

The Macroeconomics of BigTech*

Dan Su[†]

CKGSB

This Version: July 17, 2023

Abstract

This paper investigates the macroeconomics of BigTech. Compared to banks whose key characteristic is the standard collateral-based borrowing constraint, the essential feature of BigTech is the *expected-earnings-based* lending. Our model implications are twofold. First, BigTech has an efficiency-instability tradeoff as it leads to less misallocation but a higher default rate in the steady state. Second, BigTech can be interpreted as a different financial accelerator from the classical one as it generates persistent impacts on aggregate outputs through amplifying and propagating the second-moment micro-uncertainty shocks. Finally, we discuss the role of algorithm bias and optimal BigTech development.

JEL codes: E32, E44, E71, G23

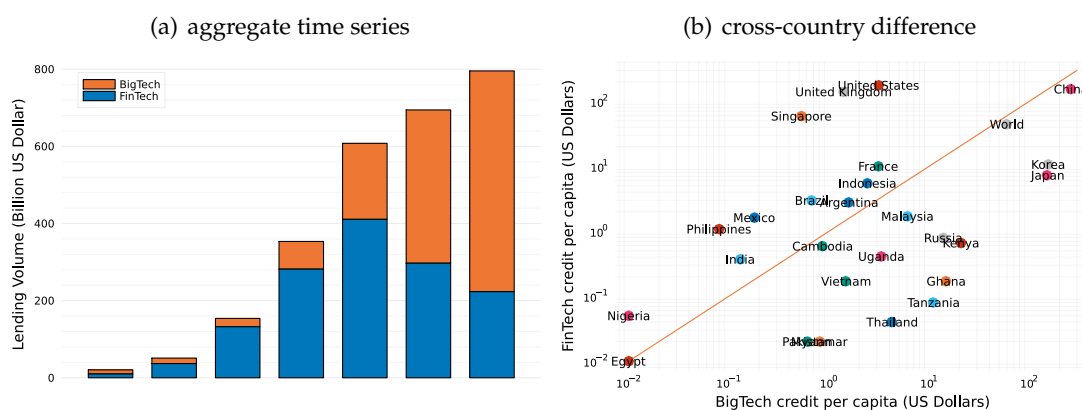
Keywords: BigTech; financial intermediation; earnings-based borrowing constraint; financial accelerator mechanism; misallocation; financial instability

*First Draft: November 2021. The previous title was “The Macroeconomics of TechFin”. For helpful comments and suggestions, I thank Tobias Berg, Diana Bonfim (discussant), Robin Döttling (discussant), Leonardo Gambacorta, Yi Huang (discussant), Minghao Li, Xiang Li, Chen Lian, Xiaoji Lin, Kjetil Storesletten, Martin Strieborny (discussant), Andrew Winton, Hong Zhang, Xingtian Zhang, Hao Zhou, Shenghao Zhu, and all the other participants at EEA-ESEM, EFA, CEBRA, China Five-star Workshop, CCER Summer Institute, and AMES China. Of course, all errors are mine own.

[†]Affiliation: Department of Finance, Cheung Kong Graduate School of Business. Oriental Plaza, 1st East Chang’An Street, Beijing, 100738, China. Email address: dansu@ckgsb.edu.cn

Over the past decade, the financial market has seen the arrival of massive new technologies, raising many debates about their consequences. The two most important ones are *FinTech* and *BigTech* (or TechFin).¹ Using the cross-country dataset provided by Cornelli et al. (2023), in Graph (a) of Figure 1, we present the world’s total lending volume in billion U.S. dollars for both Fintech and BigTech credits. As we can see, both of them have become increasingly important in our modern financial system. In addition, Graph (b) shows their relative importance across different countries in 2019. Some countries, such as the United States, the United Kingdom, and Singapore, have more development in FinTech; meanwhile, some other countries, including China, Korea, and Japan, have relatively better BigTech access. Generally speaking, these two new types of financial intermediaries have emerged at a fast pace across different credit markets around the world, which, not surprisingly, leads to a fast-growing **empirical** literature on them (e.g., Hau et al., 2018; Tang, 2019).

Figure 1: The Rise of FinTech and BigTech



In this paper, we attempt to investigate, **in theory**, the role of BigTech lending in the

¹Throughout this paper, FinTech refers to the situation where financial firms adopt new types of technology, while BigTech means that technology companies provide financial services. Typical examples of FinTech are these digital platforms facilitating peer-to-peer (P2P) lending and borrowing, while examples of BigTech include Ant Group, WeBank, and so on. The broad definition of FinTech provided by the Financial Stability Board is “technologically enabled financial innovation that could result in new business models, applications, processes, or products with an associated material effect on financial markets and institutions, and the provision of financial services.”

macroeconomy. More specifically, we explain how should we modify the existing theories of financial intermediation and business cycles so as to accommodate the rise of BigTech. In the existing macro-finance literature, theories of credit are crucial for our understanding of the macroeconomy (e.g., Buera et al., 2011; He and Krishnamurthy, 2013). However, these theories are centered on banks and the key characteristic of bank lending is this collateral-based borrowing constraint. With this financial friction, many studies find that the aggregate economy has productivity losses in the steady state because the efficient producers cannot borrow enough (e.g., Moll, 2014). In addition, a financial accelerator mechanism lies behind the macroeconomic fluctuations: small fundamental shocks can be amplified by financial frictions so that they can generate large and persistent fluctuations in aggregate economic activity (e.g., Kiyotaki and Moore, 1997; Bernanke and Gertler, 1989; Bernanke et al., 1999). Therefore, when introducing BigTech, we focus on the following two research questions: first, what is the key difference between banks and BigTech in terms of their lending behaviors; second, based on the answer to the first question, how different are those previous macro-finance implications with a new type of financial intermediation.

To begin with, we argue that the fundamental difference between bank credit and BigTech credit lies in the specific type of borrowing constraints. When large technology firms lend to the market, they have less demand for collateral and corporate borrowing is subject to an earnings-based borrowing constraint. The reason why we use this assumption is related to how BigTech companies reduce asymmetric information or agency problems in practice. In the traditional banking sector, banks use covenants or collateral to mitigate agency frictions. However, for BigTech, their technology advantages such as data, algorithms, and platforms can help reduce these agency frictions. In Section 1, we present a simple model to show if the costs of state verification can be significantly reduced with technology or data advantages, then these firms strongly prefer incomplete-collateralized contracts to fully-collateralized ones. Besides, there exists some empirical support for our assumption of earnings-based borrowing constraint on BigTech credit.

For instance, Gambacorta et al. (2023) find that BigTech credit does not correlate much with shocks to local business conditions and house prices when controlling for demand factors, but it does react strongly to changes in firm-specific characteristics, such as the transaction volumes and network scores used to calculate firm credit ratings. They argue that the critical feature of BigTech is allowing firms to borrow without any collateral.

After that, based on this assumption, we introduce both a banking sector and a BigTech sector into a continuous-time general equilibrium model with heterogeneous entrepreneurs and defaultable debt. These two financial sectors are identical except for the types of borrowing constraints faced by entrepreneurs. Entrepreneurs borrowing from banks are subject to the standard collateral-based borrowing constraints. In contrast, technology advantages allow BigTech companies to resolve agency costs and perform *expected-earnings*-based lending. With this macroeconomic framework, we show that compared to the traditional banking sector, this BigTech sector has two new macro-finance implications. First, there is an efficiency-instability tradeoff associated with BigTech development. On one hand, as it allows more productive firms to use more leverage and grow faster, BigTech credit is more efficient in resource allocation compared to traditional bank loans. On the other hand, due to the difference between *expected* and *realized* earnings, BigTech credit leads to overlending issues and hence a higher default probability in the steady state. This efficiency-instability tradeoff implies that BigTech cannot fully replace the role of traditional banks. Second, as for its impacts on the business cycles, BigTech can be interpreted as a different financial accelerator. We show that a transitory micro-uncertainty shock can lead to amplified and persistent effects on allocation efficiency and aggregate productivity. This new financial accelerator mechanism, associated with a new type of financial intermediation, differs from the classic one (e.g., Kiyotaki and Moore, 1997; Bernanke and Gertler, 1989) in three aspects: the second-moment micro-uncertainty instead of the first-moment aggregate productivity is the primitive shock; financial friction comes from earnings-based borrowing constraints instead of collateral-based ones; and the feedback loops happen between net worth in-

equality, instead of net worth level, and asset prices.

