; r, \alpha\right) = B\left(\Omega^B\right) - \int_{A(r,\alpha)} \left[B^S\left(\Omega^B; r\left(\alpha\right), \alpha\right) s\left(\alpha\right) + B^{NS}\left(\Omega^B; r\left(\alpha\right), \alpha\right)\left[1 - s\left(\alpha\right)\right]\right] dW\left(\alpha\right) \tag{8}$$

These equations start from the original measure and subtract all the borrowers that are served by lenders with lower technological precision.

Equations (7) and (8) describe the pool of good and bad borrowers that a lender with precision α faces. These equations impose a specific "clearing order" among lenders: The markets start clearing from the lowest interest rate, and within each market the less precise lenders choose their portfolio before the more precise lenders. This clearing order seems opposite the common creme skimming intuition in financial markets. However, in Lemma 1 we demonstrate that given the nested information structure, this clearing ordering is necessary for creme skimming. In the Appendix, we show that this clearing order endogenously emerges in equilibrium among all possible clearing algorithms. Furthermore, we assume markets are non-exclusive.

3 Equilibrium Structure

We start this section by defining a "hockey stick interest rate schedule" which we will extensively use throughout the paper.

Definition 3 (**Hockey Stick Interest Rate Schedule**). *A hockey stick interest rate schedule is a weakly increasing schedule defined by thresholds $\beta \leq \omega_1 \leq \omega_2 \leq 1$ and $\omega_{1,b} \geq 1 - \beta$ on borrower opacity and $0 \leq \alpha_0 \leq \alpha_1 \leq \alpha_2 \leq 1$ on lender screening precision. It consist of (at most) three segments:*

1. **Region I:** *A unique common low interest rate r_p.*

   Every easy-to-recognize good borrower, (g, ω) with $\omega \leq \omega_1$, every not-easy-to-recognize bad borrower, (b, ω) with $\omega > \omega_{1,b}$, and some easy-to-recognize bad borrowers, (b, ω) with $\omega < \omega_{1,b}$ borrow at the same common low interest rate r_p.

   Lenders with intermediate degrees of technological precision, $\alpha \in [\alpha_0, \alpha_1]$, lend at this low interest rate.

2. **Region II:** *An increasing interest rate schedule.*

   Each moderate-to-recognize good borrower, (g, ω) with $\omega_1 < \omega \leq \omega_2$, borrows at his individualized interest rate. No bad borrower borrows.

   Each lender with a high degree of technological precision, $\alpha \in [\alpha_1, \alpha_2]$, lend at their individualized interest rate.

3. **Region III:** *A common high interest rate r_{NS}.*

   Every hard-to-recognize good borrower, (g, ω) with $\omega > \omega_2$ and some easy-to-recognize bad borrowers, (b, ω) with $\omega < \omega_{1,b}$ borrow at the same common high interest rate r_{NS}.

   Lenders with the lowest technology level, $\alpha = 0$, along with any remaining very-high technology lenders, $\alpha > \alpha_2$, lend at this high interest rate.

We goal is to understand how information technologies impact the lending market, and the changes that follow from adoption of big data technologies in the credit market. To clarify the exposition, we use the following terminology throughout the paper.

Definition 4 (Incumbent Equilibrium and Entry Equilibrium).
An incumbent equilibrium is an equilibrium where there is a single group of lenders in the credit market. They can be endowed with an arbitrary expertise distribution or with a cost function to choose their expertise.

An entry equilibrium is an equilibrium where the market is populated with a group of lenders with a given arbitrary distribution. A second group of lenders, the new entrants, are endowed with a cost function to choose their expertise and enter the credit market.

Finally, we make the following assumptions throughout the main text.

Assumption 1. *Assume*

1. *If the incumbent wealth distribution, $w(\alpha)$, is exogenous, it is continuous almost everywhere and decreasing in α, with a potential mass point at $\alpha = 0$.*

2. *$g(\omega)$ and $b(\omega)$ are uniform.*[3]

3.1 Incumbent Equilibrium

In order to build some intuition, we start with an exogenous capital distribution of lenders $w(\alpha)$, i.e. an *incumbent equilibrium with exogenous distribution of lenders*. The first proposition shows that this equilibrium structure is akin to Definition 3.

Proposition 1 (Incumbent Equilibrium: Exogenous Distribution of Lenders).
Assume the wealth distribution of lenders $w(\alpha)$ is exogenous.

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 hockey-stick interest rate schedule, described in Definition 3

Proposition 1 provides a tight description of the equilibrium market structure when the wealth distribution of lenders with different degrees of technological precision is exogenous. It will serve as a stepping stone to characterize the emergence of this heterogeneous market structure even when lenders are ex-ante homogeneous, as well as with entry of new lenders.

The most important feature of this credit market is the segmentation that endogenously arises in equilibrium. This segmentation has two distinct features. First, borrowers with

[3]This assumption considerably simply the exposition, but it is not necessary for the results.

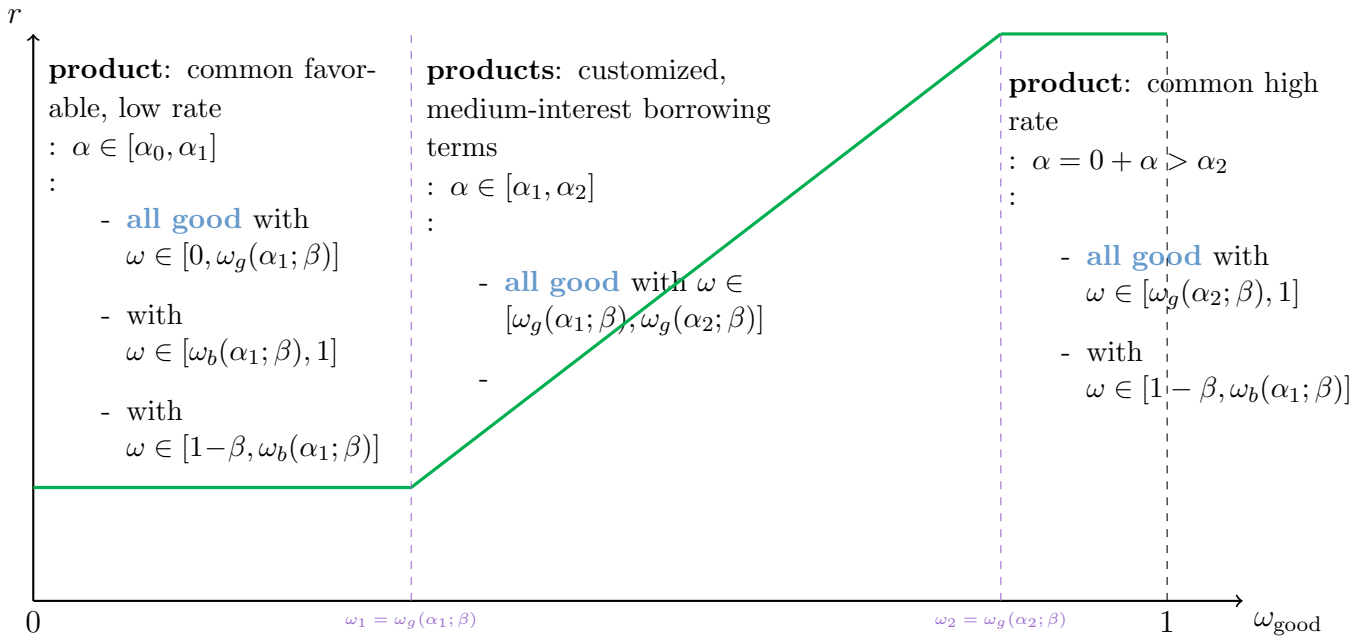


Figure 2: Equilibrium structure of the credit market in Proposition 1.

the same quality are segmented and can only borrow at different interest rates; and second, degree of segmentation varies across the distribution of borrower opacity: more transparent good borrowers are less differentiated than less transparent ones. To be more precise, the market structure features both pooling and separating market segments. The concurrent existence of different market segments is the key difference between our framework and one where lenders have a non-nested information structure. Section 6 provides the analogues of Proposition 1 with a non-nested information structure and highlights the differences.

