HONG KONG INSTITUTE FOR MONETARY AND FINANCIAL RESEARCH

BIG TECHS VS BANKS

Leonardo Gambacorta, Fahad Khalil and Bruno M. Parigi

HKIMR Working Paper No.29/2021

December 2021

HKIMR

HONG KONG
ACADEMY OF
FINANCE
金融學院

Hong Kong Institute for Monetary and Financial Research

香港貨幣及金融研究中心

(a company incorporated with limited liability)

Big Techs vs Banks

Leonardo Gambacorta
Bank for International Settlements

Fahad Khalil
University of Washington

Bruno M. Parigi
University of Padova

December 2021

Abstract

We study an economy in which large technology companies, Big Techs (BTs), provide credit to firms operating on their platforms. We focus on two advantages that BTs have with respect to banks: better information on their clients and better enforcement of credit repayment since BTs can exclude a defaulting firm from their ecosystem. When BTs have only a limited information advantage they enter the credit market and they are both more efficient than banks in screening firms ex ante, and more effective in reducing strategic defaults. When BTs have both superior enforcement and complete and private information of the firm type BTs can enter banks' turf only if they guarantee some *privacy* to firms by refraining from collecting some information and leaving some rents to them. BTs may share information by providing public information to banks or selling credit scoring to banks with different outcomes in terms of efficiency.

Keywords: Big Techs, credit markets, privacy, information sharing
JEL classification: E51, G23, O31.

* Email: (Corresponding author) Parigi: brunomaria.parigi@unipd.it, Gambacorta: leonardo.gambacorta@bis.org and Khalil: khalil@uw.edu

* The authors would like to thank Leonardo Madio, Han Qiu, the audiences at the University of Padova, the 2020 CefES Conference on European Studies, the Hong Kong Institute for Monetary and Financial Research, and the BIS for useful comments and suggestions. Bruno Maria Parigi gratefully acknowledges financial support from Hong Kong Institute for Monetary and Financial Research. This paper represents the views of the authors, which are not necessarily the views of the Hong Kong Monetary Authority, Hong Kong Institute for Monetary and Financial Research, or its Board of Directors or Council of Advisers, or those of the Bank for International Settlements. The above-mentioned entities except the authors take no responsibility for any inaccuracies or omissions contained in the paper.

1 Introduction

In the last decade large technology companies, also known as Big Techs (BTs), have entered in the provision of financial services. [1] Big Techs have become substantial players in payments in several advanced and emerging market economies (BIS, 2019). For example, Big Techs have come to account for 94% of mobile payments in China in the space of just a few years (Carstens et al, 2021). Big Tech credit grew by 40% in 2020 alone, to a global total of over US$700 billion. In some jurisdictions Big Techs participated in government credit schemes during the Covid-19 pandemic period (Cornelli et al, 2021).[2] Recently BTs have started competing with banks especially in the market for loans to small and medium sized enterprises (SMEs). This paper is about the competition between BTs and banks in the loan market for SMEs where adverse selection and difficulty to enforce repayment cause frictions. In this paper we will focus on loans to firms, in particular SMEs, because BTs mainly provide credit to small vendors in their online platforms or small firms using their payment apps via QR code.[3]

Big Techs learn massive amounts of data about the firms that sell through their online platforms or use the payment apps. While this information is valuable to improve the assessment of the credit risk, it can also be exploited by the lender as we know from the relationship banking literature (e.g. Sharpe, 1990). In the case of Big Tech lending, this problem is compounded by the fact that firms are somewhat captive in

[1] Big Techs are major digital players like Google, Amazon, Facebook, Apple and Microsoft in Europe and United States, and Baidu, Alibaba, Tencent and Xiaomi in Asia. Some venture into offering also financial services (e.g. Alibaba, Tencent). On the contrary, fintechs start as financial companies that rely heavily on technology to deliver financial services (e.g. P2P lending as in LendingClub). The distinction between Big Techs and fintech often disappears.

[2] The data show that globally, BTs credit is booming, overtaking fintech credit (Cornelli et al., 2019). The largest markets for BTs credit in absolute terms are China, Japan, Korea and the United States. China is the biggest market with BTs giants such as Ant Group operating also in the provision of wealth management and insurance products. In Japan, e-commerce firm Rakuten and social media company LINE are notable lenders. BTs credit is more developed when banking services are more expensive (higher banking sector mark-ups) and also where there is a larger un(der)met demand for financial services, as proxied by fewer bank branches per capita (Cornelli et al, 2020).

[3] As examples of Big Techs providing credit to small vendors in their online platforms see Alibaba's Taobao platform in China or Mercado Pago for Mercado Libre in Mexico. We neglect loans to large companies (not developed so far) and to households (granted mainly in the form of consumer credit). Although we do not target a specific institutional environment, our work is mainly related to Asia, Africa and Latin America, as regulation has somewhat limited the financial footprint of the Big Techs in Europe.

the Big Tech ecosystem. In fact, a default on a Big Tech loan may lead not only to the exclusion from future loans as in the case of bank lending but also to the exclusion from the platform's e-commerce (and thus from future sales) or from the payment system run by the same Big Tech. We show how competition between banks and Big Techs in attracting borrowers can lead to greater *privacy* of borrowers as Big Techs have an incentive to temper their drive to collect information about firm characteristics. However, we also find that privacy may come at the cost of increased costly defaults and a loss of investment in profitable opportunities. One way to mitigate these inefficiencies is for Big Techs to share their data with the banks that make loans funded with cheap deposits. Importantly, we show that it is preferable to put in place mechanisms that guarantee privacy in the collection of information to share rather than simply provide more information to the public.

As Frost et al (2019) argue, Big Techs present a distinctive business model due to the combination of two key features: (i) network effects, generated by e-commerce, messaging applications, search engines, payment services, etc., and (ii) technology, e.g. artificial intelligence using big data and machine learning. Networks effects and technology lead to two characteristics — superior enforcement and superior information — that differentiate Big Tech lending from bank lending and will constitute two building blocks of our model.

First, Big Techs offering loans to firms that sell their products on their online platforms (or use their payment apps) have an advantage over banks in enforcing loans repayments and avoid voluntary defaults. The threat of exclusion — or even of a downgrade of reputation within a "captive" ecosystem — upon default provides Big Techs with an extra-legal but powerful contract enforcement tool.[4]

Second, Big Techs gain additional information about the firms from the huge amount of data that they collect on the platform (sales, product quality, reputation with clients) something that the banks cannot do. While a bank would learn imperfectly the firm's probability of being able to repay the loan also through the history of repayments, as is typical in relationship banking, a Big Tech would learn this probability much faster and much more accurately and with no human intervention.[5] This information enables

[4] Superior debt enforcement need not bring efficiency improvements if as Fong et al (2021) argue it leads to costly liquidation of assets.

[5] Big data obtained directly from Big Tech platforms typically include: i) transactions (sales volumes and average selling prices); ii) reputation (claim ratio, handling time and complaints); and iii) industry-specific characteristics (sales seasonality, trend and macroeconomic sensitivity). See Hau et al (2018)

Big Techs to screen clients more effectively than banks.

We have in mind an environment with limited enforcement of loan repayment. For example, it could be that collateral is not available, and/or the efficiency of the judicial system is low, and/or there are prohibitive costs to enforce repayments, and/or the loan size is too small to make the fixed cost of enforcement worthwhile. Model-wise, this implies a scope for *strategic defaults*, i.e., a firm may choose to default even when it has enough cash flow to repay a loan. Thus, the requirement to induce firms to repay imposes an upper limit on the interest rate that the lender (bank or Big Tech) can charge. Strategic defaults by solvent firms are a key measure of inefficiency as such firms could profitably self-finance their investment if not for the fact that they require external funding.

We focus on the trade-off between data privacy and efficiency. Data privacy refers to the information that Big Techs learn about borrowers, and what they do with it, including sharing it with other agents. Efficiency has two dimensions: lending to solvent firms only and reducing their strategic defaults.

We consider cases such that the Big Tech has different degrees of information advantage over the bank. It may enjoy a limited information advantage by knowing only whether the firm is solvent before lending. It enjoys a stronger information advantage when, besides solvency information, after lending it receives sufficient information on the firm — technically a better signal about the credit worthiness of the firm — that allows learning the probability of success of its investment. In the latter case, once it has a captive user base, the Big Tech can jack up the price of its financial services for its users to extract a larger share of customer surplus. Of course, potential clients anticipate this scope to abuse information, which introduces an incentive compatibility constraint on a Big Tech based on the information that it has received from data processing.

Our model allows us to shed some light on how the Big Tech could be made to

---

and Frost et al (2019) for more details.

The Big Tech will know if the retailer or manufacturer enjoys low or high product return margins and be able to infer from customer reviews the quality of products or service supplied (Zetzsche et al, 2017). As Frost et al (2019) argue, due to their extensive use of artificial intelligence, Big Techs may be able to better organise and process the data, relative to banks. The superiority of Big Techs in organising the data from different sources allow them to construct comprehensive databases to assess customers' preferences and behaviours. Big data can then be processed through machine learning algorithms that establish correlations between client-specific characteristics/preferences and creditworthiness, so as to provide a much more precise assessment of credit-worthiness than traditional banks do.

share the vast reams of data they have collected on firms. This aspect is important because there is an ongoing debate on trade-offs and limits of alternative information arrangements between Big Techs and banks. Options also include the possibility of special separate credit scoring joint-venture that would be partly state-owned (Yu and McMorrow, 2021). In particular, we investigate two information-sharing arrangements. In one, the Big Tech makes data public for any bank that wants to use them, e.g., by conferring the firm type information to a public credit bureau. In the other arrangement, the Big Tech gives privately the firm type information to the bank, e.g., by selling it credit scoring services.