Finally, we extend our baseline model setup to discuss the role of algorithm bias and optimal BigTech development. In the first extension, we assume that BigTech's predicted future earnings have an extrapolative algorithm bias component: they tend to overestimate the likelihood of a positive future state when the current news is favorable, and vice versa when negative. With the presence of algorithm bias, the financial instability concern of BigTech becomes severer. In our second model extension, we make the optimal size of BigTech endogenous. We show that if the government cares about both resource allocation efficiency and financial stability, our theory indicates that there exists an optimal degree of BigTech development in the whole economy.

Related literature Our paper is related to four different branches of literature. First, this paper builds on the extensive literature on financial frictions and business cycles. Two seminal works in this field are Kiyotaki and Moore (1997) and Bernanke and Gertler (1989). Examples of further quantitative explorations include Carlstrom and Fuerst (1997), Bernanke et al. (1999), and many others. Recent studies, especially those done after the 2008-2009 global financial crisis, are mainly focused on analyzing the global dynamics and nonlinear effects of shocks with continuous-time models. Examples include but are not limited to Brunnermeier and Sannikov (2014), Di Tella (2017), He and Krishnamurthy (2013), and Fernandez-Villaverde et al. (2019).² However, most of these studies are focused on banks as the financial intermediation and hence the collateral-based borrowing constraint as the financial friction. In contrast, our work introduces BigTech, which becomes increasingly important in the new economy, and further explores its different macro-finance implications.

Second, our work closely relates to the growing literature on investigating the difference between FinTech and the traditional banking sector. As mentioned before, most

²Brunnermeier et al. (2013) provide an excellent and detailed survey on the discrete-time models of macroeconomics and financial frictions, and Brunnermeier and Sannikov (2017) introduce the fundamental tools used in this field.

of these studies are empirical. For instance, by using the US Peer-to-Peer (P2P) lending data, Tang (2019) finds that FinTech lending works as a complement to bank lending for small-scale loans. Similarly, Cornelli et al. (2023) find that BigTech lending complements rather than substitutes other forms of lending with a cross-country panel dataset for 79 countries during 2013-2019. In addition, Hau et al. (2018) show that the existence of FinTech credit in China improves the credit access condition for firms with lower credit scores. Liu et al. (2022) empirically show that the essential feature of BigTech lending comes from serving borrowers' short-term liquidity rather than their long-term financing needs. However, FinTech loans are not always found to be complements to bank loans. For example, Cornaggia et al. (2019) document that different from low-risk loans, high-risk FinTech lending crowds out bank lending. In terms of theoretical studies, a closely related paper to ours is Manea et al. (2023), which investigates the different monetary policy implications between banks and BigTech lending. In their paper, the key difference between these two financial intermediations is the borrowers' opportunity cost of default. If firms default on bank loans, they lose their collateral. However, if firms default on BigTech loans, they lose future profits as they will be excluded from BigTech's e-commerce platform.

Third, this paper connects to the literature on the macroeconomics of earnings-based borrowing constraints. There have been numerous theoretical papers written about the factors that determine corporate borrowing. Some of these papers, such as Stiglitz and Weiss (1981) and Holmstrom and Tirole (1997), suggest that corporate earnings should determine the debt capacity of corporations. However, others, including Hart and Moore (1994), Kiyotaki and Moore (1997), and Bernanke and Gertler (1989), argue that corporate borrowing should be closely tied to the liquidation value of assets. In terms of empirical evidence, Lian and Ma (2021) find that 80% of the corporate debt value in the US is closely linked to the firm's cash flows from their operations instead of the asset liquidation value. Their work intrigues an increasing number of studies that investigate the role of earnings-based borrowing constraints in aggregate fluctuations. For instance,

Drechsel (2023) studies macroeconomic fluctuations through the interaction between earnings-based borrowing constraints and investment shocks. In contrast, Greenwald (2019) mainly focuses on investigating how the transmission of monetary policy shocks differs across firms with different types of covenants. Besides, Li (2022) uses the Japanese firm-level dataset to quantitatively investigate the misallocation implication when corporate earnings can be pledgeable.

Finally, some of the key elements used in our model are borrowed from the distributional macroeconomics and diagnostic expectation literature. The former refers to macroeconomic theories where the relevant state variable is a distribution and the Kolmogorov Forward equation instead of the Euler equation lies at the heart of the analysis. For instance, Moll (2014) studies the impacts of wealth-based borrowing constraints on misallocation and aggregate productivity. Kaplan et al. (2018) investigate monetary policy transmission mechanism in a Heterogeneous Agent New Keynesian (HANK) framework. In addition, Fernandez-Villaverde et al. (2019) extend the Krusell and Smith (1998) method and explore the relationship between financial frictions and wealth distributions with aggregate shocks.³ In terms of the diagnostic expectation literature, some recent studies have attempted to investigate the new macroeconomics and financial implications of this behavioral bias under different settings (e.g., Bordalo et al., 2018, 2020, 2021, 2022; L’Huillier et al., 2023). We find that this irrational component also plays an important role in the economy with expected-earnings-based constraints and defaultable bonds.

Layout The rest of the paper is organized as follows. Section 1 provides the microfoundation for our key assumption on the fundamental difference between BigTech lending and bank loans. Based on this assumption, Section 2 introduces a macroeconomic model to discuss the new macro-finance implications of BigTech compared to the traditional banking sector. After that, in Section 3, we extend our baseline model to investigate

³For a concrete introduction to the tools used in this literature, please refer to Achdou et al. (2022) for details.

the role of algorithm bias and the optimal Bigtech development level. Finally, Section 4 concludes.

1. BigTech versus Bank: Microfoundation

In this section, we present a simple model to demonstrate the coexistence of two different types of borrowing constraints, which serves as the microfoundation for the most crucial assumption used in our subsequent macroeconomic framework. In our model, this coexistence could stem from differences in either information technology or data advantages among lenders or the utilization of intangible capital among borrowers.

On a side note, although Lian and Ma (2021) document the prevalence of earnings-based borrowing constraints, in Section B in the appendix, we show that the distribution of borrowing constraints is in fact **bimodal** in the data. It means that there is no “representative” borrowing constraint in the data. Instead, both types coexist in the real economy.

1.1 Basic setup

Consider an economy where an entrepreneur with a capital stock of k needs to borrow money b from lenders for a certain investment project. All agents are assumed to be risk-neutral, and the entrepreneur has a linear consumption preference. Besides, the lender’s opportunity cost is fixed at r , and the unit liquidation value of capital is set to be l .

There are two potential outcomes for the investment project. The entrepreneur could earn $z_G k$ with a probability of p , or $z_B k$ with a probability of $1 - p$. Without loss of generality, we assume that $z_G > z_B > l > 0$.

Lenders have the option to choose between two types of lending. The first type is known as a full-collateralization contract, as defined by Bernanke and Gertler (1989). This type of contract requires the entrepreneur’s net worth to be sufficiently large such that they are able to repay the lender even in the worst state. More specifically, it means

that in the worst-case scenario, the lender seizes the entrepreneur's capital and resells it in the market. Mathematically, it can be shown as follows:

$$(1 + r) b \leq lk \quad (1)$$

One advantage of this full-collateralization lending is that lenders do not need to verify the entrepreneur's actual earnings. Instead, they use a contract directly linked to the liquidation value of the productive capital, which suffers less from issues such as asymmetric information. As a result, investors impose a standard collateral-based constraint on borrowers.