To clarify this asymmetric segmentation, it is useful to divide the credit market into three distinct segments. In the lowest interest rate segment, **Region I**, lenders who have intermediate levels of technological precision, $\alpha \in [\alpha_0, \alpha_1]$, lend selectively at a common favorable rate r_p to good borrowers who are easy-to-recognize, $\omega \in [0, \omega_g(\alpha_1;\beta)]$, where $\omega_g(\alpha_1;\beta) = \omega_1 = \beta + \alpha_1(1-\beta)$. However, as their screening technology is not precise, they mistake hard-to-recognize bad borrowers with good borrowers and lend to them as well. As such, bad borrowers with $\omega \in [\omega_b(\alpha_1;\beta), 1]$ and some with $\omega \in [1-\beta, \omega_b(\alpha_1;\beta)]$, where $\omega_b(\alpha_1;\beta) = (1-\beta) + \alpha_1\beta$, can also borrow at this favorable rate due to the type II errors of lenders with intermediate degrees of screening precision. The interest rate r_p is set such that the lenders with lowest precision, $\alpha = \alpha_0$, make (weakly) positive profits, while lenders with better technology make more profits.

For each lender α with $\alpha \in [\alpha_0, \alpha_1]$, consider the good borrowers in continuous interval

$\omega \in [0, \omega_g(\alpha;\beta)]$ and bad ones in $\omega \in [\omega_b(\alpha;\beta), 1]$. Equations (2) and (3) simplify to:

$$G(\alpha;\alpha_0) = \int_0^{\beta+\alpha(1-\beta)} g(\omega;\alpha)\, d\omega \tag{9}$$

$$B(\alpha;\alpha_0) = \int_{1-\beta+\alpha\beta}^{1} b(\omega;\alpha)\, d\omega \tag{10}$$

Lender α perceives every borrower in the above two sets as a good borrower and lends to a representative pool of all of them. Furthermore, the total demand from all of the borrowers in $G(\alpha_1;\alpha_0) + B(\alpha_1;\alpha_0)$ has to be absorbed by lenders $\alpha \in [\alpha_0, \alpha_1]$.

In this segment lenders use their signals to choose selectively, i.e. they lend only to borrower from whom they receive a good signal. Let $T(\alpha;\alpha_0)$ denote the set of all of the acceptable borrowers to lender with precision α,

$$T(\alpha,\alpha_0) = B(\alpha;\alpha_0) + G(\alpha;\alpha_0) \tag{11}$$

Use this notation in Equations (4) and (6) to simplify the quality of borrowers for lender with precision α to

$$\gamma(\alpha) = \frac{G(\alpha)}{T(\alpha)}.$$

Let $\gamma_0(\alpha_0)$ denote the quality of borrowers acceptable to the lowest precision lender in Region I:

$$\gamma_0(\alpha) = \frac{G([0, \beta+\alpha(1-\beta)])}{G([0, \beta+\alpha(1-\beta)]) + B([1-\beta+\alpha\beta, 1])}. \tag{12}$$

This quality follows Equations (4) and (6) as well, with the simplification that no borrowers have exited the pool yet. As such, it only depends on the original distribution of borrower types. It follows that $\gamma_0(\alpha_0) < 1$.

As lenders α_0 are those with the least precise screening technology in this market, they makes the most type I and II errors and faces the most defaults in their portfolio. Type I (false positive) error imply that they lend to some borrowers who will ex-post default and not pay back, while type II errors imply that they give up on some good investment opportunities. As every lender makes at least π profits, the interest rate in the pooling segment has to be set to satisfy the indifference condition of α_0 lenders who face most defaults, $\gamma_0(\alpha_0)(1+r_p) = 1+\pi$. As such, the interest rate in Region I is given by:

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

As both good and bad borrowers are served at a common interest rate in this region, we call Region I the *pooling segment* of the credit market.

Lenders with screening precision α absorb equal share of all the borrowers who they lend to, proportional to their wealth, $w(\alpha)$. As such, throughout Region I, borrower demand is absorbed by lenders with screening precision α at rate $\theta(\alpha)$ given by:

$$\theta(\alpha) = \frac{w(\alpha)}{d(r_p)\, T(\alpha)}. \tag{14}$$

Recall that Equations (7) and (8) imply that when lenders with different screening precision lend at the same interest rate, the lowest precision lenders choose their portfolio first.[4] This order implies that the quality of borrowers faced by lenders improves as lender screening precision increases, as formalized by Lemma 1. As $\gamma(\alpha)$ is bounded above by one, i.e. the perfect lending portfolio, the last group of lenders who lend at the pooling interest rate r_p, α_1, must have $\gamma(\alpha_1) = 1$. Put differently, the depletion of borrower demand in the pooling market continues by lenders of increasing screening precision until all the borrowers identified as good by the highest precision lender in this market, i.e. lender α_1, are exhausted,

$$T(\alpha_1, \alpha_0) = 0.$$

Proof of Proposition 1 further shows that existence and uniqueness of equilibrium requires

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

These conditions are effective smooth pasting conditions at the transition point between **Region I** and **Region II,** i.e. at the lender screening precision level α_1. At this point the pooling segment of the credit market ends and a fragmented segment starts in which hard-to-recognize good borrowers borrow form lenders who have a high-precision screening technology and are able to identify these good borrowers, $\alpha \in [\alpha_1, \alpha_2]$.

In particular, consider borrowers $(g, \omega_g(\alpha; \beta))$. These borrowers can only be identified by lenders with precision screening technology α and higher, but by no lender $\alpha' \in [\alpha_1, \alpha)$. Lenders α use this advantage and engage in extracting rents from $(g, \omega_g(\alpha; \beta))$ borrowers. As such, each group of good borrowers ω are served at a different interest rate determined by cash-in-the-market pricing (CIM):

$$r(\alpha) = d^{-1}\left(\frac{w(\alpha)}{g\left(\omega_g(\alpha; \beta)\right)(1-\beta)}\right).$$

In this fragmented segment of the credit market, the rent extraction becomes more pronounces and interest rates rise as lenders' screening technology becomes more precise, as we assume that expert wealth is in short supply– Assumption 1. It follows that **region II** constitutes a continuum of markets with an increasing interest rate schedule. The interest rate in each market is determined to equalize lender capital and borrower demand. As borrowers are served at different interest rates determined by through cash-in-the-market pricing in this region, we call Region II the *CIM segment* of the credit market.

Finally, the most opaque good borrowers, whose type is extremely difficult to recognize, can only be identified by lenders with very precise screening technology whose capital is very scarce, if it exists. As such, these lenders require a very high interest rate to lend. At the same time, lenders with the least precise screening technology, i.e. $\alpha = 0$, can always enter and use a non-selective strategy, lend to a representative pool of borrowers and hedge their

[4]We assume this pattern in the main text and prove in the appendix that it endogenously emerges as an equilibrium outcome among all possible clearing ordering. The proof follows Kurlat (2016).

exposure to default risk of bad borrowers by charging a relatively high interest rate, the "non-selective interest rate":

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

This non-selective interest rate puts an upper bound on how much the very high-tech lenders can charge the very-hard-to-recognize good borrowers. At the lender precision level α_2 such that $r_{NS}(\alpha_2) = r(\alpha_2)$, the credit market transitions from **Region II** to **Region III**. In this region, lenders with very high and very low screening precision jointly lend to all the remaining good borrowers and very bad borrowers. The very high precision lenders make abundant profits as the quality of their portfolio is perfect, while the low precision lenders suffer defaults and only make the required profit $\pi \geq 0$. As a group of lenders who lend non-selectively participate in this segment of the credit market, we call **Region III** the *non-selective segment* of the credit market.

It is important to observe that in this equilibrium, lenders make different profit levels as they are exogenously endowed with different levels of screening precision. The least precise lenders, those choosing non-selectively ($\alpha = 0$) and the least precise selective lenders ($\alpha = \alpha_0$) make minimum profit π. Every other lender makes strictly higher profits, and profits are increasing in lender precision.

The first proposition presents the equilibrium structure when new lenders enter a market with no incumbents.

Proposition 2 (**Incumbent Equilibrium: Endogenous Distribution of Lenders).** *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 hockey-stick interest rate schedule, described in Definition 3.

Note that there is a one-to-one mapping between the total lender wealth W, and the profit that lenders make in equilibrium, $\pi(W)$. Conjecture a profit level π. With that conjecture, within each segment of the hocket stick equilibrium structure of Definition 3, the determination of equilibrium in Propositions 2 and 1 share common features as they both involve a single group of lenders.

Consider Region I, the pooling segment. When lender wealth is determined endogeneously, every lender $\alpha \in [\alpha_0, \alpha_1]$ has to make the same profits:

$$\gamma(\alpha)(1+r_p) - C(\alpha) = 1+\pi, \tag{15}$$

where $\gamma(\alpha)$ follows Equation (6), which in turn uses Equations (16) and (17) to determine the measure of good and bad borrowers acceptable to lender α. The evolution of these measures is in turn impacted by wealth of all the less precise lenders in the pooling segment via function $\theta(.)$ in Equation (14). As such, sufficient lender wealth has to enter at each precision level to ensure that the above indifference condition hold for every $\alpha \in [\alpha_0, \alpha_1]$.