Although, apparently similar, these two ways to share information lead to different outcomes. When banks compete for firms, providing public information to banks, they end up rationing solvent firms (if the judicial system is not able to fully avoid strategic defaults), while sharing information privately exploits all gains from trade.

The different outcomes stem from the fact that when information is learned privately (credit scoring) cross-subsidization between solvent firms takes place and the bank breaks even only on average. With a credit bureau, on the contrary, information becomes public and competition based on public information destroys cross-subsidization so that a bank must break even on each type to which it lends. This implies that with a credit bureau (in case of limitations of the judicial system) risky but solvent firms would default strategically because their break-even rate would exceed the no-default rate; hence they are rationed.

To preview our results, more powerful retaliation after default increases welfare as it reduces strategic defaults by solvent firms (Lemma 1). When BTs have only a limited information advantage they enter the credit market, and they are both more efficient in screening firms ex-ante — which translates in fewer defaults — and more inclusive as they reduce strategic defaults by solvent firms (Proposition 1). When, on the contrary, Big Techs have both superior enforcement and complete and private knowledge of the firm type, no firm will borrow from them anticipating extraction of the continuation value (Proposition 2). Big Techs can enter bank's turf if they guarantee some *privacy* to firms by refraining from collecting some information and leaving some rents to them (Proposition 3). Finally, it matters how Big Techs may share information with banks: providing more public information to banks does not guarantee that all solvent firms will continue to invest, while sharing information privately will exploit all gains from trade (Proposition 4).

The remainder of the paper is organised as follows. In Section 2 we review the related literature, in Section 3 we set up the model, in Section 4 a representative bank and a representative Big Tech compete when the Big Tech has a mild information advantage, in Section 5 the Big Tech has full information advantage and collects noisy information to be able to enter bank's turf. In Section 6 we study information sharing from the Big Tech to the bank. Section 7 discusses some extensions and concludes.

2 Related literature

Our paper relates to the literature on "relationship" vs "transactional" lending. The various strands of this literature focus on different and interconnected roles for relationship banks (or R-bank for short) and transactional banks (T-bank). Big Tech lending has characteristics of both lending types, because the loan offer to the client follows a period of interaction on the platform (similarly to R-bank), but at the same time the cost (for the Big Tech) of a termination of the relationship with the client is quite limited (T-bank).

Our paper based on a learning mechanism is very much related to the stream of the relationship literature that emphasises (soft) information acquisition about borrowers' types over time (Sharpe, 1990; Rajan, 1992; Von Thadden, 1995; Bolton et al, 2016). This strand of theories puts the R-bank in the position of offering continuation lending terms that are better adapted to the specific circumstances in which the firm may find itself in the future. We also add to the relationship banking literature the dimension of the superior enforcement of loan repayments that follows from the fact that borrowers are somewhat captive in the Big Tech ecosystem.

Another point of contact of our paper is with the literature on fintech lending and the capacity of credit scoring based on machine learning and big data to better assess firms' credit worthiness. Fintech credit is typically based on peer to peer (P2P) platforms that facilitate the direct matching between a borrower and a lender (See Belleflamme et al, 2016 for a review of the literature). This kind of credit is different from Big Tech credit offered to firms operating on an e-commerce platform or using BT's payment app. Fintech lenders do not raise funds and do not retain credit risk, their sources of income being only the fees paid by the borrowers and the lenders. However, fintech credit is based on credit scoring models that use machine learning and non-traditional data as in the case of Big Tech credit. In particular, a few studies have analysed how credit supplied by fintech firms, and their scoring models perform compared with traditional bank lending. Jagtiani and Lemieux (2018) compare loans made by a large fintech lender and similar loans that were originated through traditional banking channels. Specifically, they use account-level data from LendingClub and the Y-14M data reported by bank holding companies with total assets of US$50 billion or more. They find a high correlation between interest rate spreads, LendingClub rating grades (that use non-traditional data) and loan performance.

A number of papers have studied competition between Big Techs and traditional banks. Hau et al (2019) construct a model where Big Techs use data from vendors and consumers online trading for credit analysis. Their main prediction is that Big Tech credit is relatively more attractive for borrowers with low credit scores who are often excluded from bank credit. This prediction is supported by the empirical analysis based on credit data from Ant Financial. Ant Financial uses the transaction data on its retail site Taobao (China's largest) to generate credit scores for the online vendors. They also find that there are substitution effects between Big Techs and traditional bank credit, but they vanish for low-quality borrowers without bank access.

Parlour et al (2020) study how the information from payment services allows BTs to enter the credit market previously dominated by incumbent banks. A common theme is that when BTs use consumer payment data to assess credit risk and provide lending, bank's pricing of loans becomes less informative about credit risk, and the quality of bank loans worsens. Consumers with weak bank relationships benefit (from cheaper access to electronic payment services), whereas consumers with strong bank relationships could benefit or hurt (depending on the change in banks' pricing of payment services).

In a spatial model of bank competition Vives and Ye (2021) study how the diffusion of information technology brought about by the entry of fintechs and Big Techs in credit markets affects competition. Improvements in information technology increase welfare if they weaken the influence of bank–borrower distance on monitoring/screening costs, which happens if banks have local monopolies.

Finally, our model is also related to the growing field of the economics of privacy (see Acquisti et al 2016 for a survey). This literature studies the economic value and consequences of protecting and disclosing personal information, and the trade-offs associated with the privacy and the sharing of personal data. We stress three dimensions that are relevant for our work. First, the rapid advance in information technology makes it feasible for sellers to price discriminate by conditioning their price offers on consumers' prior purchase behaviour. However, as Acquisti and Varian (2005) argue consumers are far from defenceless and it is likely that sellers will have to offer buyers some benefits to induce them to reveal their identities, to the point that under certain conditions sellers do not want to condition current price offers on past behaviour. Second, one theme of the line of research on privacy and price discrimination is that firms often benefit from committing to privacy policies. For example, Taylor (2004) argues that a company's privacy-intrusive strategies are counterproductive. He shows that even in the presence of tracking technologies that allow merchants to engage in price discrimination, regulation may not be necessary. If consumers are aware of how merchants may use their data and adapt their behaviours accordingly, it is in a company's best interest to protect customers' data. In line with this strand of literature our work shows that Big Techs have an incentive to commit to protect firms' data to compete against banks.[6] However, as He et al (2020) point out the voluntary nature of data sharing which is at the root of open banking may not be sufficient to protect consumer's welfare in credit markets plagued by adverse selection. Welfare could be reduced when the mere sign-up decision signals the credit quality. A third issue is the concern that more stringent data-protection regulations may lead to reduced access to credit, thus creating a trade off with consumer privacy. Pagano and Jappelli (1993) and Jappelli and Pagano (2002) show that if banks share information about their customers, they would increase lending to safe borrowers, thereby decreasing default rates.

3 Model set up

3.1 Investments

There are three periods: $t = 0, 1, 2$. At $t = 0, 1$ each firm has an investment opportunity of fixed size normalised to 1. Investments are observable. Firms have limited wealth, that we assume to be zero, to finance an investment at $t = 0$ based on third party financing. This is a typical feature of SMEs, characterized by a limited amount of outside equity invested in the company and no assets to pledge as collateral. For these types of borrowers, the only potential source of funds is a loan, a feature that we will assume in the model. Banks and Big Techs provide the loans competing in the credit market. As shown above (Cornelli et al, 2019) some Big Techs have ventured into lending, mainly to SMEs and consumers. Loans offered are typically credit lines, or small loans with short maturity (typically up to one year), rolled over after repayment.

[6] Our work is also linked to the broader issue of strategic ignorance (Carrillo and Mariotti, 2000) and to the optimal collection and sale of personal data (Calzolari and Pavan, 2006).

A firm's output per period is $Y > 0$ in case of success, and it is 0 in case of failure. We assume that Y is large enough that lending is profitable.

The opportunity cost of funds for a bank is 0, while Big Techs which do not have access to deposits, face an opportunity cost of $r \geq 0$. There is no discounting across periods, and all players are risk neutral.

We capture firms' heterogeneity by assuming that they have different probabilities of success p, with $p \in [0, 1]$, density function $f(p)$, and cumulative $F(p)$; the type p is known to the firm; the lender only knows $f(p)$. We allow for insolvent firms, i.e. with Net Present Value < 0, or $pY < 1$.

3.2 Repayment Enforcement

We consider an environment with limited repayment enforcement in the spirit of Bolton and Scharfstein (1990). Output is not observable to outsiders at any cost.[7] Our setting captures situations in which the judicial system is inefficient, or there are large fixed costs to assess outcomes as in the case of SME lending. Hence, when the output is Y, either a firm repays the loan voluntarily or defaults, which we call a strategic default. As we will see, a key welfare criterion in the model is the fraction of solvent firms that strategically default.

At $t = 0$, the lender makes a loan of size 1 that specifies a repayment R by the borrower at $t = 1$. After success, if the borrower repays, it is free to self-finance an investment of 1 again to obtain an expected payoff of pY in $t = 2$. If the borrower does not repay, the lender can prevent the borrower from investing again. Thus, we assume that reinvestment is observable and the lender can prevent it if the borrower does not repay in $t = 1$.[8]

Strategic default means that in case of success the firm keeps Y and saves on repayment. A firm's value of the retained Y depends on the enforcement ability of the

[7] This is a more extreme friction than in the Costly State Verification model (Townsend, 1979) where outsiders can verify the state of nature at a finite cost.