The second type of lending involves securing ownership rather than collateral, which is known as the incomplete-collateralization contract. The features of this type of contract can be expressed mathematically as follows:

$$\max_{\{q, c_G, c_B, \tilde{c}_B\}} p c_G + (1 - p) [q c_B + (1 - q) \tilde{c}_B] \quad (2)$$

subject to the following constraints

$$(1 + r) b \leq p (z_G k - c_G) + (1 - p) [z_B k - q (c_B + f) - (1 - q) \tilde{c}_B] \quad (3)$$

$$c_G \geq (1 - q) [(z_G - z_B) k + c_B] \quad (4)$$

$$c_G, c_B, \tilde{c}_B \geq 0 \quad (5)$$

$$0 \leq q \leq 1 \quad (6)$$

In the equations provided above, c_G represents the entrepreneur's consumption when they announce the good state. \tilde{c}_B denotes the entrepreneur's consumption when he or she announces the bad state, and the lenders choose not to verify. Meanwhile, c_B represents the entrepreneur's consumption when announcing the bad state, and the lenders also choose to verify. The verification action is costly and it is assumed to be f . Equation (3) represents the participation constraint, and Equation (4) is the incentive constraint.

Besides, the last two equations above are feasibility constraints. Generally speaking, the optimal incomplete-collateralization contract, denoted by $\{p, c_G, c_B, \tilde{c}_B\}$, maximizes the entrepreneur's expected consumption in Equation (2) subject to the constraints (3) to (6).

1.2 Information Asymmetry Story

The first result we would like to present is that with a relatively low verification cost with their technology or data advantages, BigTech lenders prefer cash-flow-based over asset-based lending. This theoretical implication is based on the following assumption.

Assumption 1 *Technology or data advantages allow some lenders such as BigTech firms to reduce the cost of state verification f .*

Assumption 1 argues that certain lenders may have advantages in monitoring and predicting the future earnings of firms, while others do not. A great example is BigTech firms and traditional banks. This argument is also consistent with some recent studies. For instance, Thakor (2020) highlights that the use of blockchain technology and other technological advancements in the BigTech sector can significantly reduce the cost of verification.⁴ In practice, the Ant Group uses the Alipay system to aid their lending as it enables them to easily verify and monitor the cash flows of companies at a very low cost. In other words, the Alipay system allows the Ant Group to have both the data advantage of traditional banks and the technology advantage of Fintech companies.

With Assumption 1, we can show that lenders with advantages in monitoring and predicting companies' future earnings, strictly prefer imposing earnings-based over collateral-based constraints. We summarize the main result in the following lemma.

Lemma 1 *Lenders prefer cash-flow-based over asset-based lending when the cost of state verification f is lower than a certain threshold f^* .*

⁴Other potential changes mentioned in Thakor (2020) include reduced search costs for matching transacting parties, increased economies of scale in gathering and using large data, and cheaper and more secure information transmission.

A detailed proof can be found in Appendix A. The intuition can be briefly explained as follows: from a profit-maximization standpoint, lenders always prefer cash-flow-based lending. However, due to costs associated with verifying the borrower's financial state, collateral can be useful in mitigating agency costs. In other words, if lenders are able to find ways to decrease the cost of state verification, they will switch from asset-based to cash-flow-based lending in order to maximize their expected profits.

In addition, Lemma 1 aligns with some empirical evidence in the literature. For example, Gambacorta et al. (2023) find that BigTech credit is strongly responsive to changes in firm-specific characteristics. This evidence suggests that technology can enable firms to borrow without collateral. Besides, some other studies (e.g., Thakor, 2020; Huang et al., 2020; Boot et al., 2021) highlight the emergence of a new characteristic of FinTech where large technology firms are providing lending services to small and medium-sized entrepreneurs without requiring collateral.

Finally, our cost-of-state-verification perspective differs from existing technology-related ones such as fast data processing abilities (e.g., Fuster et al., 2019) or new credit-sorting models (e.g., Gambacorta et al., 2019). More importantly, our interpretation on the key difference between BigTech and banks can help explain cross-country differences in BigTech credit accesses (e.g., the U.S. and China). For instance, in order to easily verify borrowers' earnings, BigTech's data advantage should be considerable (e.g., high market share of the payment system). However, this case does not apply to large technology firms in the U.S.

1.3 Intangible Capital Story

Another way to interpret the variation in lending practices is to consider the distinction between intangible and tangible capital. Specifically, we assume that intangible capital has a relatively low liquidation value.⁵

⁵In theory, firms with more intangibles may have lower liquidation values simply because they have severe agency problems. However, in this paper, we do not distinguish between the asset's actual resalability and the value damage from agency frictions such as managers' business stealing behaviors.

Assumption 2 *Intangible capital has a low liquidation value l .*

Usually, when a business undergoes bankruptcy, it is liquidated. This process involves a fast sale of physical assets like machinery and facilities, whose value can be easily determined, as well as intangible assets, whose worth is challenging to determine from an external perspective. Our assumption here is consistent with the existing literature, which posits that the cost of liquidation decreases as the tangibility of investments increases (e.g., Beck et al., 2022).

By utilizing Assumption 2, we can demonstrate that it is advantageous for lenders to enforce cash-flow-based lending if the asset tangibility falls below a certain threshold.

Lemma 2 *In the event that the liquidation value (l) is below a specific threshold (l^*), it becomes preferable for lenders to adopt cash-flow-based over asset-based lending.*

The detailed proof is provided in the appendix. Lemma 2 is based on the intuition that if the assets have a non-collateralizable nature, then it is more advantageous for lenders to opt for cash flow-based lending since it yields higher profits. This implication also has support from the previous literature on the relationship between intangible capital and cash-flow-based lending. For instance, Haskel and Westlake (2018) find that financing intangibles have less association with conventional credit restrictions.

To summarize, the coexistence of different types of borrowing constraints can be explained by either information asymmetry or intangible capital story. Of course, these two are by no means the only explanations. Other possible reasons for the simultaneous existence of these constraints can be found in other related studies such as Lian and Ma (2021) and Huang (2022). However, our focus here is not on the *micro-level origins* but on exploring their *macro-finance implications*.

2. BigTech versus Bank in the Macroeconomy

2.1 Model Setup

Our goal in this section is to investigate how different is a macroeconomy with BigTech lending, compared to the one with traditional bank lending, with a focus on allocation efficiency, financial instability, and the nature of business cycles. In order to achieve this goal, we use a standard continuous-time distributional macro model à la Moll (2014) but with two different financial sectors and endogenous default. More specifically, Consider an infinite-horizon continuous-time economy that populates $S + 1$ continua of heterogeneous entrepreneurs, \bar{L} homogeneous workers, and two representative financial intermediaries (bank \mathcal{B} and BigTech \mathcal{F}). Each worker supplies one efficient unit of labor inelastically and they are hand-to-mouth consumers. Each entrepreneur interacts with *one and only one* of the two financial intermediations, and the only difference between them is the type of borrowing constraints. Without loss of generality, we normalize the size of entrepreneurs borrowing from banks as 1. In this way, the parameter S represents the relative importance of the BigTech sector in the whole economy. In the baseline model, we set S to be 1 so that both sectors are equivalently important. In Section 3.2, we allow S to be endogenous and discuss the optimal development level of BigTech.