Recall that Region II, i.e. the CIM segment, is a fragmented segment with a continuum of markets which feature different interest rates. The indifference condition of lenders in this

segment governs the change in interest rate across markets:

$$r'(\alpha) = C'(\alpha).$$

The lender wealth at each precision is determined such that the indifference condition is satisfied pointwise:

$$w(\alpha) = d\left(r(\alpha)\right) g(\beta + \alpha(1-\beta))(1-\beta). \qquad \forall \alpha$$

Finally, Region III entirely consists of lenders who only have the base technology β with $\alpha = 0$ and lend non-selectively. At any precision level α_2 where the non-selective interest rate, r_{NS}, equalizes the CIM interest rate, the latter segment ends and the former starts, with a unique interest rate of r_{NS}.

$$\gamma_{NS}(\alpha_2)\left(1 + r_{NS}(\alpha_2)\right) = 1 + \pi$$

The extra piece to determine in Proposition 2 is the minimum (strictly positive) precision of any lender who enters the market, α_0. α_0 is the precision level that minimizes the pooling interest rate r_p for a given profit level π, when $\gamma(\alpha) = \gamma_0(\alpha)$ and the pooling interest rate is determined by Equation (15). This is exactly the dual of the problem of finding the precision level that maximizes profit level π, for a given pooling interest rate r_p, when $\gamma(\alpha) = \gamma_0(\alpha)$:

$$\alpha_0 = \operatorname{argmin} \frac{\pi + C(\alpha)}{\gamma_0(\alpha)}$$

The Final step is to verify a consistency condition that ensures the endogenous wealth of lenders W adds up to the exogenous total wealth that corresponds to the conjectured profit level π.

An important implication of Proposition 2 is that when ex-ante homogeneous lenders choose the precision of their screening technology endogenously, the unique equilibrium is heterogeneous. This is even the case in the market segment where a common interest rate prevails. The ex-post heterogeneity arises due to two different forces in different market segments. The first force is gain to specialization, which leads to lenders' choice of heterogeneous precision in the cash-in-the-market pricing segment. The second force is less intuitive, and leads to heterogeneous choice of precision by lenders who all lend at the same pooling interest rate. Here, the reason is the improvement in portfolio quality that follows from spreading out screening precision.

To see the intuition, consider a market in which a positive measure of borrowers with different opacity borrow, and a positive measure of lenders with a common screening precision lend at the common interest rate. Consider an individual lender. He can pay the marginal cost $C'(\alpha)$ and choose a marginally higher precision. However, his portfolio quality improves discretely as a strictly positive measure of lenders at a lower precision level α improve his quality by asymmetrically absorbing a strictly positive measure of good and bad borrowers acceptable to this lender, as illustrated in Equations (16) and (17). Since everyone is lending at the same interest rate, each individual lender has an incentive to deviate from the mass point, and that cannot be an equilibrium outcome.

Propositions 1 and 2 illustrate that whether the incumbents are endowed with an exogenous distribution of expertise or if they are exante identical and choose their expertise endogenously with a cost function, the market structure remains the same. The next proposition formalizes this duality.

Proposition 3 (**Duality of Distribution of Lenders' Wealth & Cost of Acquiring Precision**). *For 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 $w(\alpha)$, up to an additive constant.*

3.2 Equilibrium Properties

Let $r(\omega)$ denote the interest rate that good borrower (g, ω) is served at, and $r(\alpha)$ the interest rate offered by (a non-zero measure of) lenders with precision α, i.e. lenders α. The first equilibrium property is that the interest rate schedule is increasing both in ω and α, i.e. $\frac{dr(\omega)}{d\omega} \geq 0$ and $\frac{dr(\alpha)}{d\alpha} \geq 0$. This property directly follows from Proposition 1 and is reflected in Definition 3 of the hockey stick interest rate schedule.

A key equilibrium outcome in the credit market is the quality of the portfolio that each lender lends to, $\gamma(\alpha; \beta)$. This quality captures the share of borrowers in lender's portfolio that are creditworthy and will pay back,

$$\gamma(\alpha; \beta) \equiv \Pr(\tau = g | x = g, \alpha; \beta).$$

The portfolio quality is determined in general equilibrium and depends both on the baseline technology β, lender specific screening precision, α, as well as the equilibrium market structure. In particular, it depends on the mass of good and bad borrowers who are identified as good by lender α and are not served on markets with lower interest rates, or at the same interest rate by lower precision lenders.

As such, for a lender α, the wealth of other lenders who lend at the same interest rate as him and have a less precise screening technology impacts lender α's portfolio quality. This observation directly stems from Definition 2 that implies that in equilibrium, lenders with a less precise technology choose their portfolio first, both across and within markets. This clearing ordering seems counterintuitive at the first glance as one would imagine that choosing first gives lenders a comparative advantage. However, the nesteness of information structure of lenders have precisely the opposite implication—lenders with most precise technologies prefer to choose last. The next Lemma formalizes this result.

Lemma 1 (**Improvement in Selection**). *Lender portfolio quality is improving with lender precision. The improvement is strict if $\beta > 0$.*

$$\frac{d\gamma(\alpha)}{d\alpha} > 0 \; \forall \beta > 0; = 0 \quad \textit{if } \beta = 0$$

Lemma 1 implies that despite the fact that most precise lenders wait and lend last, creme skimming is prevalent in equilibrium. In other words, each lender serves the most profitable borrowers who they can identify given the market structure.

To see this, consider any market where lenders with different precision levels lend at the same interest rate. Consider two lenders with $\alpha_1 < \alpha_2$. Information set of of the less precise lender, α_1, is nested in that of the more precise lender, α_2. This implies that α_1 misindentifies every bad borrower who α_2 misidentifies as good, but cannot correctly identify all the good borrowers who α_2 correctly identifies. Equatiosn (16) and (17) below illustrate the evolution of acceptable good and bad borrowers formally. They are the derivatives of Equations (7) and (8) in Definition 2 when applied in a continuous interval.

$$\frac{\partial g(\omega;\alpha)}{\partial \alpha} = -\theta(\alpha)\,\mathbb{I}(\omega \leq \beta + \alpha(1-\beta))\,g(\omega;\alpha) \tag{16}$$

$$\frac{\partial b(\omega;\alpha)}{\partial \alpha} = -\theta(\alpha)\,\mathbb{I}(\omega \geq 1 - \beta + \alpha\beta)\,b(\omega;\alpha) \tag{17}$$

It follows that a less precise lender impacts the portfolio quality of a more precise lender on the same market asymmetrically. The less precise lender absorbs more of bad borrowers relative to good borrowers acceptable to the more precise lender and thus, improves the more precise lender's portfolio quality.

As such, allowing α_1 to choose first improves the pool of borrowers who α_2 faces, and thus, improves the quality of his portfolio. We impose this clearly ordering as an assumption in our equilibrium definition. In the appendix, we show that it endogenously emerges as an equilibrium outcome.[5]

The evolution of lender portfolio quality is a useful statistics to think about definition of markets. Recall that each market $m \in M$ is defined by an interest rate. Thus, we use $m(r)$ to indicate a market with prevailing interest rate r. The next lemma provides an alternative way to define a market.

Lemma 2 (**Increase in Interest Rate Across Markets).** *For the interest rate to increase between market $m(r)$ and the market with the lowest interest rate above r, it must be that there are lenders α on market $m(r)$ such that $\gamma(\alpha) = 1$.*

This result provides an alternative definition of a market. In particular, it shows that if a range of lenders lend in a single market, the highest precision lender must have $\gamma(\alpha) = 1$ before the market transitions to a higher interest one. Thus, it shows that a pooling market can only "end" in $\gamma(\alpha_{\max}) = 1$. Since once $\gamma(\alpha) = 1$ in a market $m(r)$, no bad borrowers with $\omega > 1 - \beta + \beta\alpha$ has remained that have not been already served at interest rate (weakly below) r, $\gamma(\alpha) = 1$ for lenders α who choose selectively in any market with $r' > r$. Similar to interest rate that is monotonically increasing, $\gamma(\alpha)$ is also weakly increasing for selective lenders across the whole range of α chosen in equilibrium, $\frac{d\gamma(\alpha)}{d\alpha} \geq 0$. This monotonocity is violated for lenders who lend non-selectively in a market segment such as $\alpha = 0$ lenders in the non-selective segment.