[8] Observe that repaying R and being allowed to self-finance a new investment of size 1 is equivalent to repaying $R + 1$ and receiving a roll over loan of 1. In our finite horizon setting, this would require commitment to lend upon repayment as the borrower would default for sure in the last period. However, while we observe loan roll over after repayment, we do not observe contractual commitments from Big Techs, or banks, to roll over. Thus, in our model, we do not rely on such a commitment assumption. Being able to prevent reinvestment after default provides sufficient incentive to a lender to offer an initial loan.

lender. We assume that the Big Tech has superior enforcement ability with respect to the bank in that the Big Tech can exclude a defaulting firm from future trades and from access to the payment system, so that a defaulting firm can at most consume Y. By contrast with income Y, a firm defaulting from a bank can conduct trades with net return ρY, where $\rho > 1$.

Thus, the parameter ρ measures the (negative of) the enforcement ability of the bank. Observe that the parameter ρ can be interpreted as reputation or *collateral* since it captures (the negative of) what the firm loses when it defaults at $t = 1$.[9] Note that ρ can also be interpreted as the (negative of the) differential benefit of operating in the Big Tech ecosystem.[10]

In sum, if the firm repays R it is allowed to reinvest 1 with its own funds and expect to earn pY in $t = 2$. If it defaults on a bank loan, it receives ρY in $t = 1$.[11] Thus, under a bank loan, the repayment R must satisfy the following incentive compatibility constraint for a firm of type p upon receiving Y:

$$Y - R - 1 + pY \geq \rho Y. \tag{IC}$$

The LHS is the firm's payoff from repaying R and investing 1 in a new project that will yield pY. The RHS is what the firm can obtain by strategically defaulting at $t = 1$. If a firm defaults on a Big Tech loan, it receives Y in $t = 1$ and the RHS of its incentive constraint is of course just Y, while the LHS is unchanged.

From (IC), we obtain the important cut-offs for bank and Big Tech loans, denoted,

[9] This resembles the problem of why a country may not want to default on its sovereign debt. Cole and Kehoe (1995) have studied the case of a country that may not want to default on its sovereign debt to avoid losing trade agreements.

[10] The mere use of Big Tech products and services could generate some sort of discount effects. In fact it is typical for Big Tech companies to use information obtained in their ecosystem to offer targeted discounts to their customers, although in general it happens at the expense of their competitors. For example, if clients use credit or other financial products on the payment platform Alipay in China or use a credit line, clients get "points" to be used to receive money and other free products. These benefits would be lost upon exiting the Big Tech ecosystem. Finally, besides exclusion from e-commerce and payment services the superior enforcement ability of the Big Techs can also be justified because in some instances they can seize the receivables of these companies in their accounts to repay their debts (Gambacorta et al, 2020).

[11] From our setting it follows immediately that the firm is 'locked-in' a two-period relationship with the lender and has no outside option at period $t = 1$. In fact, at $t = 1$ competition among lenders is mute. That is a firm would not be able to switch lenders for a one-period contract regardless of whether it has been cut off from another lender.

respectively,[12]

$$\widehat{p}_B = \frac{R+1}{Y} + \rho - 1, \tag{1}$$

$$\widehat{p}_P = \frac{R+1}{Y}, \tag{2}$$

such that a firm will repay R if and only if p is bigger or equal than its cutoff $\widehat{p}_i$, for $i = B, P$. Firms with p smaller than their cutoff will default strategically at $t = 1$ and will be prevented from investing again.

To illustrate the welfare impact of enforcement it is convenient to consider a benchmark with a representative lender with $\rho > 1$ and $r \geq 0$ subject to a zero expected profit condition,

$$\mathbb{E}(\Pi) = \int_{\frac{R+1}{Y}+\rho-1}^{1} pRf(p)\,dp - (1+r)\int_{0}^{1} f(p)\,dp = 0, \tag{3}$$

where R denotes the break even repayment. It is inefficient if R is so high that solvent firms with $p > 1/Y$ strategically default. Thus we will focus on the conditions that induce solvent firms to repay and continue in $t = 1$. The following Lemma establishes that tougher enforcement increases welfare as it reduces the fraction of solvent firms that strategically default.

Lemma 1. In economies where enforcement is tougher (i.e. where ρ is lower) both the repayment R satisfying the lender's zero expected profit (3) and the threshold probability below which firms strategically default $\widehat{p} = \frac{R+1}{Y} + \rho - 1$, decline.

Proof. See Appendix.

This result establishes that, everything else constant, an institutional environment that allows tougher retaliations against defaulters, discourages strategic defaults, lowers the break-even repayment, and ultimately is more inclusive and efficient. On the contrary, an institutional environment that limits the Big Techs' ability to retaliate and exclude from their ecosystem on the ground of protecting firms against powerful Big Techs, has the unintended consequence of encouraging strategic defaults by solvent firms, hence limiting their investment opportunities.

[12] We assume ρ is not too high such that these cut-offs remain below 1.

The difference $\rho - 1$ also captures the degree of exclusivity of the agreements between firms and Big Techs. At one extreme are agreements that force firms to sell through only one platform («choose one from two»). At another extreme a firm is excluded only from e-commerce of the Big Tech whose loan it defaults. In the Discussion and Conclusion section we discuss policy issues related to the post default options.

4 Competitive lending with Big Techs

We now move on to consider repayment competition between banks and Big Techs. We focus on a representative bank and a representative Big Tech assuming that each type of lender makes zero expected profits.[13]

As mentioned, Big Techs have information advantages with respect to banks stemming both from the massive amount of data and from technology. For example, using data for Mercado Credito, which provides credit lines to small firms in Argentina on the e-commerce platform Mercado Libre, Frost et al (2019) find that, when it comes to predicting loss rates, credit scoring techniques based on big data and machine learning have so far outperformed credit bureau ratings. A number of studies show that even digital soft information has informational content that enhances credit scoring.[14]

Here we distinguish between two Big Tech's information sets: the ability to distinguish whether the firm is solvent (solvency information) when it enters its ecosystem before lending, and the ability to identify its type (type information) after lending but before repayments.

[13] Since the credit market is ex-ante competitive, we also rule out the possibility to tie-in access to e-commerce to credit. A Big Tech at $t = 0$ cannot exclude from its e-commerce a firm that does not want to borrow from the very same Big Tech.

[14] Dorfleitner et al (2016) study the relationship between soft factors in peer to peer (P2P) loan applications and financing and default outcomes. Using data on the two leading European P2P lending platforms, Smava and Auxmoney, they find that soft factors influence the funding probability but not the default probability. Jagtiani and Lemieux (2018) find that the ratings assigned on the basis of alternative data perform well in predicting loan performance over the two years after origination. The use of alternative data has allowed some borrowers who would have been classified as subprime by traditional criteria to be slotted into "better" loan grades, enabling them to benefit from lower priced credit. In addition, for the same risk of default, consumers pay smaller spreads on loans from LendingClub than from credit card borrowing. Berg et al (2020) show that digital footprints are a good predictor of the default rate. Analysis of simple, easily accessible variables from digital footprints is equal to or better than the information from credit bureau scores.

First, we capture solvency information in our setting by assuming that by virtue of e-commerce and/or payment services, the Big Tech ex ante has a better idea than the bank of the distribution of the firms' types. In particular the Big Tech can assign loan applicants to two groups, solvent and insolvent, i.e. firms with $p \in [0, 1/Y)$, without identifying their true type though. Thus, the Big Tech at $t = 0$ can exclude insolvent loan applicants. At $t = 0$ the bank faces types $p \in [0, 1]$, with density $f(p)$ and cumulative $F(p)$ while the Big Tech faces types $p \in [1/Y, 1]$ with conditional density $\frac{f(p)}{1-F(1/Y)}$. Second, besides detecting solvent firms, after lending, the Big Tech may also identify their true type between $t = 0$ and $t = 1$. This allows the Big Tech to demand a type-contingent repayment $R(p)$ if its private signal about the firm's type is perfectly accurate. The rationale is that the loan relationship begets a deeper knowledge of the firm. Anecdotal evidence indicates that a Big Tech understands the firm's type early on in the ecosystem, after a couple of years of knowledge in the payment platform, independently of the use of the credit line.

We assume that the Big Tech acquires privately either information set, an assumption that, as we will show, matters for the architecture of information sharing.

As they compete in repayment, the Big Tech enters banks' turf by undercutting them. We assume that if the Big Tech lends, it serves at most a "small" fraction $\alpha \in [0, 1]$ of randomly selected firms of each type p, while the bank serves a complementary fraction.[15]. There are two main institutional reasons why the Big Techs' market share is restricted from what they could achieve through price competition with banks: Big Techs' debt capacity is quite limited as they cannot raise deposits lacking a banking license, and in some jurisdictions regulation and moral suasion limit their presence to some segments of the credit market.[16] Our model mimics these institutional

[15] We assume that the Big Tech market share is small enough that the bank's profit remains positive even when the Big Tech enters.

[16] Big Techs' relatively small lending footprint so far has reflected their limited ability to fund themselves through retail deposits. They could have the possibility to establish an online bank, but regulatory authorities could restrict the opening of remote (online) bank accounts. One relevant example is China, where the two Chinese Big Tech banks (Mybank and WeBank) rely mostly on the interbank market funding and certificates of deposit rather than on traditional deposits (Bank for International Settlements, 2019). Big Techs cannot issue virtual deposits which increases substantially their cost of funding (certificate of deposits and bonds are typically more costly than deposits). A second limitation is given by the fact that Big Techs cannot adopt a full originate-to-distribute model, partnering with banks. In principle, Big Techs could provide the customer interface and allow for quick loan approval using advanced data analytics; if approved, the bank could be left to raise funds and manage the loan. This option can be attractive to Big Techs as their platforms are easily scalable at low cost and they

characteristics and wants to shed light on policy choices that affect the market share and the impact on competition between banks and a Big Tech. A similar exogenous limit on market share is also used in Hau et al (2019) in their analysis of firms borrowing from their e-commerce platforms companies.