All entrepreneurs have the same log preference and constant-returns-to-scale (CRS) production technology. In addition, they each have the option to default and exit the market permanently. The precise timing of this option is explained in Section 2.1.3. In order to maintain a constant number of entrepreneurs in each sector, the defaulting entrepreneur is being replaced by a newborn in the subsequent period, with average wealth and productivity levels. In addition, we also assume “costless default”, which means that all the default costs are repaid by a third party (e.g., the government) so that the cost of debt is always $1 + r$ despite a non-zero default probability. We use this setup for two reasons. To begin with, this assumption allows us to avoid pricing the defaultable bond and hence simply our analysis. Furthermore, the price effect also plays an important role in determining the *de facto* tightness of borrowing constraint faced by entrepreneurs. In this

paper, we exclude the endogenous price effect and focus on the fundamental difference between these two exogenous borrowing constraints.

For the simplification of notations, we suppress the individual and time subscripts unless it is necessary.

2.1.1 Preference

All entrepreneurs in this economy share the same additive utility function shown below:

$$\mathbb{E}_0 \int_0^{\infty} e^{-\rho t} \log c_t dt \quad (7)$$

c denotes the consumption of final goods and ρ represents the rate of time preference. Our choice of log utility is aimed at simplifying the optimal consumption decision. However, it is worth noting that our main conclusions remain unchanged even if employing alternative constant relative risk aversion (CRRA) utility functions.

2.1.2 Technology

Each entrepreneur possesses a private firm that employs both capital k and labor l to produce the final consumption goods y , using the same production function depicted below:

$$y = f(z, k, l) = zk^{\alpha}l^{1-\alpha} \quad (8)$$

where z denotes the idiosyncratic productivity. Equation (8) illustrates that the production technology takes the standard Cobb-Douglas form with the parameter $\alpha \in (0, 1)$. We assume CRS technology to make borrowing constraints the most crucial role in determining the individual firm size and hence the aggregate outcomes. However, our key conclusions remain valid even though the production function has a feature of decreasing returns-to-scale.

All entrepreneurs participate in a competitive market to purchase the physical capital

and hire homogeneous workers, with prices denoted by $r + \delta$ and w , respectively. Here, r represents the risk-free interest rate, δ is the capital depreciation rate, and w is the flat wage rate.

Following the existing literature (e.g. Moll, 2014; Di Tella, 2017), we assume that the idiosyncratic productivity is stochastic and follows a standard Ornstein–Uhlenbeck process:

$$dz = \frac{1}{\theta} (\bar{\mu} - z) dt + \sigma \sqrt{\frac{1}{\theta}} d\mathcal{W} \quad (9)$$

In Equation (9), the parameter $\bar{\mu}$ represents the long-run mean level of the entrepreneur’s productivity. \mathcal{W} denotes the standard exogenous Brownian shock, which is assumed to be independent and identically distributed (i.i.d.) across various firms. In other words, under the incomplete market assumption, entrepreneurs are unable to fully hedge their individual productivity risk, and σ captures the sensitivity of z to the underlying Brownian shock.

2.1.3 Timing and Expectations

To introduce defaultable debt, we assume that the actual individual productivity realizes *after* the entrepreneur has done financing and capital & labor purchases. To elaborate further, entrepreneurs are able to establish their expectations regarding their future productivity z at the outset of each period, or in the final phase of the preceding period, utilizing the following functional form:

$$\tilde{\mathbb{E}}[z] = \tilde{\mathbb{E}}[\tilde{z} + dz] = \frac{1}{\theta} [\bar{\mu} + (\theta - 1) \tilde{z}] \quad (10)$$

\tilde{z} represents the productivity realized in the preceding period. Throughout the rest of this paper, the notation \sim and $'$ represent the preceding and subsequent period, respectively. In the baseline model, we assume that the expectation has only a rational expectation component. In the model extension, we allow for additional *algorithm bias*

component. We impose this assumption as in reality, BigTech's advantage lies in using vast amounts of data instead of collateral to assess individual firm's creditworthiness. Although it is often more accurate than banks, still, it cannot perfectly predict any firm's actual earnings. It has at least some random error terms and possibly persistent irrational biases.

Once entrepreneurs have estimated their expected productivity $\tilde{\mathbb{E}}[z]$ and taken into account their current wealth, they determine the optimal amount of capital to rent $k(a, \tilde{\mathbb{E}}[z])$ and the number of workers to hire $l(a, \tilde{\mathbb{E}}[z])$. Subsequently, the actual productivity z is realized, and entrepreneurs must make a decision on whether or not to default. As mentioned before, if they choose to default, they leave the market permanently and are replaced by a new-born entrepreneur with an average wealth \bar{a} and an initial productivity of $\bar{\mu}$. If they choose not to default, then the operating profits received during this period can be expressed as follows:

$$\pi(a, z, \tilde{z}) = \pi(a, \tilde{\mathbb{E}}[z]) \equiv zk(a, \tilde{\mathbb{E}}[z])^\alpha l(a, \tilde{\mathbb{E}}[z])^{1-\alpha} - (r + \delta)k(a, \tilde{\mathbb{E}}[z]) - wl(a, \tilde{\mathbb{E}}[z]) \quad (11)$$

At the same time, the budget constraint of any individual entrepreneur during this period can be written as below:

$$c + a' \leq a(1 + r) + \pi(a, \tilde{\mathbb{E}}[z]) = a(1 + r) + \pi(a, z, \tilde{z}) \quad (12)$$

2.1.4 Borrowing constraints

This part outlines the different borrowing constraints faced by entrepreneurs as they seek loans from different financial institutions. Generally speaking, the banking sector applies the conventional collateral-based borrowing constraint, as suggested in the classical macro-finance literature, whereas the BigTech sector adopts the earnings-based borrowing constraint, as evidenced by some recent empirical studies.

Banking sector To begin with, we assume that all entrepreneurs face the same collateral-based borrowing constraint when borrowing from a traditional bank:

$$k - a \leq \lambda_B k \quad (13)$$

The difference between operating capital stock k and personal wealth a is by definition the external debt position. The parameter $0 \leq \lambda_B \leq 1$ implies that each entrepreneur's borrowing limit from the traditional banking sector is constrained by their capital stock, due to factors such as limited enforcement or information asymmetry. As a result, the magnitude of λ_B represents the degree of severity of these frictions. To elaborate further, if λ_B is set to be 0 in Equation (13), it implies that entrepreneurs can only rely on self-financing for their capital purchases. On the other hand, if λ_B is equal to 1, it indicates that their entire capital can be externally financed. Of course, these are two extreme scenarios, and we aim to examine how the value of λ_B impacts the role of the banking sector in driving macroeconomic fluctuations.

Rewriting Equation (13) gives us the standard wealth-based borrowing constraint shown as follows:

$$k \leq \frac{1}{1 - \lambda_B} a \quad (14)$$

Similarly, Equation (14) captures the common intuition that the amount of capital available to an entrepreneur is limited by his personal wealth a and again the magnitude of λ_B captures the degree of financial development in the banking system. As Equation (14) is more comparable to the borrowing constraint in the BigTech sector, therefore we will use this equation throughout the rest of the paper.

BigTech sector In contrast, all entrepreneurs in the BigTech sector face the same earnings-based borrowing constraint shown as follows:

$$k - a \leq \lambda_{\mathcal{F}} \tilde{\mathbb{E}}[\pi] = \lambda_{\mathcal{F}} \left[\left(\frac{(1-\alpha)(r+\delta)}{\alpha w} \right)^{1-\alpha} \tilde{\mathbb{E}}[z] - \frac{r+\delta}{\alpha} \right] k \quad (15)$$

where $\tilde{\mathbb{E}}[\pi]$ is the entrepreneur's expected earnings based on the one-period ahead information set. We assume that only a certain fraction of earnings can be externally financed, i.e., $0 \leq \lambda_{\mathcal{F}} \leq 1$. Similarly, the existence of $\lambda_{\mathcal{F}}$ also comes from the limited enforcement that entrepreneurs might steal a fraction of their companies' earnings. More specifically, $\lambda_{\mathcal{F}} = 0$ refers to the situation where entrepreneurs can only self-finance, and $\lambda_{\mathcal{F}} = 1$ means that all earnings can be externally financed. Rewriting this equation can give us the wealth-based borrowing constraint in the BigTech sector as follows:

$$k \leq \frac{1}{1 - \lambda_{\mathcal{F}} \left(\zeta \tilde{\mathbb{E}}[z] - \frac{r+\delta}{\alpha} \right)} a \quad (16)$$

where $\zeta \equiv \left(\frac{(1-\alpha)(r+\delta)}{\alpha w} \right)^{1-\alpha}$.