Lemma 2 also provides an alternative definition to a pooling versus a CIM market. A pooling market is one where different lenders, $\alpha \in [\alpha_{\min}, \alpha_{\max}]$ lend with different $\gamma(\alpha)$, and $\gamma(\alpha_{\max}) = 1$. A CIM market is one where only one type of lender α lends with $\gamma(\alpha) = 1$.

[5]The proof follows Kurlat (2016). We relegate it to the appendix as it considerably complicated the exposition in the main text while not being entirely novel.

It further shows that a pooling segment can only involve lower precision lenders and lower opacity good borrowers compared to a CIM segment in this market structure, as some lenders in a pooling segment have $\gamma(\alpha) < 1$ while $\gamma(\alpha) = 1$ for all lenders in the CIM segment.

The next property of this market structure is the non-assortative matching between lenders and borrowers, as formalized by the following corollary.

Corollary 1 (**Non-Assortative Lender-Borrower Matching).** *The matching between lenders' screening precision and borrowers opacity is non-assortative in the credit market equilibrium.*

Corollary 1 directly follows Proposition 2. Throughout most of the market structure, more precise lenders lend to harder-to-recognize borrowers and make higher profits. However, this pattern breaks down in Region III: There, lenders with the least precise screening technology join the ones with the most precise screening technology to lend to the hardest-to-recognize borrowers, making minimal profits. The competition between lenders with highest and lowest precision has important consequences for impact of adoption of AI in the financial sector on financial inclusion.

There are two key differences between the equilibria in Propositions 1 and 2. First, in Proposition 2, all lenders make the same profits as they are ex-ante identical, while in Proposition 1 higher precision lenders make strictly more profits. Second, the endogenous wealth distribution that arises in equilibrium in Proposition 2 is generically not decreasing in α. Despite these differences, they both have properties of the hockey stick equilibrium, described in Definition 3. The following lemma relates the main equilibrium features in these two equilibria to the exogenous elements in the economic environment.

Lemma 3 (**Equilibrium Features with a Single Group of Lenders).**
In an incumbent equilibrium with exogenous lender distribution $w(\alpha)$, across the cash-in-the-market pricing segment $r'(\alpha) \sim w(\alpha)$.

Alternatively, in the incumbent equilibrium with exogenous identical cost of expertise $C(\alpha)$ and endogenous lender distribution, $\gamma'(\alpha) \sim C'(\alpha)$ in the pooling segment while $r'(\alpha) \sim C'(\alpha)$ across the the cash-in-the-market pricing segment.

For a lender with precision α, $\gamma(\alpha)(1 + r(\alpha))$ represents the gain from lending. When the lenders choose their precision level, the marginal benefit is

$$\gamma'(\alpha)(1 + r(\alpha)) + \gamma(\alpha)r'(\alpha),$$

while the marginal cost of precision is $C'(\alpha)$. Lemma 3 illustrates that in the pooling market, all the marginal benefit comes from the improvement in selection while the interest rate is constant. As such, $C'(\alpha)$ determines $\gamma'(\alpha)$. Alternatively, across the CIM markets the marginal benefit comes from higher interest rate while the selection is constant. As such, $C'(\alpha)$ determines $r'(\alpha)$.

When α distribution is exogenous, the equilibrium profit level π determines the aggregate capital, W, which is endogenous. In this case, in the CIM segment, market clearing implies that $r'(\alpha) = w(\alpha)$.

3.3 Changes to Incumbent Equilibrium

In this section, we consider changes to the incumbent equilibrium. We start by analyzing how a change in the nature of the baseline screening technology impacts the market structure.

Lemma 4 (**Loosening of Lending standards: Increase in β**). *Consider the entry equilibrium with no existing incumbents where lenders have cost $C(\alpha)$ and aggregate wealth W. The effect of an increase in the type I relative to type II error rate, captured by technological parameter β, on the entrant equilibrium structure of Proposition 2 is given by:*

$$\frac{dr_p}{d\beta} \geq 0; \quad \frac{dr_{NS}}{d\beta} \leq 0$$
$$\frac{d\alpha_i}{d\pi} \geq 0; \quad \frac{d\omega_i}{d\pi} \geq 0 \qquad \imath = 1, 2, 3 \qquad\qquad \frac{d\alpha_0}{d\beta} \geq 0; \quad \frac{d\alpha_1}{d\beta} \geq 0; \quad \frac{d\alpha_2}{d\beta} \geq 0$$
$$\frac{d\omega_1}{d\beta} \geq 0; \quad \frac{d\omega_2}{d\beta} \geq 0; \frac{d\omega_b}{d\beta} \leq 0$$

Lemma 4 argues that as the type I (false positive) error rate increases relative to type II (false negative) error rate, the credit market equilibrium continuously moves from extreme fragmentation to full pooling and no segmentation. The largest gap between r_p and r_{NS} corresponds to $\beta = 0$, when all the errors are false negative and no bad borrower can raises funding.

As β increases, lenders become prone to type II errors and those with less precise screening technology are exposed to default of bad borrowers in their portfolio, which increases the pooling interest rate r_p. An increase in r_p begets an increase in α_0. On the other hand these lenders absorb some of the bad borrowers, which improves the portfolio of the least precise $\alpha = 0$ lenders who lend non-selectively and leads to a decline in r_{NS}.

As $\beta \to 1$, $\alpha_2 \to \alpha_2 \to 1$, $\omega_1 \to \omega_2 \to 1$ and $r_{NS} \to r_p$. At the $\beta = 1$ limit, all the errors are type I. In this case, all the good borrowers can be identified by all the lenders. As such, there cannot be any difference in the rate that they borrow at. It follows that the credit market equilibrium is pooling throughout.

Next, we would like to understand how an increase in the wealth of incumbents without any change in the beseline screening technology or cost of acquiring precision impacts the nature of the equilibrium. To understand how the market structure changes when supply of wealth increases, it is insightful to consider how the equilibrium in Proposition 2 changes as lender profit level declines.

Lemma 5 (**Increase in Lender Profit Level**). *The effect of an increase in the profit level on the entry equilibrium structure when there are no existing lenders is given by:*

$$\frac{dr_p}{d\pi} \geq 0$$
$$\frac{d\alpha_i}{d\pi} \geq 0; \quad \frac{d\omega_i}{d\pi} \geq 0 \qquad i = 1, 2$$

This lemma is intuitive. An increase in lender profits is accompanied with an increase in the lowest interest rate, which is the pooling interest rate r_p. The lowest precision lender will

be more precise, so will the highest precision lender in the pooling region. As these lenders become more precise, they can identify more opaque borrowers correctly. Thus, the opacity of both marginal borrowers increase.

4 Entry Equilibrium

We next investigate the impact of new lenders entering the market. In particular, assume an *incumbent* equilibrium, characterized by Proposition has formed at the beginning of period $t = 0$ and let $w_0(\alpha)$ denote the incumbent wealth distribution, which can be endogenous or exogenous. We consider unexpected entry of a positive measure of new lenders at the end of period $t = 0$, before borrowing and lending happens at date $t = 1$. These new entrants are endowed with the same basic technology β as incumbent, but they have access to a new cost function $C(\alpha)$ to endogenously choose the precision of their technology α. Let W denote the exogenous aggregate wealth, and $w(\alpha)$ the endogenous wealth distribution of the new entrants.

In order to study the short run impact of entry, we assume the incumbent lenders cannot change their precision, α, in response to entry of the new lenders. However, they can lend at a different interest rate or exit the credit market if they choose to, i.e. if lending in every market leads to ex-post losses.

Proposition 4 (**Entry Equilibrium**). *Consider an incumbent equilibrium where lenders make profit π_0. A measure W of ex-ante identical new lenders enter with cost $C(\alpha)$ to acquire precision α.*

The unique entry equilibrium is heterogeneous in lender precision, every new lender makes endogenous profits $\pi(W)$, and every incumbent lender makes $0 \leq \pi_0' \leq \pi_0$ profits.

The equilibrium features a hockey-stick interest rate schedule, described in Definition 3.

Proposition 4 demonstrates that the hockey stick interest rate schedule is a very robust property of equilibrium when lenders have a nested information structure. The first point to notice is that when new lenders enter, the interest rate faced by every borrower weakly decreases. This is intuitive as new entry corresponds to an increase in supply of capital, which would decrease the interest rate at least in some segment(s).

The main idea to construct the equilibrium is to conjecture a profit level π for the new entrants corresponding to their exogenous aggregate wealth. We then construct an "effective cost function" for each group of lenders, $C_0(\alpha) + \pi_0$ for incumbents and $C(\alpha) + \pi$ for new entrants, and find the minimum of the two effective cost functions across the different degrees of precision, $\alpha \in [0, 1]$.