That Big Techs serve a small, albeit growing, fraction of loans fits stylized facts: for example, Frost (2020) shows that in 2017 the Big Techs market share was around 2.7% in China, 2.2% in South Korea, 1.65% in US, and 1.1% in UK. Similarly, for fintech, Hau et al (2019) observe that in China in 2016 fintech credit represented only 0.37% of all credit to SMEs; Frost et al (2019) shows that fintech firms extend less than 1% of global private sector credit. In 2017 fintech and Big Techs combined accounted for only 0.14% of the total assets of the global financial system (Frost, 2020).

4.1 Competition with solvency information

The amount of information available to the Big Tech plays a critical role in determining whether it can enter. First consider the case where the Big Tech is only able to detect privately whether loan applicants are solvent, i.e., whether $p \geq 1/Y$. The bank and the Big Tech compete in repayments, denoted by R^B and R^P, which are determined by their respective zero expected profit conditions. We will show that, with solvency information, the Big Tech will enter if the funding cost r is not too large and will be more inclusive ex post than the bank.

As the Big Tech will only lend to solvent firms, all insolvent firms with $p \in [0, 1/Y)$ have no choice but to borrow from the bank at R^B, and they default either strategically or because they have zero output. Note that any firm defaulting on a bank loan obtains an expected payoff $\rho p Y$. Thus, any firm in $p \in [1/Y, 1]$ planning to default, prefers to borrow from the bank because of the stronger outside post default option $\rho > 1$. Any firm planning to repay either lender, will choose to borrow based only on a comparison of repayment requirements as the outside option is not relevant for them.

Thus, we present the key results for the case of competition with only solvency information for the Big Tech.

---

interface directly with the client. However, regulation could limit this practice imposing retention requirements for joint lending with banks. Commercial banks in China must jointly contribute funds to issue internet loans with a partner, and the proportion of capital from their partnership in a loan should not be less than 30%. Moreover, limits on banks' internet loans relative to tier-1 capital are also in place.

Proposition 1. If r is not too large, the Big Tech enters with $R^P < R^B$, and there exists a $p_t \in (\widehat{p}_P, \widehat{p}_B)$, with $1/Y < \widehat{p}_P < \widehat{p}_B$, where,

- the Big Tech is more inclusive ex post, i.e., there is less strategic default by solvent firms ($\widehat{p}_B > \widehat{p}_P$) on Big Tech loans than on bank loans,
- all firms in $p \in [0, p_t)$ borrow from the bank at R^B and default either strategically or because they have zero output,
- a randomly assigned fraction α of firms with $p \in [p_t, 1]$ borrow from the Big Tech and do not default strategically,
- the fraction $1 - \alpha$ of firms with $p \in [p_t, 1]$ that the Big Tech does not finance, borrow from the bank; but only those with $p \in [\widehat{p}_B, 1]$ do not default strategically.

Proof. See Appendix.

Two reinforcing factors are at work. First, the superior enforcement ability of the Big Tech translates into a lower break-even rate and a lower probability threshold for repayment. Second, solvency information improves the risk of the loan applicants of the Big Tech, which lowers the break-even rate. These two factors allow the Big Tech to enter if its funding cost is not too high.

The lower Big Tech repayment rate reduces the incentive for solvent firms to default strategically. This makes the Big Tech more inclusive ex post than the bank. The Big Tech and the bank share the safest firms, while the riskiest firms borrow from the bank only.

Furthermore, since $R^P < R^B$ absent a cap to the market share of the Big Tech all safest firms, those with $p \in [p_t, 1]$, would borrow from the Big Tech while all the riskiest firms, those with $p \in [0, p_t)$, would borrow from the bank, if the bank could break even.

4.2 Competition with type information

Let us specify the timing. At $t = 0$ the bank faces applicants with types $p \in [0, 1]$ while the Big Tech faces only applicants with types $p \in [1/Y, 1]$ as it identifies and excludes insolvent firms. If the firm borrows from the BT, between $t = 0$ and $t = 1$, the

Big Tech also identifies its true type and demands a type-contingent repayment $R(p)$. Thus, anticipating $R(p)$, at $t = 0$ a firm of type p decides to borrow from the Big Tech or from the bank at a fixed repayment.

We assume that the Big Tech cannot commit at $t = 0$ to a repayment based on the type of the firm it will learn later. This implies that the Big Tech cannot commit to set a cap to the repayment. Since the Big Tech learns the information privately, firms understand that it may gain by overstating the probability of success if it can make a firm repay more.

We will first show that firms will not borrow from an all too powerful Big Tech that learns the firm type p perfectly and privately. Then, in the following section, we will introduce noisy private learning to show that the Big Tech will optimally choose not to learn the firms' type perfectly. This would be to counteract the effect of its strong enforcement ability. Imperfect learning aims to leave sufficient information rent to firms to enable the Big Tech to compete against the bank and draw clients.

The $t = 0$ contract provides for a type-contingent repayment $R(p)$, where p will be announced after the Big Tech learns a firm's type. Without loss of generality, we assume that Big Tech is induced to announce p truthfully. Thus, the repayment function $R(p)$ for the Big Tech satisfies the following Principal's Incentive Compatibility constraint, or (PIC).[17] It captures the fact that a firm anticipates that the Big Tech cannot stop itself from using the information it has acquired privately about the firm and that it will charge the highest repayment it can. It characterizes $R(p)$ as the maximum repayment that the Big Tech can demand from any type $p \geq 1/Y$. It is derived by replacing $R(p)$ for R in the firm (IC) :

$$Y - R(p) - 1 + pY = Y \Leftrightarrow pY - 1 = R(p), \text{ for } p \geq 1/Y. \qquad (PIC)$$

It follows that a firm with $p \geq 1/Y$, borrowing from the Big Tech, would be indifferent between repaying $R(p) = pY - 1$ at $t = 1$, or strategically defaulting. In other words, the Big Tech uses its available information to fully extract the firm's continuation value at $t = 1$ leaving the firm an expected payoff of pY from the first period of the relationship. Again, insolvent firms with $p < 1/Y$ can only borrow from the bank, defaulting strategically if successful, and they earn an expected return ρpY. Types planning to default, choose to borrow from the bank as the outside option ρ after default is

[17] See, e.g., Laffont and Martimort (2002, chapter 9.1), or Khalil et al (2015).

greater. Types planning to repay either type of lender compare the repayments of Big Tech and bank to realise that the bank provides a greater expected return.

To interpret the contract $R(p)$, observe that even if it may sound at odds with reality the fact that the borrower ignores the requested repayment when offered a loan, a borrower of type p will be able to anticipate at $t = 0$ what its repayment will be. Thus, we have the following result showing that too much information prevents Big Tech entry.

Proposition 2. If a Big Tech learns the firm type p *perfectly* and *privately,* all firms will borrow from the bank at the repayment $\underline{R}$ such that

$$\mathbb{E}\left(\Pi^{B}\right) = \int_{\frac{\underline{R}+1}{Y}+\rho-1}^{1} p\underline{R}f\left(p\right)dp - \int_{0}^{1} f\left(p\right)dp = 0. \tag{4}$$

Proof. See Appendix.

Two factors are at work here. First, the Big Tech cannot stop itself from using all the available information to demand a repayment rate to fully extract the firm's continuation value at $t = 1$. Model-wise this is captured by PIC. Second, the better post default options on a bank loan hurts the Big Tech. The combination of these two factors turns out to be too costly to the Big Tech ex ante.

Thus, we next move to studying the case where the Big Tech can commit to data privacy by choosing the precision of learning.

5 Noisy Signal

In this section, we argue that the Big Tech may want to learn the firm's type imperfectly to compete for the bank's clients. The precision with which the Big Tech learns about its clients depends on systems in place. We show that while the Big Tech needs to protect customer privacy and comply with relevant laws and regulations that limit the use of some information, it could be in its own interest to limit even further what it learns about its clients.

Financial intermediaries cannot include certain client specific characteristics in the information set to be used to train credit scoring models out of concern for issues of privacy and discrimination. For example, the US Fair Housing Act (FHA) and Equal Credit Opportunity Act (ECOA) prohibit credit scoring agencies from considering information like race, colour, religion, gender, marital status. The consumer credit scoring FICO voluntarily also excludes age, salary, occupation, title, employer, employment history, address. In either case the rationale is to avoid discrimination against applicants based on characteristics of groups with lower average scores. In our model instead, the rationale is to avoid extracting continuation values, i.e., strategically avoid price discrimination against, firms with high probabilities of success.[18]

To make our point in a stark manner, we assume zero (direct or physical) cost for the precision of learning. It would be straightforward to introduce costly learning without affecting our key insights.

We rely on a simple setting to illustrate our point. The intuition is that, to compete against the bank, the Big Tech chooses to acquire a limited amount of information at $t = 0$ that induces itself to offer a low enough repayment. In particular, limited learning enables the Big Tech to assure a client that it will not fully extract the firm's continuation value at $t = 1$. On this dimension, the Big Tech's advantage over the bank is given by $\rho - 1 > 0$. Thus to compete with the bank the Big Tech aims to leave a rent at least $(\rho - 1) Y$ to each type to which it lends. Thus, at $t = 0$, it sets up a technology to collect information that will generate a noisy signal ε about the firm type. This makes the Big Tech uncertain about the true type and prevents it from charging the maximum that a firm of that type can repay.