With Equations (14) and (16), we are able to discuss the similarity and differences between these two types of constraints.

Similarity and difference In the existing literature, some papers argue that these two types of borrowing constraints are fundamentally different because earning is a concept of flows while collateral is a stock variable. However, here we argue that if we investigate the relationship between debt capacity and net worth, then there is some similarity between these two borrowing constraints. More specifically, based on the previous model setup, by the time entrepreneurs need to repay their debt at $t + dt$, earnings at t have already become part of the net worth at $t + dt$. Therefore, both borrowing constraints link debt capacity to verifiable net worth as follows:

$$\text{debt capacity} = \phi \times \text{expected verifiable net worth}$$

Therefore, if entrepreneurs' net worth can be observed, then these two types of bor-

rowing constraints are similar in the sense that they impose a requirement that the maximum amount of debt entrepreneurs can borrow is a fraction of their net worth. In other words, the classical financial mechanism (e.g., Kiyotaki and Moore, 1997; Bernanke and Gertler, 1989) still exists. Whether the arrival of new financial intermediation weakens the classical financial accelerator mechanism depends on the choice of parameters.

Comparing Equations (14) and (16), we can clearly see that the fundamental difference between these two types of borrowing constraints lies in their **cross-sectional difference**. For the entrepreneurs faced with collateral-based borrowing constraints, given their wealth, productive firms do not face any additional advantages because the tightness of borrowing constraints is the same for all entrepreneurs. However, as for entrepreneurs with earnings-based borrowing constraints, given their wealth, firms that are *expected* to be highly productive have additional advantages because the tightness of constraints is decreasing in *expected* productivity. As we will see in the following sections, such a difference is crucial for our understanding of the different macroeconomic implications of these two borrowing constraints. On one hand, BigTech lending is more efficient and it relies less on collateral value, which is not related to individual productivity. On the other hand, as the expectation is not exactly the same as the realization, BigTech credit could bring more financial instability to the whole economy.

Caveats Two caveats are worth noting. First, here we assume that entrepreneurs can borrow against their *expected future* earnings. This setting is different from Lian and Ma (2021), which assumes that a fraction of the *current* earnings can be externally financed. We adopt our modeling approach for mainly two reasons. On one hand, the setup of borrowing against the future is closer to what happens in reality. On the other hand, it allows us to further explore how it affects not only allocation efficiency but also financial instability.

Second, in our framework, we assume that the only difference between the banking sector and the BigTech sector lies in the type of borrowing constraints. However, it does

not mean that in reality, this is the only difference between these two sectors. For example, with the US loan-level data on mortgage applications and origination, Fuster et al. (2019) show that FinTech lenders originate mortgages faster and screen borrowers more effectively compared to other lenders. Philippon (2016) suggest that FinTech can lower the costs of financial services provided by financial intermediations. Thakor and Merton (2019) have developed a theory of bank and non-bank lending in which banks have an endogenous advantage over non-bank lenders when it comes to being trusted to make good loans because banks possess an advantage in developing investor trust due to their unique access to low-cost deposit funding. However, the reason why we focus on this specific difference is that, in the existing macro-finance literature, the type of corporate borrowing constraint is essential to our macroeconomic analysis of financial frictions.

Therefore, the state of the economy can be summarized by the joint distributions of wealth and (current and future) productivity $\omega^j(t, a, z, \bar{z})$, where $j \in \{\mathcal{B}, \mathcal{F}\}$. At period t , each entrepreneur within sector $j \in \{\mathcal{B}, \mathcal{F}\}$ is indexed by their current productivity z , past performance \bar{z} , and personal wealth a . In addition, we assume that the aggregate shocks being investigated are “M.I.T. shocks”, which are unexpected shocks that occur when the economy is in its steady state and lead to a transition towards a new one. Specifically, shocks to the parameter $\bar{\mu}$ are interpreted as shocks to aggregate productivity, and those to σ are considered shocks to the level of micro-level uncertainty.

2.1.5 Definitions

Equilibrium definition A stationary recursive competitive equilibrium consists of prices $\{r_t, w_t\}_{t=0}^{\infty}$ and resource allocations $\left\{ (l_{it}^{\mathcal{B}}, k_{it}^{\mathcal{B}}, c_{it}^{\mathcal{B}})_{i \in [0,1]}; (l_{it}^{\mathcal{F}}, k_{it}^{\mathcal{F}}, c_{it}^{\mathcal{F}})_{i \in [0,S]} \right\}_{t=0}^{\infty}$ that satisfy the following three conditions:

1. **Optimization:** given market prices $\{r_t, w_t\}_{t=0}^{\infty}$, resource allocations

$$\left\{ \left(l_{it}^{\mathcal{B}}, k_{it}^{\mathcal{B}}, c_{it}^{\mathcal{B}} \right)_{i \in [0,1]} ; \left(l_{it}^{\mathcal{F}}, k_{it}^{\mathcal{F}}, c_{it}^{\mathcal{F}} \right)_{i \in [0,S]} \right\}_{t=0}^{\infty}$$

maximize each entrepreneur's life-time utility (7) subject to constraints (8), (12), (14), (16), and initial endowment $\left\{ \left(a_{i0}^{\mathcal{B}}, k_{i0}^{\mathcal{B}}, z_{i0}^{\mathcal{B}} \right)_{i \in [0,1]} ; \left(a_{i0}^{\mathcal{F}}, k_{i0}^{\mathcal{F}}, z_{i0}^{\mathcal{F}} \right)_{i \in [0,S]} \right\}$.

2. **Market clearance:** market prices $\{r_t, w_t\}_{t=0}^{\infty}$ satisfy the following conditions

- labor market clears at any time t

$$\iiint l^{\mathcal{B}}(t, a, z, \tilde{z}) \omega^{\mathcal{B}}(t, a, z, \tilde{z}) dadzd\tilde{z} + \mathcal{S} \iiint l^{\mathcal{F}}(t, a, z, \tilde{z}) \omega^{\mathcal{F}}(t, a, z, \tilde{z}) dadzd\tilde{z} = \bar{L} \quad (17)$$

- capital market clears at any time t

$$\sum_{j \in \{\mathcal{B}, \mathcal{F}\}} \mathcal{S}^j \iiint b^j(t, a, z, \tilde{z}) \omega^j(t, a, z, \tilde{z}) dadzd\tilde{z} = \sum_{j \in \{\mathcal{B}, \mathcal{F}\}} \mathcal{S}^j \iiint a^j(t, a, z, \tilde{z}) \omega^j(t, a, z, \tilde{z}) dadzd\tilde{z} \quad (18)$$

where $\mathcal{S}^{\mathcal{B}} = 1$ and $\mathcal{S}^{\mathcal{F}} = \mathcal{S}$.

3. **Stationary distribution:** the wealth distributions $\omega^j(t, a, z, \tilde{z})$ obey entrepreneur's optimal decision and the exogenous productivity process (9), and they are stationary, i.e., $\frac{\partial \omega^j(t, a, z, \tilde{z})}{\partial t} = 0$.