At points α where the minimum corresponds to the new entrants' effective cost function, the entry equilibrium should feature new entrants. The new entry will lead to a reduction in interest rate in that market. The existing incumbents with this α will continue lending in the same market unless if at the new interest rate, they make negative ex-post profits, for instance due to excessive defaults in their portfolio. Recall that by assumption, these incumbents cannot change their precision of screening. Furthermore, they endogenously choose not to switch to a market with a different interest rate. They do not switch to a

lower interest rate market because they have a higher precision than all the lenders in such markets, so they will be the last one to choose a portfolio, and the interest rate is lower than what they can get in the current market where they are first to choose. Alternatively, they will not switch to a higher interest rate market as they cannot recognize any good borrowers in such market.

The construction is straight forward when the new entry is in the cash-in-the-market pricing region of the incumbent equilibrium. The evolution of the new interest rate schedule follows the marginal cost of the new entrants. The incumbents at each α still make positive ex-post profit, as $\gamma(\alpha) = 1$ in the CIM region, so they continue lending. The wealth of new entrants should be such that along with the wealth of incumbents, it absorbs all the (higher) borrower demand at the new lower interest rate and the market clears.

Entry into the pooling region of the incumbent equilibrium is more complicated. In particular, depending on the extent of new entry, two different possibilities can arise. First, recall that by Lemma 2, the highest precision lender who lends in the pooling market must have $\gamma(.) = 1$, i.e. he must lend to a perfect portfolio of borrowers.

Next, note that new entry into the pooling regions at $\hat{\alpha}$ has two opposing effects on $\gamma(\alpha')$ for all the incumbents lender with $\alpha' > \hat{\alpha}$. First, new entry implies an increase in supply of capital in this market, which decreases the interest rate. The decline in interest rate pushes $\gamma(\alpha)$ down for all the lenders who lend on the pooling market due to the demand elasticity. However, Lemma 1 implies that entry of new lender wealth at any precision α increases $\gamma(\alpha')$ for $\alpha > \alpha'$. The relative position of the entrant and incumbent effective cost function determines the levels of α in the pooling region at which new lenders enter, and their indifference condition governs the new entrants' wealth. If it is the case that sufficient new entrant wealth enters that there is new entry at the level of $\tilde{\alpha}$ where $\gamma(\tilde{\alpha}) = 1$ for the first time in the new pooling region, then the pooling region switches to the cash-in-the-market pricing at $\tilde{\alpha}$ and the interest rate schedule in the CIM region starts by following the marginal cost of the new entrants.

Alternatively, if the new entrant wealth is limited and/or the improvement in cost function corresponds to a limited range of precision in the pooling region, the pooling region will have to include incumbents with higher precision than the last new entrant before reaching $\gamma(.) = 1$ and switching to CIM region. Note that in this case, the pooling segment will shorten. As no new lenders are entering at $\tilde{\alpha}$ where $\gamma(\tilde{\alpha}) = 1$ and the equilibrium segment switches, the CIM region in the entry equilibrium has to start with incumbents who used to lend in the pooling segment of the incumbent equilibrium—the *switching incumbents.* We call this new cash-in-the-market pricing range the "connecting CIM range" as it is constitutes the beginning of Region II in the entry equilibrium, and connects the pooling segment to the old cash-in-the-market pricing segment.

The wealth distribution of the switching incumbents determines the slope of the connecting CIM range in the enry equilibrium. If the wealth distribution of incumbents was determined endogenously, then the interest rates will be ironed in the connecting CIM range in order to ensure that less opaque good borrowers face lower interest rates. If the incumbent wealth distribution was exogenous, this condition is automatically satisfied.

The final step is to make sure that the total wealth of new entrants required to sustain the proposed equilibrium is consistent with their exogenous aggregate wealth, and reiterate

on the conjectured profit level.

We end this section by the benchmark entry equilibrium where entrants are endowed with the same cost function as incumbents, $C(\alpha) = C_0(\alpha)$.

Corollary 2 (**Entry Equilibrium with Homogeneous Technology**). *Consider an incumbent equilibrium with total incumbent wealth W_0 and cost function $C_0(\alpha)$, and measure W of ex-ante identical new lenders who enter with the same cost function $C_0(\alpha)$ and acquire precision α.*

The unique entry equilibrium is equivalent to an equilibrium with a single group of lenders with aggregate wealth $W + W_0$ and cost function $C_0(\alpha)$.

It has a hockey stick structure, described in Definition 3.

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.

We can also use Proposition 4 to explore short run change in the market structure when the screening technology of incumbents improves. To capture this situation, we assume a subset of incumbent lenders with total wealth W are randomly selected and can choose their screening precision α with the new cost function. In equilibrium, the rest of the lenders cannot choose their screening precision. However, they can lend in the different market or stop lending. We assume that the randomly selected incumbents get reimbursed for their upfront cost.[6] Lemma 6 illustrates that the market structure is robust to a chance in the incumbent screening cost.

Lemma 6 (**Improvement in Incumbents' Screening Cost: Short Run Equilibrium**). *Consider an incumbent equilibrium where lenders make profit π_0. A measure W of randomly selected incumbents get access to a new cost function $C(\alpha)$, are reimbursed for their old cost and and adjust their screening precision.*

The unique equilibrium is heterogeneous in lender precision and every incumbent who can adjust their screening precision makes endogenous profits $\pi(W)$. The incumbent lenders stuck with their old precision make potentially variable $\pi_0'(\alpha)$ profits, where $0 \leq \pi_0'(\alpha) \leq \pi_0$ $\forall \alpha$.

The equilibrium features a hockey-stick interest rate schedule, described in Definition 3.

In the next section, we explore more directed changes to the technology of new entrants.

[6]This assumption is to allow the incumbents to choose their new precision freely, without being contaminated with the fixed cost they have already paid. The result is robust to changing the assumption that they have to pay the cost difference. However, that exercise implies that no lender switches to a lower precision α even if the corresponding cost has substantially decreased.

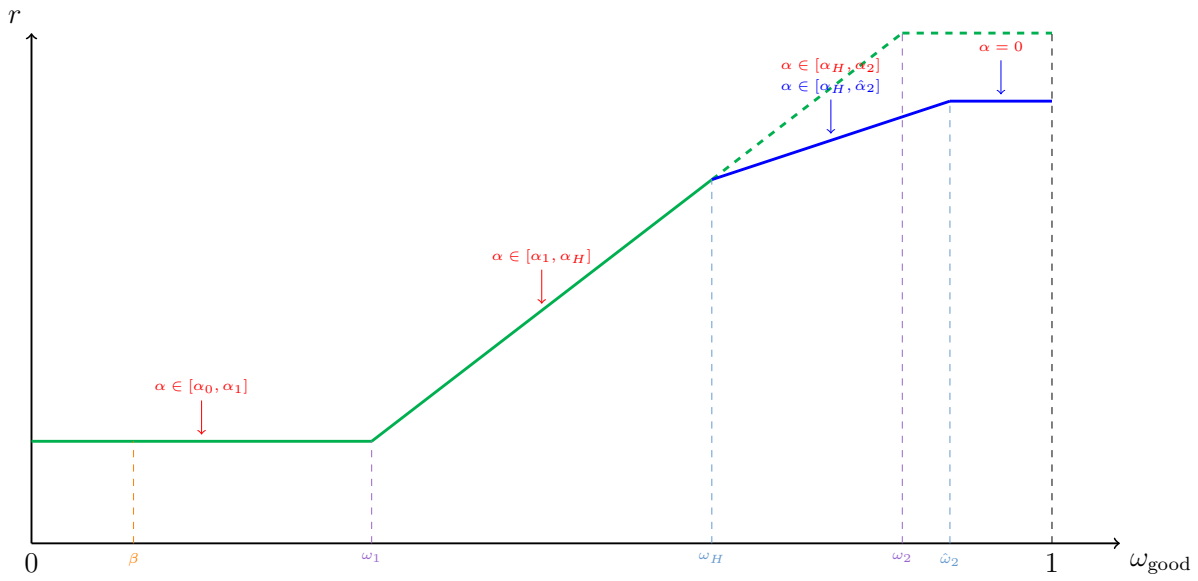


Figure 3: Impact of AI adoption in financial services on the equilibrium structure of the credit market.

5 Application: Equilibrium Spillover of Big Data

In this section, we use our model to study the consequences of growth in big data technologies as well as 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. As information asymmetry is a prevalent friction in the credit markets and screening is an important task of financial institutions, we believe that our model is particularly fit to investigate the impact of big data technologies and policies on the credit market. We would like to understand which borrowers and lenders benefit from each such type of technological progress, and in particular what forms of directed big data technological progress or policy leads to financial inclusion. We call the extent of this benefit the *equilibrium spillover of big data.*

Consider a reduction in the cost of a specific range of screening precision, i.e. a directed technological improvement that does not benefit all lenders equally. We focus on situations where the incumbents equilibrium features a continuous wealth distribution $w_0(\alpha)$ and new entrants have a limited supply of capital.