A model of noisy learning would work as follows. Each type $p \geq 1/Y$ knows that the Big Tech will draw a signal s about the firm from the interval $[p - \varepsilon, p + \varepsilon]$, such that $\mathbb{E}[s|p] = p$.[19] As in the case with perfect learning of type, the Big Tech's signal-contingent repayment $R(s)$ can be derived from a principal's incentive constraint, but modified to allow for noisy learning. The constraint is now denoted as (PIC'). Again,

[18] A related topic is unfair price discrimination. Sophisticated machine learning algorithms may not be as neutral as their mathematical nature suggests at first glance. Even though artificial intelligence and machine learning algorithms are neither trained nor fed with protected characteristics such as race, religion, gender, or disability, they are able to triangulate such information. Using data on US mortgages, Fuster et al (2019) find that black and Hispanic borrowers are disproportionately less likely to gain from the introduction of machine learning in credit scoring models, suggesting that the algorithm may develop differential effects across groups and increase inequality.

[19] For completeness we specify the expectation of signals at the top extreme by assuming that $E[s|p] = 1$ for all $p \in [1 - \varepsilon, 1]$. These latter types will not be relevant for the Big Tech in equilibrium. They will strictly prefer to borrow from the bank unless the noise is very large.

it captures the maximum repayment $R(s)$ such that no solvent type has an incentive to default strategically

$$Y - R(s) - 1 + (s - \varepsilon)Y \geq Y, \tag{PIC'}$$

which yields

$$R(s) = \max(0, (s - \varepsilon)Y - 1).$$

In addition to ensuring that a firm of type p does not default strategically, the repayment $R(s)$ must also leave no incentive to borrow from the bank. Thus, the expected repayment must have a large enough discount to attract as many solvent firms as possible from the bank.

Since the repayment $R(s)$ cannot be negative, some firms in $p \in [1/Y, 1/Y + \varepsilon]$ will prefer to borrow from the bank and strategically default. That is, there exists a firm of type $p_0 \in (1/Y, 1/Y + \varepsilon)$ which is indifferent from borrowing from the Big Tech at $R(s) = 0$, and borrowing from the bank and then defaulting strategically:

$$\underbrace{p(Y + pY - 1)}_{\text{borrowing from BT at } R(s)=0} = \underbrace{pY\rho}_{\text{borrowing and defaulting from bank}} \Leftrightarrow$$

$$p_0 = \frac{1}{Y} + \rho - 1. \tag{5}$$

This allows us to establish the following result:

Proposition 3. Under noisy and private type learning by the Big Tech, there exists both a $p_0 = \frac{1}{Y} + \rho - 1 \in (1/Y, 1/Y + \varepsilon)$, and a $p_b = \frac{R^B+1}{Y} + \varepsilon \in (\widehat{p}_B, 1)$ such that firms choose to borrow from the bank or the Big Tech depending on their types as follows:

- Firms with $p \in [0, p_0)$, (that we label Group 1) borrow from the bank.
- A randomly chosen fraction α of the firms with $p \in [p_0, p_b]$, (Group 2) borrow from the Big Tech at a signal-contingent rate

  $$R(s) = \max(0, (s - \varepsilon)Y - 1),$$

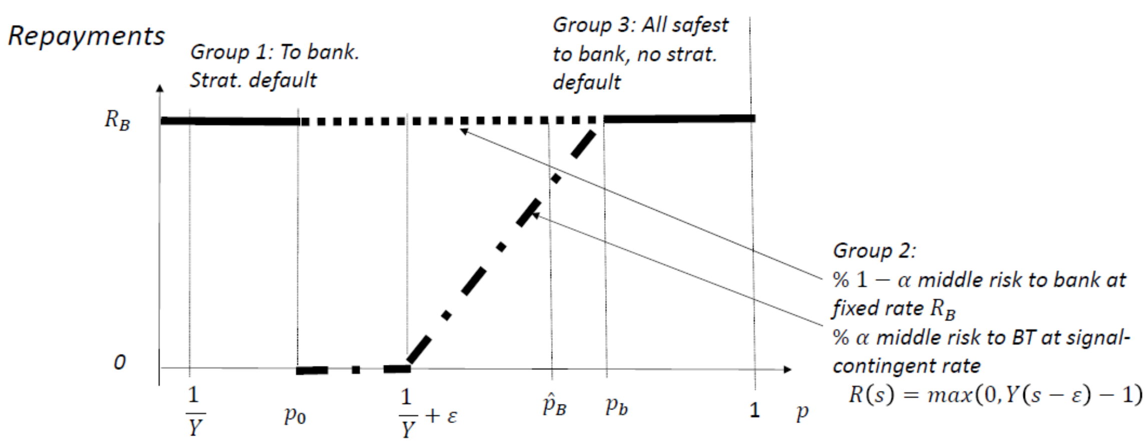
  and do not default strategically. The complementary fraction $1 - \alpha$ firms of Group 2 borrow from the bank at the fixed break-even rate R^B, which is determined by the solution of expected profit condition of the Big Tech and bank ((15) and (14) in the Appendix) but only those with $p \in [\widehat{p}_B, p_b]$, do not default strategically.

- All firms with $p \in (p_b, 1]$, (Group 3) borrow from the bank at R^B and do not default strategically.

- In equilibrium $\varepsilon > \rho - 1$, where ε is determined by the zero expected profit conditions of the Big Tech and bank ((15) and (14) in the Appendix).

Proof. See Appendix.

For an illustration of Proposition 3 see Figure 1.



$$p_0 = \frac{1}{Y} + \rho - 1 \qquad \hat{p}_B = \frac{R_B + 1}{Y} + \rho - 1 \qquad p_b = \frac{R_B + 1}{Y} + \varepsilon$$

Figure 1: Proposition 3. Riskiest and safest firms borrow from the bank. Middle risk firms split between Big Tech and bank.

Several comments are in order. First, our result is linked to one of the themes of the research on privacy and price discrimination, namely that firms often benefit from committing to privacy policies (Acquisti et al, 2016). Collecting anonymous data with aggregate, market-level information prevents the seller to set personalized prices (Bergemann et al, 2021). Federated machine learning plays a similar role by filtering information without revealing the identity.

Second, Proposition 3 points to inefficiencies both in initial lending and in strategic defaults by solvent firms. In particular, for firms in Group 3 and for a fraction α of the firms in Group 2, we have an efficient outcome whereby solvent firms do not strategically default. On the contrary, of the $1-\alpha$ firms in Group 2 that borrow from the bank, those with $p \in [\frac{1}{Y}+\rho-1, \widehat{p}_B)$, strategically default even though solvent. Furthermore, the bank funds insolvent firms in Group 1.

Third, our results fit stylized industry facts. It is widely accepted that the Big Techs are able to identify firm's types, which allows them to classify firms better than banks, particularly the riskiest firms (Frost et al, 2019). In Proposition 3 the riskiest firms are excluded from the Big Tech loans at $t=0$ but not from bank loans.

Finally, another stylized fact is that the safest firms borrow from the banks.[20] Indeed Group 3 firms prefer to borrow at a fixed rate from the bank rather than suffering the extraction of the continuation value from the Big Tech.

We now move on to explore information sharing from Big Techs to banks.

6 Information sharing

Big Techs and banks have complementary advantages. From e-commerce, Big Techs receive troves of data for free, while banks largely fund themselves with cheap deposits that the Big Techs cannot access. Lack of access to deposits makes Big Techs funding more costly than banks and it is a factor limiting their size. As we have shown in Proposition 3, this in turn, is responsible for inefficiencies: funding insolvent firms and strategic defaults by solvent firms. Thus, it seems natural to investigate whether it is possible to exploit these relative advantages so that the Big Techs gather and share data with the banks and the banks make loans funded with deposits.

To this end we consider two information sharing arrangements. We maintain that banks and Big Techs are subject to zero expected profits and that Big Techs are able to exclude insolvent firms before lending and learn the firm type fully and privately after lending. In one arrangement, the Big Tech makes data *public* for any bank that wants to use them, e.g., by conferring the type information that it gathers to a public credit bureau. In another one, the Big Tech gives *privately* the type info to the bank, e.g., by selling it credit scoring. Data are non-rival goods and have a zero marginal cost as a by-product of other digital services (Feyen et al, 2021). Thus, we assume that the

[20] Jagtiani et al (2019) find that fintech lenders in the United States tend to supply more mortgages to consumers with weaker credit scores than do banks; they also have greater market shares in areas with lower credit scores and higher mortgage denial rates.

transaction between the Big Tech and the receiver of the data in either arrangement does not alter the Big Tech incentive to gather and transfer the data.

Consider the case where the Big Tech gives its type info to a credit bureau which makes it *public*, that is the type p is publicly observed. In this model there are two sources of asymmetric information: about the firm type and about the firm output. A public bureau that makes type information public eliminates the first source of asymmetric information, but not the lack of output information. The latter forces the bank to continue to restrain the requested repayment to avoid strategic default. However, once type information is available to any bank, under perfect competition, each bank must charge a rate that breaks even for each type, thus riskier firms must be charged more. It turns out that the break even repayment requested to riskiest firms exceeds the rate that induces them not to default strategically. Hence, riskiest firms are rationed.