Aggregate productivity Following the standard literature, aggregate and sectoral productivity are defined as follows:

$$\mathcal{Z} \equiv \frac{\mathcal{Y}}{\mathcal{K}^{\alpha} \bar{\mathcal{L}}^{1-\alpha}} = \frac{\sum_{j \in \{\mathcal{B}, \mathcal{F}\}} \mathcal{S}^j \iiint y^j(t, a, z, \tilde{z}) \omega^j(t, a, z, \tilde{z}) dadzd\tilde{z}}{\left(\sum_{j \in \{\mathcal{B}, \mathcal{F}\}} \mathcal{S}^j \iiint a^j(t, a, z, \tilde{z}) \omega^j(t, a, z, \tilde{z}) dadzd\tilde{z} \right)^{\alpha} \bar{\mathcal{L}}^{1-\alpha}}$$

$$\mathcal{Z}^j \equiv \frac{y^j}{(\mathcal{K}^j)^{\alpha} (\mathcal{L}^j)^{1-\alpha}} = \frac{\iiint y^j(t, a, z, \tilde{z}) \omega^j(t, a, z, \tilde{z}) dadzd\tilde{z}}{\left(\iiint k^j(t, a, z, \tilde{z}) \omega^j(t, a, z, \tilde{z}) dadzd\tilde{z} \right)^{\alpha} \left(\iiint l^j(t, a, z, \tilde{z}) \omega^j(t, a, z, \tilde{z}) dadzd\tilde{z} \right)^{1-\alpha}}$$

where $j \in \{\mathcal{B}, \mathcal{F}\}$.

Allocation efficiency If there were no financial frictions, in our model framework, only the most productive firm produces in the equilibrium. We use Z^U to denote the upper bound of the exogenous productivity distribution. Therefore, the aggregate and sectoral allocation efficiency can be calculated as follows:

$$\mathcal{E} = \frac{Z}{Z^U}, \mathcal{E}^j = \frac{Z^j}{Z^U}$$

Default probability Aggregate and sectoral default probability can be calculated as follows:

$$\begin{aligned} \mathcal{P} &\equiv \frac{1}{S+1} \sum_{j \in \{\mathcal{B}, \mathcal{F}\}} S^j \iiint \mathbb{1}_{V^j(t, a, z, \tilde{z}) < 0} \omega^j(t, a, z, \tilde{z}) da dz d\tilde{z} \\ \mathcal{P}^j &\equiv \iiint \mathbb{1}_{V^j(t, a, z, \tilde{z}) < 0} \omega^j(t, a, z, \tilde{z}) da dz d\tilde{z} \end{aligned}$$

where V represents the value function representation of the individual entrepreneur's optimization problem.

2.2 Characterizing Equilibrium

2.2.1 Individual optimal decisions

To begin with, we characterize the optimal policy functions for each individual entrepreneur. The optimal decisions of capital holdings can be summarized as in Lemma 3.

Lemma 3 *Given the market prices r and w , there is the same productivity cutoff for being active \tilde{z} for entrepreneurs in both sectors. More specifically, the optimal capital holdings for entrepreneurs in the banking sector are*

$$k^B(a, z, \tilde{z}) = \begin{cases} \frac{a}{1-\lambda_B} & \tilde{\mathbb{E}}[z] \geq \tilde{z} \\ 0 & \tilde{\mathbb{E}}[z] < \tilde{z} \end{cases} \quad (19)$$

Meanwhile, the optimal capital holdings for entrepreneurs in the BigTech sector are

$$k^F(a, z, \tilde{z}) = \begin{cases} \frac{1}{1-\lambda_F(\zeta\tilde{\mathbb{E}}[z]-\frac{r+\delta}{\alpha})} a & \tilde{\mathbb{E}}[z] \geq \tilde{z} \\ 0 & \tilde{\mathbb{E}}[z] < \tilde{z} \end{cases} \quad (20)$$

where $\tilde{z} = \left(\frac{r+\delta}{\alpha}\right)^\alpha \left(\frac{w}{1-\alpha}\right)^{1-\alpha}$.

Given our assumption on the constant return-to-scale technology and frictionless labor market, the marginal product of capital is always linear in expected productivity $\tilde{\mathbb{E}}[z]$. In this case, the optimal capital choice is a corner solution: it is zero for entrepreneurs with expected productivity lower than some threshold \tilde{z} , and the maximal amount of borrowing for entrepreneurs with higher expected productivity. At the same time, these inactive entrepreneurs lend all their wealth to the market so that they can get a constant return r . The cutoff \tilde{z} is the same for the two sectors simply because we assume the production technology in these two sectors is the same. In addition, as the BigTech companies do not have perfect foresight on entrepreneurs' future earnings, they can only base their lending behaviors on the *expected* earnings, instead of the realized ones.

Another important observation from Lemma 3 is that compared to entrepreneurs in the traditional banking sector, productive firms in the BigTech sector get to use more leverage and thus have additional advantages in lending and capital accumulation. As shown in the following lemma, these advantages eventually reflect on the wealth evolution dynamics.

Lemma 4 *With preference assumption (7), the entrepreneur's wealth a in both sectors evolves*

according to the following equations:

$$da_{\mathcal{B}} = \left\{ 1_{\tilde{\mathbb{E}}[z] \geq \tilde{z}} \times \left[\frac{\zeta z - \frac{r+\delta}{\alpha}}{1 - \lambda_{\mathcal{B}}} + r - \rho \right] + 1_{\tilde{\mathbb{E}}[z] < \tilde{z}} \times (r - \rho) \right\} a_{\mathcal{B}} dt \equiv \Gamma^{\mathcal{B}}(z, \tilde{z}) a_{\mathcal{B}} dt \quad (21)$$

$$\begin{aligned} da_{\mathcal{F}} &= \left\{ 1_{\tilde{\mathbb{E}}[z] \geq \tilde{z}} \times \left[\frac{\zeta z - \frac{r+\delta}{\alpha}}{1 + \lambda_{\mathcal{F}} \left(\frac{r+\delta}{\alpha} - \zeta \tilde{\mathbb{E}}[z] \right)} + r - \rho \right] + 1_{\tilde{\mathbb{E}}[z] < \tilde{z}} \times (r - \rho) \right\} a_{\mathcal{F}} dt \\ &= \left\{ 1_{\tilde{\mathbb{E}}[z] \geq \tilde{z}} \times \left[\frac{\zeta z - \frac{r+\delta}{\alpha}}{1 + \lambda_{\mathcal{F}} \left[\frac{r+\delta}{\alpha} - \zeta \left(\frac{1}{\theta} - \gamma \right) \bar{\mu} - \frac{\zeta(\theta-1+\theta\gamma)}{\theta} \tilde{z} \right]} + r - \rho \right] + 1_{\tilde{\mathbb{E}}[z] < \tilde{z}} \times (r - \rho) \right\} a_{\mathcal{F}} dt \\ &\equiv \Gamma^{\mathcal{F}}(z, \tilde{z}) a_{\mathcal{F}} dt \end{aligned} \quad (22)$$

For simplicity, we rewrite them as follows:

$$da_j = \Gamma^j(z, \tilde{z}) a_j dt \quad (23)$$

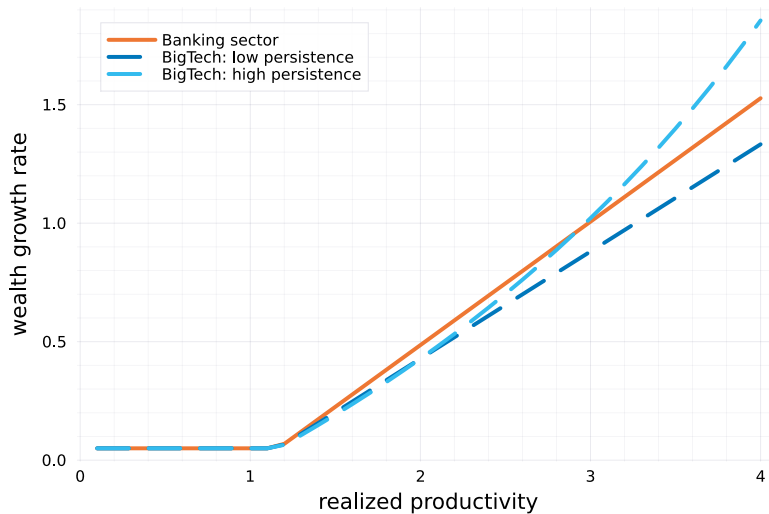
where Γ is the wealth growth rate function, and it depends on the entrepreneur's idiosyncratic productivity z, \tilde{z} and the sector $j \in \{\mathcal{B}, \mathcal{F}\}$ he belongs to.