Let $C_0(\alpha)$ denote the cost function that corresponds to the incumbent wealth distribution. Recall that when new lenders enter, the incumbent lenders cannot change their precision, α. They continue lending as long as they make weakly positive ex-post profit. These results best represent the short/medium-run impact on equilibrium, as we assume the rest of the lenders are not able to change the level of technological sophistication α that they have already chosen. Or first exercise considers the impact of adoption of AI by financial institutions.

Adoption of AI in financial services When financial institutions such as fintech firms utilize artificial intelligence in lending, they use non-traditional data sources for borrowers to assess their creditworthiness above and beyond what the commonly used indicators such as credit score captures. In other words, use of artificial intelligence in financial services particularly improves lenders' inference about quality of borrowers who lack the commonly used measures of credit quality. It follows that in our framework, adoption of AI by financial firms translates to a reduction of the cost of screening the high ω borrowers who are highly opaque. These borrowers can only be successfully identified by lenders with a highly precise screening technology. As such, we model the use of a novel AI technology in financial services as a directed technological improvement which leads to a reduction in the marginal cost of precision for a high level of precision, $\alpha \in [\alpha_H, \bar{\alpha}_H]$, where $\bar{\alpha}_H \leq 1$.

Corollary 3 (**Adoption of AI in Financial Services).** *Adoption of AI technology in financial services improves the financial inclusion of under-served borrower population.*

Figure 3 illustrates the result of Corollary 3 graphically. We are interested in the case where the capital of the new lenders is limited, so their profit level is high in equilibrium. Furthermore, the reduction in the marginal cost of screening is concentrated in the interval $\alpha \in [\alpha_H, \bar{\alpha}_H]$. It follows that the new entrants' effective cost function can constitute the lower bound of the effective cost function in the entry equilibrium only in the range $\alpha \in [\alpha_H, \bar{\alpha}_H]$, which means the good borrowers with $\omega \in [\omega_g(\alpha_H; \beta), \omega_g(\bar{\alpha}_H; \beta)]$ will benefit. More importantly, the incumbent low-tech lenders with $\alpha = 0$ who were competing on the non-selective segment with the high-tech lenders to lend to very opaque good borrowers now face a more fierce competition, which forces them to offer better rates in equilibrium. As such, the rate faced by the most financially-excluded borrowers with $\omega \in (\omega_g(\bar{\alpha}_H; \beta), 1]$ also reduces, despite the fact that the technological improvement does not directly target them.

As such, the benefits of AI adoption in the financial sector spills over to the most opaque borrowers and improves financial inclusion.

Mandatory data sharing policy A mandatory data sharing policy requires financial institutions to share the data of their customers with other institutions if the customer requests so. In other words, data sharing policies reduce the cost of data collection for borrowers whose data is already in the financial system, albeit at a different institution. Open Banking regulations are a prime example of this type of policy. They are implemented in multiple countries including the UK, Brazil and India and discussed in the US. In our framework, these borrowers most closely correspond to the more transparent borrowers whose data is widely available—low ω borrowers. As in our model these borrowers are identified by many lenders but end up borrowing from the low precision lenders, we model the implementation of a mandatory data sharing policy in financial services as a directed technological improvement which leads to a reduction in the marginal cost of precision for low levels of precision, $\alpha \in [\alpha_L, \alpha_H]$, where $\alpha_0 < \alpha_L < \alpha_H < \alpha_2$.

Corollary 4 (**Mandatory Data Sharing Policy).** *The equilibrium spillover of implementation of a mandatory data sharing policy in the credit market is limited to well-served borrowers. Such policy has no financial inclusion benefits and instead exacerbates the inequality in financial access.*

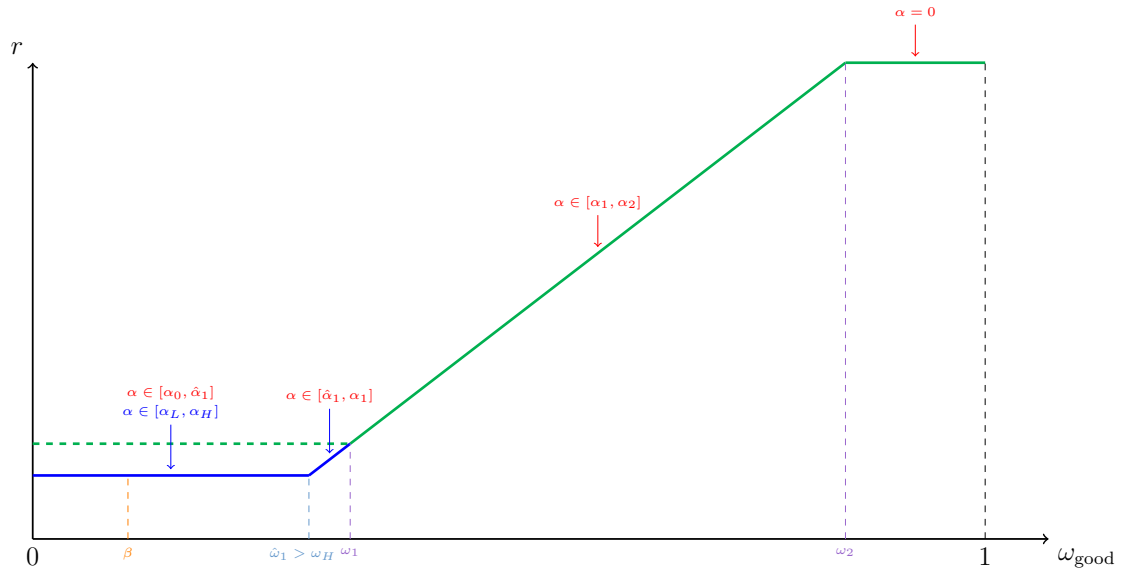


Figure 4: Impact of mandatory data sharing policy in credit markets on the equilibrium structure of the credit market.

Figure 4 illustrates this result graphically. We again focus on limited entry of lender capital, which implies that new entry happens only within the interval $\alpha \in [\alpha_L, , \alpha_H]$. The increase in capital supply in this region leads to a decline of the pooling interest rate. Furthermore, as the wealth of new entrants is limited, it follows from Proposition 4 that some incumbents with $\alpha \in [\alpha_H, \alpha_2]$ participate in the new pooling region with the lower interest rate as well. The pooling segment then transitions into a "connecting CIM range," financed by switching incumbent lenders, and extends all the way to the old CIM segment.

All the highly transparent good borrowers who were already served in the pooling segment of the credit market benefit, as evident in Figure 4. This is independent of whether the marginal cost of screening them has reduced or not. The reason is that lenders with moderate technology levels compete for them and pool them with direct beneficiaries of the data sharing policy. Some of these borrowers are now served in a fragmented segment of the market, the connecting CIM range. However, the rate that they face still declines.

Alternatively, such policies have no spillover to the financially excluded borrowers. Intuitively, these borrowers are only served by non-selective lenders who have not generated much useful data about them. As such, they will not benefit from a data sharing policy.

It is most useful to compare Corollary 4 with Corollary 2, which considers the entry of new lenders without any technological improvement. Corollary 2 demonstrates that without any technological improvements, all of the borrowers benefit from an increase in the supply of capital. Contrasting this finding with the impact of a mandatory data sharing policy in Corollary 4 reveals an adverse ex-ante effect of such policies. They divert investment in screening technology from serving the financially excluded population towards the less-costly, already well-served borrowers, and thus amplify the inequality in financial access.

Lastly, note that lenders with high levels of screening precision happily join such manda-

tory data sharing policies and share data of highly transparent, low-ω, borrowers who they can well identify. The reason is that the high-tech lenders do not serve that segment in equilibrium, so data sharing does not expose them to any competitive pressure. Furthermore, as Corollary 4 shows, there is not spillover to the borrowers who they serve.

6 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:*

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

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,

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

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 1 and 2 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 fragmantion 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.

7 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

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.

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

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.