More formally, the bank can now lend and contract on a repayment $R(p)$ without facing PIC. Competition between banks leads to a break-even condition for each type of firm: $R(p) = 1/p$. However, Incentive Compatibility requires that $R(p) \leq p(Y - (\rho - 1)) - 1$. Thus, there exists $p^* > 1/Y$, such that for solvent firms with $p < p^*$ the break-even $R(p)$ exceeds the maximum repayment consistent with no strategic default.[21] Since the bank has no interest to set $R(p) > R(p^*)$, it would not lend to types $p \in [1/Y, p^*)$ at $t = 0$. Hence public information leads to an outcome whereby some solvent firms are rationed. Figure 2 illustrates this point.

On the contrary, with the sale of credit scoring, the information remains *private*. Hence, the banks cannot contract on it and still face a principal's incentive constraint or PIC. As in the case of Proposition 3 with a Big Tech learning the type privately, banks must also compete by offering privacy to their clients in the form of noisy learning in our model.[22] A bank that learns the type with a lower level of noise ends up demanding a higher repayment, risking being undercut by another bank offering greater privacy. In equilibrium, banks learn type information from Big Techs with a noise denoted by δ, where a firm of type p anticipates its incentive compatible rate $R(p) = \max(Y(p - \delta) - 1, 0)$, and δ is determined by the bank zero profit condition:

$$\mathbb{E}\left(\Pi^B\right) = \int_{\frac{1}{Y}+\delta}^{1} p\left(Y(p-\delta) - 1\right) f(p)\, dp - \int_{\frac{1}{Y}}^{1} f(p)\, dp = 0.$$

[21] The cut-off level p^* is obtained from the condition $p^*(Y - (\rho - 1)) - 1 = 1/p^*$.

[22] As mentioned, the algorithm of the consumer credit scoring FICO voluntarily excludes some information.

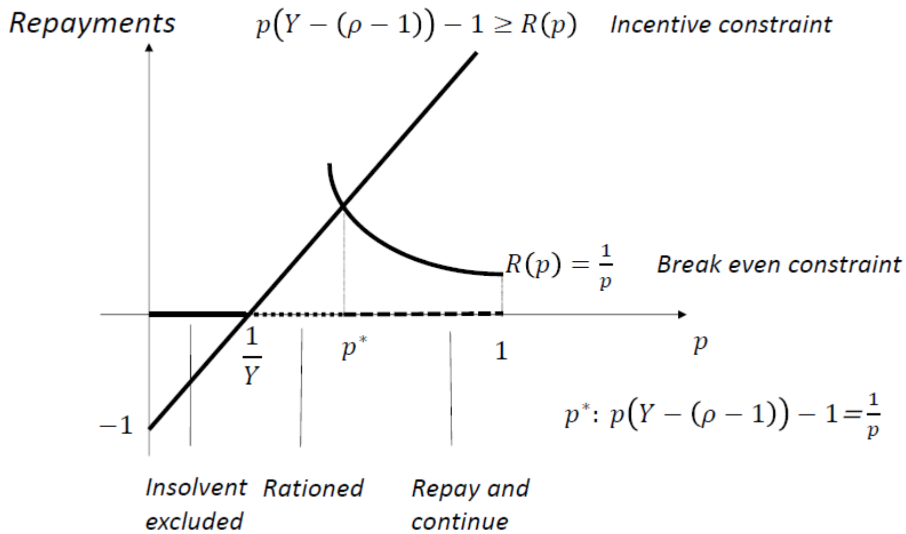


Figure 2: Big Techs share information publicly with banks, as in a credit bureau.

As competitive banks lend to all solvent firms the repayment simply splits the surplus between bank and firms and it is therefore welfare irrelevant. Figure 3 illustrates this point.

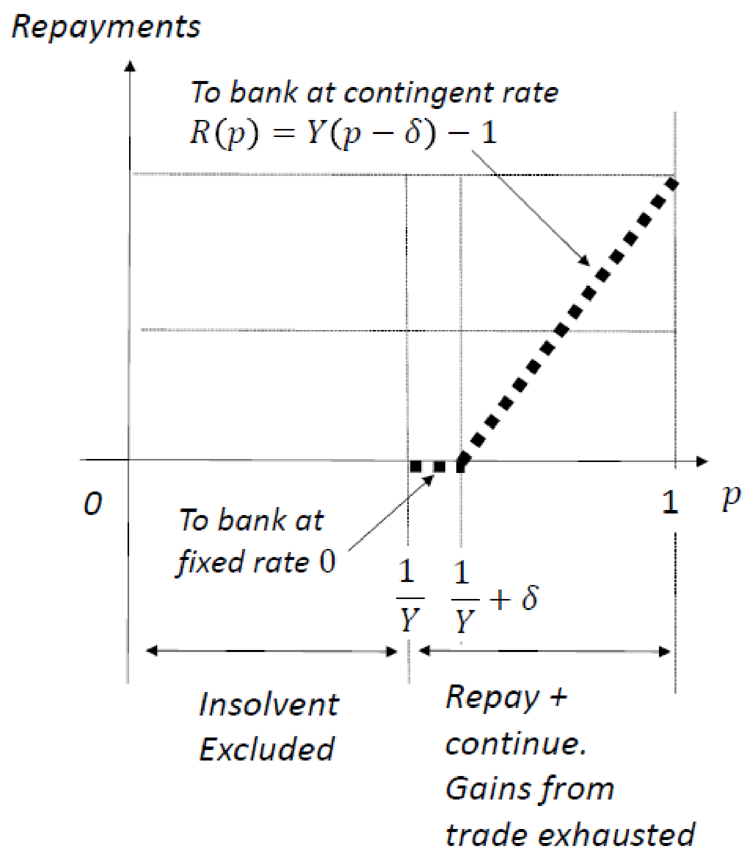


Figure 3: Big Techs share information privately with banks, e.g. selling credit score.

We collect these observations in the following proposition:

Proposition 4. It matters how Big Techs share information with banks: when banks compete for firms providing more public information to banks end up rationing solvent firms, while sharing information privately exploits all gains from trade.

The different outcomes in the two arrangements stem from the fact when information is public, competition between banks forces them to charge a rate that breaks even for each type. This conflicts with the maximum rate that a bank can charge without inducing strategic default and leads to rationing of the riskiest types. When instead, the information that the Big Tech shares remains private, as in a credit scoring service, that information cannot be used by that bank or other banks to commit to a rate. Hence, competition on type-contingent rates is mute and the zero expected profit condition holds on the entire portfolio of loans not on each loan. This cross-subsidy allows to lend to all solvent firms.

Therefore, in this model information sharing between Big Techs and banks is beneficial because it excludes insolvent firms and allows to tailor rates to types to exploit all gains from trade. However, selling credit scoring is preferable to making information public and contractible which prevents cross-subsidies. Hence it is preferable to put in place mechanisms that guarantee privacy in the collection of information to share rather than simply provide more information to the public.

7 Discussion and Conclusion

We have modeled competition between banks and Big Techs in a credit market where adverse selection and difficulty to enforce repayment cause frictions. We have obtained three main results.

First, more powerful retaliation after default increases welfare as it lowers the fraction of solvent firms that strategically default (Lemma 1). This result implies an extreme form of exclusivity where a firm cannot access e-commerce through one Big Tech after defaulting on a loan from another Big Tech. Lemma 1 has the counterintuitive implication that regulations limiting a Big Tech's ability to exclude defaulting firms from its ecosystem encourages inefficient strategic defaults by solvent firms. However, the exclusivity of the agreements between Big Techs and firms may lead to monopoly distortions in access to e-commerce. For this reason, for example, regulators in China have dismantled the policy "choose one from two" whereby an e-commerce platform can prevent a firm from selling through another platform. Our model doesn't capture the monopoly distortion of exclusive dealing as we don't price access to e-commerce. Thus, if extending credit is the main policy objective (for example for an urgent need of financial innovation) there is a case for exclusivity. If, on the contrary, the main policy objective becomes preventing monopoly distortions in access to e-commerce there is a case to eliminate exclusivity. That is as Big Techs' credit footprints grow, exclusivity to enforce repayment loses importance.

Second and quite intuitively, their superior information and their superior enforcement allows Big Techs to enter in banks' turf with lower rates (Proposition 1). However, the general theme from Propositions 2 and 3 is that the ex-ante competitive threat of banks prevent Big Techs from charging very high rates to the safest firms. To enter bank's turf Big Techs must self-regulate by credibly committing to data privacy to limit firms' exploitation (Proposition 3).

Furthermore, our model points to information sharing as a way to mitigate the tension between privacy and efficiency. We have explored two information-sharing arrangements between Big Techs and banks (Proposition 4). This aspect is particularly relevant in light of the ongoing debate on trade-offs and limits of alternative information arrangements. Our conclusion is in the tradition of the literature that shows the peril that competition destroys solutions based on cross-subsidies. In the presence of limitations in the judicial system, we show that it is preferable not to make information public, rather it is better to regulate privacy to permit the cross-subsidies that allow exploiting all gains from trade.

In order to have a workable model of the trade-off between data privacy and efficiency we remain silent on many related issues. First, we have taken as given the network externalities stemming both from the e-commerce and the payments systems run by the Big Techs. Second, we did not model data externalities. Data externalities refer both to the fact that information about an individual helps understanding characteristics of other (Hagiu and Wright, 2020; Brunnermeier et al, 2020; Bergemann et al, 2021) and to the fact that information about others help the individual like in Google traffic data. Third, we did not address the regulations of the multidimensional issues raised by the presence of the Big Techs in the financial world. In particular, on the funding side we do not model capital regulation and deposit insurance and the related financial stability concerns, issues that are growing in importance with the financial footprint of

Big Techs. Fourth, we model Big Techs and banks as direct competitors. However, the analysis of information sharing captures the business model of a Big Tech that is mainly an information intermediary collecting service fees. As such, even if we don't consider fees as they are welfare irrelevant in this setting, our model nonetheless captures some of the elements of the Ant business model where banks provide the bulk of the funding and take almost all credit risk.