The detailed proof is provided in the appendix. The key messages in Lemma 4 are threefold. First, the wealth growth rates of firms with relatively low *expected* productivity are the same in both sectors. With the assumption of log-utility, the optimal consumption choice is always a constant ρ fraction of wealth, where ρ represents the time value. For entrepreneurs with relatively low *expected* productivity, i.e., $\tilde{\mathbb{E}}[z] < \tilde{z}$, they do not produce anything on the market. Instead, they lend all their wealth to those with relatively high *expected* productivity, i.e., $\tilde{\mathbb{E}}[z] \geq \tilde{z}$. As a result, their growth rate is always set to be $r - \rho$, and this number does not depend on which financial sector they belong to.

Second, the wealth growth rates of highly productive entrepreneurs are different in these two sectors. As we can see from Equation (21), for any active producer borrowing from the banking sector, his wealth growth rate is a linear function in the actual productivity z . In contrast, the wealth growth rate of any active firm with productivity z borrowing from the BigTech sector is a function of z and \tilde{z} both. If the individual pro-

ductivity is highly persistent, then the wealth growth rate for highly productive firms is convex in (realized and expected) productivity. It means that for the productive firms in the BigTech sector, their wealth grows faster not only because their productivity is high, but also because they get to use more leverage with this earnings-based borrowing constraint. The key implications can be better illustrated in Figure 2, where we plot the relationship between the wealth growth rate and productivity for both sectors. The degree of convexity is jointly determined by the tightness of the borrowing constraint $\lambda_{\mathcal{F}}$ and the persistence of the productivity shocks θ . As we will see later, the convexity of the wealth growth rate in the BigTech sector is the underlying reason why uncertainty, i.e., the second-moment shocks, matters for the aggregate quantities both in the steady state and over the business cycles.

Figure 2: Wealth Growth Rate in Banking and BigTech



Third, we can see a clear efficiency-instability tradeoff for the BigTech based on Equations (21) and (22). The advantage of BigTech credit is that if the productivity persistence is high or the BigTech does predict the earnings accurately, then more resources can be efficiently allocated towards productive firms in the BigTech sector. However, if the productivity has some random components or the prediction is biased, then BigTech credit

could lead to severe overlending or underlending problems. As we will see later, this special feature of BigTech as a new financial intermediation makes the whole economy relatively more unstable than the traditional economy with banks only.

2.2.2 Dynamics of wealth distributions

After discussing the optimal policy function for any individual entrepreneur, now we turn to characterize how the wealth and productivity joint distributions in both sectors evolve over time. With the exogenous productivity process (9) and the endogenous wealth process (23), the distribution in each sector evolves according to the following equations:

$$\begin{aligned} \frac{\partial \omega^j(t, a, z, \tilde{z})}{\partial t} = & - \frac{\partial [\Gamma^j(t, z, \tilde{z}) a \omega^j(t, a, z, \tilde{z})]}{\partial a} - \frac{\partial [\frac{1}{\theta} (\bar{\mu} - z) \omega^j(t, a, z, \tilde{z})]}{\partial z} - \frac{\partial [\frac{1}{\theta} (\bar{\mu} - \tilde{z}) \omega^j(t, a, z, \tilde{z})]}{\partial \tilde{z}} \\ & + \frac{\sigma^2}{2\theta} \frac{\partial^2 [\omega^j(t, a, z, \tilde{z})]}{\partial \tilde{z}^2} + \frac{\sigma^2}{2\theta} \frac{\partial^2 [\omega^j(t, a, z, \tilde{z})]}{\partial z^2} \text{ where } j \in \{\mathcal{B}, \mathcal{F}\} \end{aligned} \quad (24)$$

Generally speaking, what determines the evolution of wealth distribution in this economy is a system of high-dimensional partial differential equations (PDEs). Previous works, including Kiyotaki (1998), Caselli and Gennaioli (2013), and Moll (2014), use wealth shares to characterize aggregates so that we can save one state variable. However, this method is not applicable here as we have two different sectors in this economy. Scaling the wealth by using aggregate capital and getting wealth share cannot reduce the number of state variables. Therefore, we follow Raissi et al. (2019) and use a Physics-informed neural network (PINN) approach to numerically solve the dynamics of ω^j . This deep learning method utilizes the advantages of deep neural networks to solve high-dimensional PDEs and it has significantly reduced time and memory costs compared to those classical methods such as finite difference and finite element. The advantage comes from the fact that the algorithm does not require interpolation and coordinate transformation because it is a universal nonlinear approximator (Bach, 2017) and thus avoids the

curse of dimensionality. In addition, it can overcome the local optimization problem by introducing some penalty factors or stochasticity into the loss function.

2.3 Parameterization

Table 1 summarizes the values of different parameters used in our macroeconomic framework. To remain consistent with the previous studies, we choose the rate of time preference ρ , capital share α , and labor market size \mathcal{L} to be 0.05, 0.33, and 1.0, respectively. The average capital depreciation rate is computed using the Bureau of Economic Analysis (BEA) Fixed Assets Tables dataset, resulting in a value of 6%. In the baseline model, the values for $\bar{\mu}$, productivity persistence, and idiosyncratic productivity are determined to be 0, 0.85, and 0.56, respectively, following Asker et al. (2014). These productivity-related parameter choices align with the actual firm-level TFP measure in the data. With these parameter values, we are able to explore the macro-finance implications of BigTech.

Table 1: Parameterization

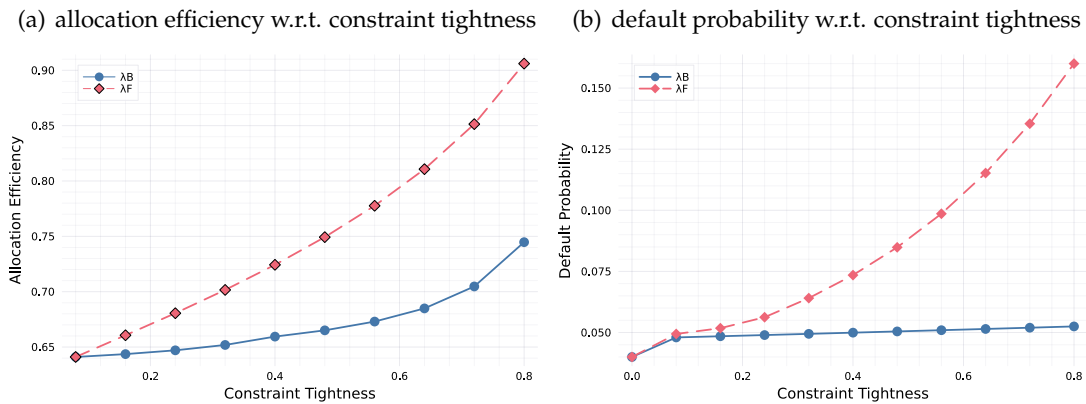
Parameter	Description	Value	Source/Reference
ρ	rate of time preference	0.05	
α	capital share	0.33	Moll (2014)
\mathcal{L}	labor market size	1.0	
δ	capital depreciation rate	0.06	BEA-FAT
\mathcal{S}	size of BigTech	1.0	
$\bar{\mu}$	log idiosyncratic productivity mean	0.0	
θ	autocorrelation $e^{-\theta}$	0.16 (corr = 0.85)	Asker et al. (2014)
σ	log idiosyncratic productivity s.d.	0.56	

2.4 Implications

2.4.1 Steady state: efficiency-instability tradeoff

Our first key model implication is that there is an efficiency-instability tradeoff associated with BigTech development in the steady state. Figure 3 summarizes our main results on the equilibrium aggregate allocative efficiency and default probability. In these two graphs, we use the tightness of constraints in the banking sector λ_B and that in the BigTech sector λ_F as proxies for the development of each financial institution, and then present the corresponding impacts on the aggregate productivity and default probability defined as in Section 2.1.5. Based on these two graphs, we can clearly see that BigTech is efficient in resource allocation but also associated with a higher default probability. In contrast, the traditional banking sector is less efficient but also less risky.