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

Proof of Proposition 1.
A hockey-stick stick equilibrium where lender distribution is exogenous is defined by the following set of equations:

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

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 Definition 3 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 Definition 3, 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 (9), (10) and (11), respectively. Furthermore, Equation (6) defines the quality of borrowers for lender with precision α and Equation (14) defines th 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 (16) and (17), 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 $b\left(1-\beta+\alpha\beta;\alpha\right)$ in the last step is **not** the original density. Finally, add up to get $T'(\alpha)$:

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

$$T'(\alpha)=-\frac{w\left(\alpha\right)}{d\left(r_p\right)}+\left(1-\beta\right)g(\beta+\alpha\left(1-\beta\right))-\beta b\left(1-\beta+\alpha\beta;\alpha\right) \tag{18}$$

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 (12). Since γ_0 is increasing in α_0, then r_p is decreasing in α_0.

Define $T\left(\alpha,\alpha_0\right)$ as is the solution to the differential equation (18) with initial condition α_0. $T\left(\alpha,\alpha_0\right)$ 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 $\alpha_1\left(\alpha_0\right)$ denote the lowest solution to

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

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, $T\left(\alpha_1\left(\alpha_0\right),\alpha_0\right)=0$.

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. *$T\left(\alpha,\alpha_0\right)$ 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 (9), (10), (11), (14), (16) and (17).

Base step The base step is for $\alpha_h = \alpha_0^m$.[7] 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 (9), (10), (11). 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 (16) and (17), 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 (9), (10) and (11) 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.[8]

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 (16) and (17), 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

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

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

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 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 $T'(\alpha_1) \not< 0$. 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 1, $\frac{dw(\alpha)}{d\alpha} < 0$ 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:

$$dW(\alpha) = w(\alpha)d\alpha = d\left(r(\alpha)\right) g(\beta + \alpha(1-\beta))(1-\beta)d\omega.$$
$$r(\alpha) = d^{-1}\left(\frac{w(\alpha)}{g(\beta+\alpha(1-\beta))(1-\beta)}\right) \tag{19}$$

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 2, 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

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

□

Proof of Proposition 2.
A hockey-stick stick equilibrium where lender distribution is endogenous is governed by the same following set of equations as when the lender wealth distribution is exogenous, with two changes.

First, the last equation that is replaced by

$$r'_{CIM}(\alpha) = C'(\alpha) \qquad \alpha \in [\alpha_1, \alpha_2].$$

Second, for any exogenous profit level π, at each α where positive lender wealth enters, the lenders have to break even

$$\gamma(\alpha)(1 + r(\alpha)) - C(\alpha) = \pi \qquad \forall \alpha | w(\alpha) > 0$$

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

$$\begin{aligned}
&\alpha_0 = \arg\min_{\alpha} r_p(\alpha) + 1 = \frac{\pi + C(\alpha) + 1}{\gamma_0(\alpha)} \\
&\gamma_0(\alpha_0)(1 + r_p) = \pi + C(\alpha_0) + 1 \\
&r_p - C(\alpha_1) = \pi \\
&\alpha_2 = \arg\min_{\alpha} \gamma^{NS}(\alpha)\left(1 + r_{CIM}(\alpha)\right) \geq 1 + \pi \\
&\omega_i = \beta + \alpha_i(1 - \beta) \qquad i = 0, 1, 2
\end{aligned}$$

The first two equations need to be solved together. The next five equations are then solved one-by-one in order.

We construct the three equilibrium regions consecutively

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 (12).

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
   $$\alpha_1 = \begin{cases} 1 & \text{if } (1+r_p) - 1 - C(\alpha_1) > \pi \text{ for all } \alpha \\ C^{-1}(\pi - r_p) & \text{otherwise} \end{cases}$$
5. 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)
     - 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}. [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}]
   - 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 $W(A) = \lim_{\Delta \to 0} W_\Delta(A)$
6. For $\alpha > \alpha_1$, find
   - $r(\alpha)$ by the indifference condition
   $$r(\alpha) - C(\alpha) = \pi$$
   - The density $w(\alpha)$ by the condition
   $$w(\alpha) = d(r(\alpha))\, g(\beta + \alpha(1-\beta))(1-\beta)$$
7. Find the point of non-selective entry, if it exists
   - 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:
  - W has a mass point ot $\alpha = 0$, 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)
  - $r(0) = r_{NS}$ (i.e. unskilled lenders go to market r_{NS})
  - $s(0) = 1$ (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.

8. Compute the total mass of entrants W
9. Steps 1-8 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.

The equilibrium is therefore:

1. The measure W defined by the construction above
2. Choice of markets and selectiveness

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

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

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

Finally, 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

□

Proof of Lemma 5.

First, we show that the lower bound of pooling, α_0, and r_p are increasing in π. Let

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

$R(\alpha, \pi)$ is the answer to the question: "if you are first in line and need to make profits π, what interest rate do you need?" For a given π, α_0 solves

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

with FOC and SOC

$$R_1(\alpha_0(\pi), \pi) = 0$$
$$R_{11}(\alpha_0(\pi), \pi) > 0$$

Now use 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, implies

$$\alpha_0'(\pi) > 0.$$

What's driving this is that profits are

$$\pi(r, \alpha) = (1 + r)\gamma_0(\alpha) - 1 - C(\alpha)$$

The cross-partial of the profit function with respect to $r$and α is positive. Furthermore, the profit function is supermodular, i.e., raising the interest rate benefits higher-α lenders more than lower-α lenders. Therefore as we increase r the profit-maximizing α increases

Second, we show that the upper bound of pooling, α_1, is also increasing in π The upper bound is given by:

$$\underbrace{R\left(\alpha_0\left(\pi\right),\pi\right)}_{\text{pooling interest rate}} - C\left(\alpha_1\left(\pi\right)\right) = \pi$$

Take derivatives on both sides:

$$\underbrace{R_1\left(\cdot\right)}_{=0 \text{ by FOC}} \alpha_0'\left(\pi\right) + R_2\left(\cdot\right) - C'\left(\cdot\right)\alpha_1'\left(\pi\right) = 1$$

so

$$\begin{aligned}\alpha'\left(\pi\right) &= \frac{R_2\left(\cdot\right)-1}{C'\left(\cdot\right)} \\ &= \frac{\frac{1}{\gamma_0(\alpha)}-1}{C'\left(\cdot\right)} > 0\end{aligned}$$

Since the pooling interest rate goes up by more than the profit level (to make up for $\gamma < 1$ for the first entrant), the old α_1 does not need $\gamma = 1$ to remain indifferent, so the pooling region expands to the right

□

Proof of Proposition 4.
Suppose we have constructed an equilibrium with endogenous entry giving the measure $W\left(\alpha\right)$, implied by cost function $C\left(\alpha\right)$ and resulting in profit π. Here we consider the construction of a new equilibrium where a group of entrants with cost function $C^E\left(\alpha\right)$ and in total mass $\bar{W}^E$ enter at various markets determining the equilibrium measure $W^E\left(\alpha\right)$. (Note that the equilibrium construction is the same regardless whether the incumbents' $W\left(\alpha\right)$ is exogenously given or constructed as above.) We proceed in steps.

1. Fix π^E
2. Define $\gamma\left(\alpha; r, \tilde{W}, \tilde{W}^E\right)$ as the $\gamma\left(\cdot\right)$ function at α implied by an arbitrary pooling interest rate r, when the measure $\tilde{W}$ specifies the entrants for each market α from the incumbent group and $\tilde{W}^E\left(\alpha\right)$ specifies the entrants from the new group. For instance, $\gamma\left(\alpha; r, 0, 0\right) = \gamma_0\left(\alpha\right)$ while $\gamma\left(\alpha; r_p, W, 0\right)$ is the equilibrium $\gamma\left(\cdot\right)$ we have constructed in the incumbent equilibrium. (Zeros for the last two arguments stand for the zero measure.) That is, $\gamma\left(\alpha; r_p, W, 0\right) = 1$ for all $\alpha_{NS} \geq \alpha \geq \alpha_1$.
3. Find the (potential) marginal entrant in the pooling region:

$$\alpha_0^E = \arg\min_{\alpha \leq \alpha_1} \frac{\pi^E + C^E\left(\alpha\right) + 1}{\gamma\left(\alpha; r_p, W, 0\right)} - 1$$

4. Find the (potential) pooling interest rate at entry:

$$r_p^E = \frac{\pi^E + C^E\left(\alpha_0^E\right) + 1}{\gamma\left(\alpha; r_p, W, 0\right)} - 1.$$

If $r_p^E \geq r_p$ there is no entry in the pooling region: $W^E(\alpha) = 0$ for $\alpha \leq \alpha_1$. In this case, $\alpha_1^E = \alpha_1$, and the construction continues the previous step.

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

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

for all α. (Note that for $\alpha \geq \alpha_1$ $\hat{r}(\alpha) = C(\alpha) + \pi$ by definition.)