Finally, this work could be extended to Big Techs with market power. For example, in Proposition 3 the problem that a monopoly Big Tech faces when setting the repayment rate is like that under perfect competition. In fact, regardless of market power the repayment that an informed Big Tech can demand is capped by the possibility that the firm defaults strategically minus a rent to induce the firm not to borrow from the bank. Therefore, when the Big Tech maximizes its expected profit subject to the zero expected profit condition of the bank the only changes will be in the size of the regions of the firm types that borrow from Big Tech and bank.

8 References

Acquisti, A., and H.R. Varian (2005). "Conditioning Prices on Purchase History," *Marketing Science*, 24, 367-381.

Acquisti, A., C. Taylor, and L. Wagman (2016). "The economics of privacy," *Journal of Economic Literature*, 54, 442-492.

Belleflamme, P., N. Omrani, and M. Peitz (2016). "The Economics of Crowdfunding Platforms," CESIfo DICE Report, Volume 14.

Bergemann, D. Bonatti, A., and T. Gan (2021). "The Economics of Social Data," mimeo, Yale University, January.

Berg, T., V. Burg, A. Gombović, and M. Puri (2020). "On the rise of fintechs – credit scoring using digital footprints," *The Review of Financial Studies*, 33, 2845-2897.

Bank for International Settlements (2019). "Big tech in finance: opportunities and risks," *Annual Economic Report*, Chapter III, June.

Bolton, P., and D. Scharfstein (1990). "A theory of predation based on agency problems in financial contracting," *The American Economic Review*, 80, 93-106.

Bolton, P., X. Freixas, L. Gambacorta, and P. E. Mistrulli (2016). "Relationship and Transaction Lending in a Crisis," *The Review of Financial Studies*, 29, 2643–2676.

Calzolari, G. and A. Pavan (2006). "On the Optimality of Privacy in Sequential Contracting," *Journal of Economic Theory*, 130, 168-204.

Carrillo, J.D. and T. Mariotti (2000). "Strategic Ignorance as a Self-Disciplining Device," *Review of Economic Studies*, 67, 529-544.

Carstens, A. (2018). "Big tech in finance and new challenges for public policy," Keynote address, FT Banking Summit, London, 4 December 2018.

Carstens, A., S. Claessens, F. Restoy, and H.S. Shin (2021). "Regulating big techs in finance," BIS Bulletin, No 45.

Cole, H. and P. Kehoe (1995). "The role of institutions in reputation models of sovereign debt," *Journal of Monetary Economics*, 35, 45-64.

Cornelli, G., J. Frost, L. Gambacorta, R. Rau, R. Wardrop, and T. Ziegler (2020). "Fintech and big tech credit: a new database," BIS, Working paper No 887.

Cornelli, G., J. Frost, L. Gambacorta, C. Mu, and T. Ziegler (2021). "Big tech credit during the Covid-19 pandemic," mimeo.

Dorfleitner, G., C. Priberny, S. Schuster, J. Stoiber, M. Weber, I. de Castro, and J. Kammler (2016). "Description-text related soft information in peer-to-peer lending – Evidence from two leading European platforms," *Journal of Banking and Finance*, 64, 169-87.

Feyen, E., J. Frost, L. Gambacorta, H. Natarajan, and M. Saal (2021). "Fintech and the digital transformation of financial services: implications for market structure and public policy," BIS, Working paper, No 117.

Fong, Y., K. Liu, X. Meng, and K.Y. Tam (2021). "FinTech Lending and Financial Inclusion," mimeo, March.

Frost, J., L. Gambacorta, Y.L. Huang, H.S. Shin, and P. Zbinden (2019). "Big Tech and the changing structure of financial intermediation," *Economic Policy*, 34, 761-99.

Frost, J. (2020). "The economic forces driving fintech adoption across countries," BIS, Working paper, No 838.

Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther (2019). "The effect of machine learning on credit markets," VoxEU, 11 January 2019.

Gambacorta, L., Y. Huang, Z., Li, H. Qiu, and S. Chen (2020). "Data vs collateral," BIS, Working paper, No 881.

Hagiu, A. and J. Wright (2020). "Data-enabled learning, network effects and competitive advantage," mimeo.

Hau, H., Y. Huang, H. Shan, and Z. Sheng (2019). "How FinTech Enters China's Credit Market," *AEA: Papers and Proceedings*, 109, 60-64.

He, Z., J. Huang, and J. Zhou (2020). "Open Banking: Credit Market Competition When Borrowers Own the Data," Working paper, No 168, Becker Friedman Institute, University of Chicago.

Huang Y., C. Lin, Z. Sheng, and L. Wei (2018). "FinTech Credit and Service Quality," mimeo.

Jagtiani, J., and C. Lemieux (2018). "The roles of alternative data and machine learning in fintech lending: evidence from the LendingClub consumer platform," Federal Reserve Bank of Philadelphia, Working Papers, 18–15, April.

Jagtiani J., L. Lambie-Hanson, and T. Lambie-Hanson (2019). "Fintech lending and mortgage credit access," Federal Reserve Bank of Philadelphia Working Papers, 19-47, November.

Jappelli, T., and M. Pagano (2002). "Information sharing, lending and defaults: Cross-country evidence," *Journal of Banking and Finance*, 26, 2017-2045.

Khalil, F., J. Lawarrée, and T.J. Scott (2015). "Private monitoring, collusion, and the timing of information," *RAND Journal of Economics*, 46, 872-890.

Laffont, J-J., and D. Martimort (2002). *The Theory of Incentives*, Princeton University Press.

Myers, S., and R.G. Rajan (1998). "The Paradox of Liquidity," *Quarterly Journal of Economics*, 113, 733-771.

Pagano, M., and T. Jappelli (1993). "Information sharing in credit markets," *The Journal of Finance*, 48, 1693-1718.

Parlour, C.A., H. Rajan, and H. Zhux (2020). "When FinTech Competes for Payment Flows," mimeo February.

Philippon, T. (2020). "On fintech and financial inclusion," BIS, Working paper No 841, Monetary and Economic Department.

Sharpe, S.A. (1990). "Asymmetric Information, Bank Lending and Implicit Contracts: A Stylized Model of Customer Relationships," *The Journal of Finance*, 45, 1069–1087.

Taylor, C. R. (2004). "Consumer privacy and the market for customer information," *RAND Journal of Economics*, 35, 631-650.

Townsend, R. M. (1979). "Optimal contracts and competitive markets with costly state verification," *Journal of Economic Theory*, 21, 265–293.

Vives, X. and Z. Ye (2021). "Information Technology and Bank Competition," mimeo.

Von Thadden, E. L. (1995). "Long-term contracts, short-term investment and monitoring," *Review of Economic Studies*, 62, 557–575.

Yu, S. and R. Mc Morrow (2021). "Beijing to break up Ant's Alipay and force creation of separate loans app," Financial Times, 13 September.

Zetzsche, D.A., R.P. Buckley, D.W. Arner, and J.N. Barberis (2017). "From FinTech to TechFin: The Regulatory Challenges of Data-Driven Finance," EBI Working Paper Series – no. 6.

9 Appendix

Proof of Lemma 1.

The zero expected profit condition for a representative lender determines the break-even repayment rate R:

$$\mathbb{E}(\Pi) = \int_{\frac{R+1}{Y}+\rho-1}^{1} pRf(p)\,dp - (1+r)\int_{0}^{1} f(p)\,dp = 0. \tag{6}$$

Differentiating (6) with respect to the parameter ρ we have

$$\frac{dR(\rho)}{d\rho}\left[\left\{\int\limits_{\frac{R(\rho)+1}{Y}+\rho-1}^{1} pf(p)\,dp\right\} - \frac{R(\rho)}{Y}\left\{\left(\frac{R(\rho)+1}{Y}+\rho-1\right)f\left(\frac{R(\rho)+1}{Y}+\rho-1\right)\right\}\right]$$
$$= \underbrace{R(\rho)\left[\frac{R(\rho)+1}{Y}+\rho-1\right]f\left(\frac{R(\rho)+1}{Y}+\rho-1\right)}_{>0}.$$

Observe that the coefficient of $\frac{dR(\rho)}{d\rho}$ on the LHS is positive as $\frac{R(\rho)}{Y} < 1$, and term $\left\{\left(\frac{R(\rho)+1}{Y}+\rho-1\right)f\left(\frac{R(\rho)+1}{Y}+\rho-1\right)\right\}$ is the value of the function $pf(p) > 0$ in the integral $\int_{\frac{R(\rho)+1}{Y}+\rho-1}^{1} pf(p)\,dp$ evaluated at its lower limit. Hence we have $\frac{dR(\rho)}{d\rho} > 0$, from which we also have that $\widehat{p} = \frac{R(\rho)+1}{Y} + \rho - 1$ increases with ρ. End proof.

Proof of Proposition 1. The repayments R^P and R^B are to be determined by corresponding ZEP conditions for the Big Tech and the bank. These are provided at the end of the proof. Types $p < 1/Y$ borrow from the bank and default; they are excluded by the Big Tech based on solvency. Recall that the cutoff levels $\widehat{p}_i$, for $i = B, P$, from the firm IC given success:

$$\widehat{p}_B = \frac{R^B+1}{Y} + \rho - 1,$$
$$\widehat{p}_P = \frac{R^P+1}{Y}.$$

Types $p < \min\{\widehat{p}_B, \widehat{p}_P\}$ will strategically default from both lenders. Thus, they strictly prefer to borrow from the bank as they expect to earn $p\rho Y > pY$ if they strategically default with a bank rather than a Big Tech loan. Also, note that types $p > \max\{\widehat{p}_B, \widehat{p}_P\}$ will repay both lenders and self-finance a reinvestment. Thus, they decide on whom to borrow simply comparing the repayments R^B versus R^P.

First, we prove that the Big Tech cannot enter if $R^P > R^B$. We already know that types $p < 1/Y$ and $p > \max\{\widehat{p}_B, \widehat{p}_P\}$ borrow from the bank. We will show next that the remaining types between the two cutoffs also borrow from the bank. If $\widehat{p}_B > \widehat{p}_P$, then all $p \in (\widehat{p}_P, \widehat{p}_B)$ prefer to borrow from the bank and strategically default as:

$$p\rho Y > p(Y - R^B - 1 + pY) > p(Y - R^P - 1 + pY).$$

If $\widehat{p}_B < \widehat{p}_P$, then all $p \in (\widehat{p}_B, \widehat{p}_P)$ prefer to borrow from the bank and repay as:

$$p(Y - R^B - 1 + pY) > p\rho Y > pY > p(Y - R^P - 1 + pY).$$

Thus, it must be that $R^P < R^B$ when the Big Tech and bank coexist. Then, it must be that $\widehat{p}_B > \widehat{p}_P$. We also know that types $p > \widehat{p}_B$ strictly prefer to borrow from the Big Tech and repay. To show how the types $p \in (\widehat{p}_P, \widehat{p}_B)$ self-select lenders, we define a cut-off p_t by

$$p_t \equiv \frac{R^P + 1}{Y} + \rho - 1,$$

where the type p_t is indifferent between strategically defaulting on a bank loan or borrowing from the Big Tech and repaying it.

$$p_t \rho Y = p_t(Y - R^P - 1 + p_t Y).$$

Then, types $p < p_t$ strictly prefer to borrow from the bank and strategically default as:

$$p\rho Y > p(Y - R^P - 1 + pY) > p(Y - R^B - 1 + pY),$$

whereas, types $p > p_t$ prefer to borrow from the Big Tech and repay.

Thus, we have proved that when the Big Tech and the bank compete with $R^P < R^B$, for all types $p \in [0, 1]$, those with $p < p_t$ strictly prefer to borrow from the bank and default, while types $p > p_t$ strictly prefer to borrow from the Big Tech and repay. However, the latter types are limited by the Big Tech's capacity α. Therefore, R^P solves the BT's ZEP:

$$\mathbb{E}\left(\Pi^P\right) = \frac{\alpha}{1 - F\left(1/Y\right)} \left[\int_{p_t}^{1} \left(pR^P - (1 + r)\right) f\left(p\right) dp\right] = 0, \tag{7}$$

which can be written also as

$$\int_{p_t}^{1} \left(pR^P - (1+r)\right) f(p)\, dp = 0. \tag{8}$$

There exists $R^P \in (1+r, \frac{1+r}{p_t})$ that solves the Big Tech's ZEP condition.[23]

The bank's repayment R^B solves the bank's ZEP :

$$\mathbb{E}\left(\Pi^B\right) = (1-\alpha)\left[\int_{\frac{R^B+1}{Y}+\rho-1}^{1} pR^B f(p)\, dp - \int_{p_t}^{1} f(p)\, dp\right] - \int_{0}^{p_t} f(p)\, dp = 0, \tag{9}$$

which can be written also as

$$\int_{\frac{R^B+1}{Y}+\rho-1}^{1} pR^B f(p)\, dp - \int_{p_t}^{1} f(p)\, dp - \frac{1}{(1-\alpha)}\int_{0}^{p_t} f(p)\, dp = 0. \tag{10}$$

To see that there exists r such that $R^P < R^P$ can occur, assume for the time being that $r = 0$ also for Big Tech. Observe that $R^P : \mathbb{E}\left(\Pi^P\right) = 0$ with $r = 0$ also solves

$$\int_{p_t}^{1} \left(pR^P - 1\right) f(p)\, dp = 0. \tag{11}$$

When $r = 0$, observing that $p_t \leq \frac{R^B+1}{Y} + \rho - 1 = \widehat{p}_B$, we can see that $R^B > R^P$ by comparing the repayment rates R^P and R^B that solve (9) and (11). By continuity, $R^B > R^P$ continues to hold if we increase r up to the point where either the expected profit for the Big Tech is less than 0 or $R^B \leq R^P$. End proof.

Proof of Proposition 2. Consider each type of firm's choice between the bank and the Big Tech. Again, firms with $p \in [0, \frac{1}{Y})$ will be detected and excluded by the Big Tech while they can borrow from the bank. If successful they will strategically default at $t = 1$, and they can expect to obtain ρpY at $t = 0$ from a bank loan, with $\rho > 1$. If firms with $p \in \left[\frac{1}{Y}, 1\right]$ borrow from the Big Tech they are fully exploited as they will have to repay $R(p) = pY - 1$ with an expected return pY at $t = 0$.[24] What they obtain if they borrow from the bank depends on their type: firms with $p \in \left[\frac{1}{Y}, \frac{R+1}{Y} + \rho - 1\right]$ will strategically default on a bank loan to obtain $p\rho Y$; firms with $p \in \left[\frac{R+1}{Y} + \rho - 1, 1\right]$ enjoy a rent from the bank (zero rent only for types $p = \frac{R+1}{Y} + \rho - 1$) at the reinvestment

[23] If $R^P = 1 + r$, then $\left(pR^P - (1+r)\right) \leq 0$ for all p, and if $R^P = \frac{1+r}{p_t}$, then $\left(pR^P - (1+r)\right) > 0$ for all $p > p_t$.

[24] It follows from $pY = p\left[Y - (pY - 1) + (pY - 1)\right]$.

stage. Hence, firms with $p \in \left[\frac{1}{Y}, 1\right]$ have no incentive to borrow from the Big Tech from which they are fully exploited at the reinvestment stage. End proof.

Proof of Proposition 3.

We analyze the behavior of the three groups of firms separately and then we derive the zero expected profit conditions of the bank and the Big Tech.

Group 1. Recall that $p_0 = \frac{1}{Y} + \rho - 1$. Firms $p \in [0, \frac{1}{Y})$ are cutoff from the Big Tech. From (5) firms $p \in [\frac{1}{Y}, p_0)$ prefer to borrow from the bank and default strategically.

Group 2. Let us start with the riskiest firms in Group 2. For $p \in [p_0, \frac{1}{Y} + \varepsilon)$, the Big Tech's expected repayment is 0.[25] Consider now the safest firms in Group 2. Define a marginal type $p_b \geq \widehat{p}_B = \frac{R^B+1}{Y} + \rho - 1$ by

$$p_b = \frac{R^B + 1}{Y} + \varepsilon, \tag{12}$$

where the type p_b is indifferent between between borrowing from either lender with no strategic default:

$$\underbrace{Y - R^B - 1 + p_b Y}_{\text{from bank}} = \underbrace{Y - (p_b - \varepsilon)Y + p_b Y}_{\text{from BT}}. \tag{13}$$

Recalling that firms with $p \in [\widehat{p}_B, 1]$ will repay both lenders after success, types $p \in [\frac{1}{Y} + \varepsilon, p_b)$ prefer to borrow from the Big Tech. A fraction α of them borrows from the Big Tech, and the complement from the bank. Thus for each $p \in [\frac{1}{Y} + \varepsilon, p_b)$, the expected repayment to the Big Tech is $\mathbb{E}\left[R(s)|p\right] = Y\left(p - \varepsilon\right) - 1$.

Group 3. From (13) firms with $p \in (p_b, 1]$ strictly prefer to borrow from the bank at R^B.

Finally, the bank repayment R^B, and the Big Tech noisy learning ε, are jointly

[25] From $R(s) = \max\left(0, (s - \varepsilon)Y - 1\right)$,we have, for $p \in [\frac{1}{Y} + \rho - 1, \frac{1}{Y} + \varepsilon)$

$$\mathbb{E}\left[R(s)\right] = \max\left(0, (p - \varepsilon)Y - 1\right) = 0.$$

determined by the two $ZEPs$, where $p_b = \frac{R^B+1}{Y} + \varepsilon$,

$$\mathbb{E}\left(\Pi^B\right) = \underbrace{\int_{p_b}^{1} \left[pR^B - 1\right] f(p)\, dp}_{\text{Group 3}} + \underbrace{(1-\alpha)\left[\int_{\widehat{p}_B}^{p_b} pR^B f(p)\, dp - \int_{p_0}^{p_b} f(p)\, dp\right]}_{\text{Group 2}} - \underbrace{\int_{0}^{p_0} f(p)\, dp}_{\text{Group 1}} = 0, \tag{14}$$

$$\mathbb{E}\left(\Pi^P\right) = \underbrace{\frac{\alpha}{1-F(1/Y)}\left[\int_{\frac{1}{Y}+\varepsilon}^{p_b} p\left(Y(p-\varepsilon)-1\right) f(p)\, dp - \int_{p_0}^{p_b} (1+r) f(p)\, dp\right]}_{\text{Group 2}} = 0, \tag{15}$$

where the first zero expected profit condition is for the bank and the second is for the Big Tech.

Finally, $\varepsilon > \rho - 1$. This is because if $\varepsilon = \rho - 1$, in which case $p_b = \widehat{p}_B = \frac{R^B+1}{Y} + \rho - 1$, the solution for R^B generically would not satisfy (15) and (14). End of proof.