6. find $\alpha_1^{E,\min}$, by the indifference condition:

$$r_p^E - C^E\left(\alpha_1^{E,\min}\right) = \pi^E.$$

7. find α_1^E , and W^E by discretizing, iterating and then taking the limit. In particular, for any fixed Δ we are specifying steps to build up a series of mass points $\left\{w_0^E, w_1^E ... w_n^E\right\}$ at the corresponding markets $\{\alpha_1, \alpha_2, ...\alpha_n\}$ which define a measure W_Δ^E, and an end point for pooling $\alpha_{1,\Delta}^E$. Taking the limit $\Delta \to 0$ gives α_1^E , and W^E.

Note that by assumption,

$$\left(1 + r_p^E\right)\gamma\left(\alpha'; r_p^E, W, 0\right) - 1 - C^E(\alpha') \leq \pi^E \quad \text{for all } \alpha' > \alpha_0^E$$

(i.e. no one can make π^E with the incumbent distribution). If a $\tilde{W}^E$ has a mass point $\tilde{w}^E$ at α, then for all $\alpha' > \alpha$, the quantity $\gamma\left(\alpha'; r_p^E, W, \tilde{W}^E\right)$ is increasing in w, and reaches $\gamma\left(\alpha'; r_p^E, W, \tilde{W}^E\right) = 1$ for some finite w^E.

Now, fix Δ. For each n, starting from $n = 0$, suppose W_Δ^E has a mass point w_n^E at α_n. Find w_n^E such that

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

where we define $W_\Delta^{E,n}$ as the measure where $w_0^E, w_1^E, ...w_{n-1}^E, w_n^E$ is specified by steps up to n, and $W_\Delta^{E,n}(\alpha) = 0$ for all $\alpha > \alpha_n$. Call the argmax α_{n+1}. 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 $\alpha_{n+1}]$

In each step, check if there is an $\alpha'' \geq \alpha_{n+1}$ for which $\gamma\left(\alpha''; r_p^E, W_\Delta^{E,n}, 0\right) = 1$. If yes, let this define $\alpha_{1,\Delta}^{E,n} = \alpha''$. If there is any $\alpha \in \left[\alpha_{1,\Delta}^{E,n}, \alpha_1\right]$ for which $\hat{r}(\alpha) < r_p^E$, then $\alpha_{1,\Delta}^E = \alpha_{1,\Delta}^{E,n-1}$ and $W_\Delta^E = W_\Delta^{E,n-1}$ and stop the process. (For sufficiently low Δ there will not be such α for $n = 0$.) Otherwise, continue until $\alpha_{n+1} \geq \alpha_1^{E,\min}$ and let $\alpha_{1,\Delta}^E = \alpha_1^{E,\min}$ and $W_\Delta^E = W_\Delta^{E,n+1}$.

For any subset $A \subseteq [\alpha_0, \alpha_1]$, let $W^E(A) = \lim_{\Delta \to 0} W_\Delta^E(A)$ and $\alpha_1^E = \lim_{\Delta \to 0} \alpha_{1,\Delta}^E$.

If $\alpha_1^{E,\min} < \alpha_1^E < \alpha_1$ 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 . With an abuse of notation, let $\hat{r}(\alpha)$ be the ironed version.

For $\alpha > \alpha_1^E$, let

$$r_{CIM}(\alpha) = \min\left(\hat{r}(\alpha), C^E(\alpha) + \pi^E\right).$$

Find the point of non-selective entry, if it exists. Let α_2^E defined by

$$\frac{1}{d\left(r_{CIM}\left(\alpha_2^E\right)\right)}\left[G\left(\left[0,1\right];r_{CIM}\left(\alpha_2^E\right),0\right)+B\left(\left[0,1\right];r_{CIM}\left(\alpha_2^E\right),0\right)\right]=$$
$$=\frac{1}{d\left(r^{NS}\right)}\left[G\left(\left[0,1\right];r^{NS},0\right)+B\left(\left[0,1\right];r^{NS},0\right)\right]$$

where r^{NS} is the non selective price at the incumbent equilibrium. The right hand side is the mass of non-selective entrants defined by the incumbent equilibrium. This condition pins down the upper limit of entry for new entrants by finding the interest rate at which iff incumbent non-selective lenders come in, they clear the remaining group of good lenders. If there is such an α_2^E then the new non sellective price is $r^{NS,E}=r_{CIM}\left(\alpha_2^E\right)$.

Entry in the CIM region, $\alpha\in[\alpha_1^E,\alpha_2^E]$ is

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

Finally, compute the total mass of entrants $W^E\left(1\right)$ for the given π^E. Then the previous steps define a decreasing function mapping π^E to total mass of entrants $W^E\left(1\right)$. The last step is to invert this function to find the level of π^E that is consistent with the exogenous total wealth of lenders $\bar{W}^E$.

The equilibrium with entry (for predetermined measure $W,\alpha_0,\alpha_1,r\left(\alpha\right)$)is therefore:

1. The measure W^E defined by the construction above
2. Choice of markets and selectiveness

$$r^E\left(\alpha\right)=\begin{cases}r^{NS,E} & \text{if } \alpha=0\\ r_p^E & \text{if } \alpha\in\left[\alpha_0^E,\alpha_1^E\right]\\ r_{CIM}\left(\alpha\right) & \text{if } \alpha\in\left[\alpha_0^E,\alpha_2^E\right]\end{cases}$$

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

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

□

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

- For each α
  - $G\left(\alpha\right)$ and $B\left(\alpha\right)$ are sizes of remaining pools when it's the turn of α
  - $r\left(\alpha\right)$ is interest rate in market that α visits
- Define $z\left(\alpha\right)=\frac{G(\alpha)}{G(\alpha)+B(\alpha)}$

- The quality of the pool faced by $\alpha$is:

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

  - For future reference, 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, $\gamma(\alpha) = \gamma(\alpha, \tilde{\alpha})$

- Lowest entrant: $\underline{\alpha}$

  - $z(\underline{\alpha}) = \frac{G}{G+B}$

- Profits for lender α:

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

- Evolution of G and B:

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

- Optimality condition for type α:

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

  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:

$$r'(\alpha) = -\frac{\left.\frac{\partial \gamma(\alpha, \tilde{\alpha})}{\partial \tilde{\alpha}}\right|_{\tilde{\alpha}=\alpha} (1 + r(\alpha))}{\gamma(\alpha, \alpha)} \tag{20}$$

- Terminal conditions: $\underline{\alpha}$ and $r(\underline{\alpha})$ have to be such that:

  - Indifference for marginal lender:

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

  - Exhaust good types:

$$G(1) = 0$$

□

Proof of Proposition 2-iid.

Construct the equilibrium as follows

1. Fix π
2. Find the lender who can lend cheapest and still make π
   - $z_0 \equiv \frac{G}{G+B}$
   - $\gamma_0(\alpha) \equiv \frac{z_0[\beta+\alpha(1-\beta)]}{z_0[\beta+\alpha(1-\beta)]+(1-z_0)\beta(1-\alpha)}$

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

3. Lender α_0 goes to market

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

4. Find W by discretizing
   - Fix Δ
   - For each n, suppose that W has a mass point $w(\alpha_n)$ at α_n
     - Careful about this: because things are iid, it's not the same to suppose there is one lender with a positive measure of wealth $w(\alpha_n)$ or many lenders with total mass $w(\alpha_n)$, because if there are many of them you have to figure out how to order them, and they won't be happy about it. I think it's easier to think of one lender with mass $w(\alpha_n)$. Then when we take the limit this shouldn't matter.
   - At the next market $\alpha_{n+1} = \alpha_n + \Delta$, 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}. To ensure this, let:
$$\gamma(\alpha, \alpha_{n+1}) = \frac{z(\alpha_{n+1})[\beta + \alpha(1-\beta)]}{z(\alpha_{n+1})[\beta + \alpha(1-\beta)] + (1 - z(\alpha_{n+1}))\beta(1-\alpha)}$$
    (i.e this is the quality you get if you have skill α and choose market α_{n+1}). Indifference requires:
$$\gamma(\alpha_{n+1}, \alpha_{n+1})(1 + r(\alpha_{n+1})) - 1 - C(\alpha_{n+1}) = \pi \tag{21}$$
  - lenders who already chose skill α_n are indifferent between visiting market α_n and visiting market α_{n+1}. Indifference requires:
$$\gamma(\alpha_n, \alpha_{n+1})(1 + r(\alpha_{n+1})) - 1 - C(\alpha_n) = \pi \tag{22}$$
  - Equations (21) and (22) are a system of two equations in two unknowns: $w(\alpha_n)$ and $r(\alpha_{n+1})$.
  - Interpretation: 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.

□