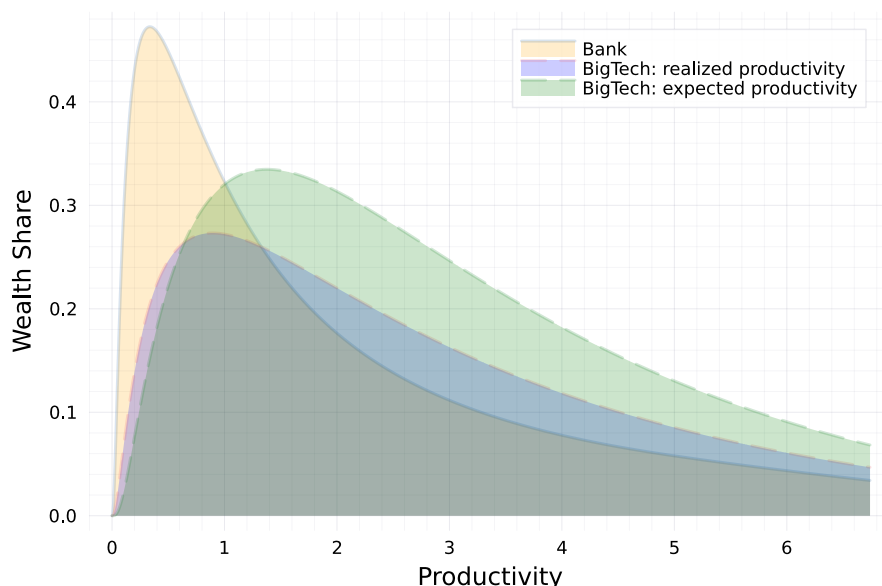
Figure 3: Efficiency-Instability Tradeoff



The underlying mechanism is closely related to the different types of borrowing constraints associated with these two financial institutions. In the banking sector, the maximum debt that can be borrowed is linked to the capital stock, whereas in the BigTech sector, the upper limit of debt financing is directly related to expected productivity. As a result, borrowing constraints for highly expected productive firms are looser in the BigTech sector, resulting in a more concentrated equilibrium wealth towards productive firms and lower capital misallocation. This result can be shown in Figure 4, where we plot the equilibrium wealth share distributions of firms with different (*realized or expected*)

productivity in both sectors. As we can see, the equilibrium wealth share of right-tail firms is higher in the BigTech sector. In other words, the development of BigTech is crucial in reducing the wedges between marginal products of capital compared to the development of the traditional banking sector. As a result, the aggregate allocative efficiency increases as the BigTech sector grows.

Figure 4: Wealth Concentrations in Banking and BigTech



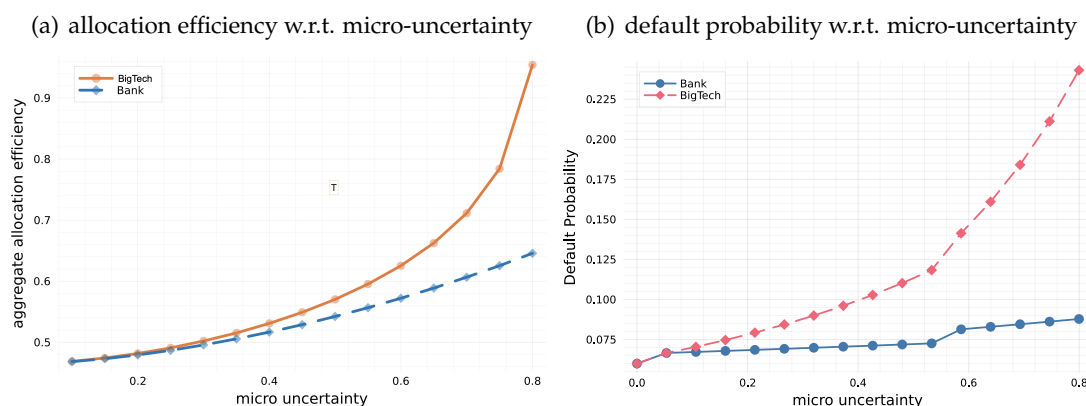
However, at the same time, as BigTech lending is linked to *expected* instead of *realized* productivity, there could be some overlending issues. Similarly, in Figure 4, we present the equilibrium wealth share distributions of firms with respect to both *expected* and *realized* productivity. Compared to that of *realized* productivity, wealth share is more concentrated in terms of *expected* productivity in BigTech, especially on the right tail distribution. It indicates that there exist some severe overlending issues in BigTech, which contributes to the increasing default probability alongside the BigTech development.

Generally speaking, we observe an efficiency-instability tradeoff associated with this new financial institution. In addition, this tradeoff is different from what is documented in the shadow banking system literature. The efficiency-instability tradeoff within the

banking sector usually comes from risk-taking incentives and/or changes in regulatory policies. However, the tradeoff in our paper comes from this new type of lending practice associated with BigTech.

After that, we examine the link between micro-uncertainty and earnings-based borrowing constraints, which turns out to be crucial for exploring the amplification and propagation mechanism in Section 2.4.2. The two graphs in Figure 5 show that within the BigTech sector, both aggregate allocative efficiency and default probability increase as the dispersion of underlying firm-level productivity grows.

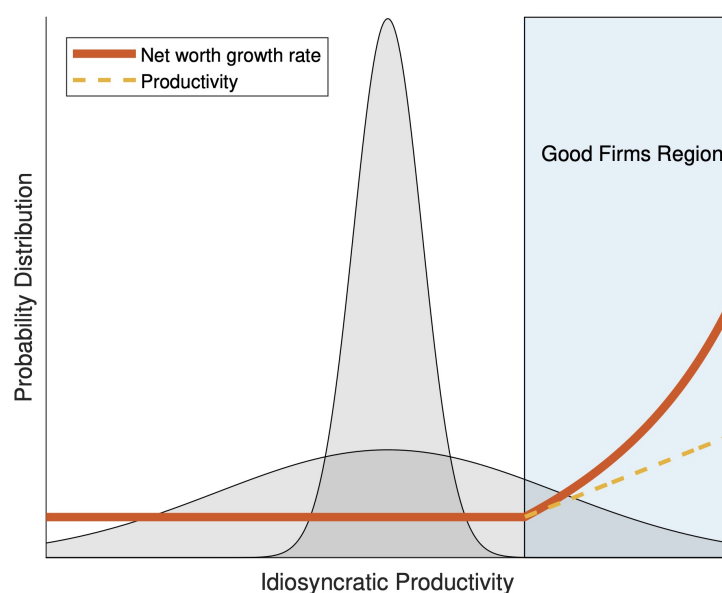
Figure 5: Efficiency-Instability Tradeoff: the role of micro-uncertainty



This conclusion may initially seem counter-intuitive, but it can be explained by the “expected good-firms principle” in Figure 6. The intuition can be explained as follows. In an economy with greater micro-level uncertainty, there is a wider range of highly unproductive and productive firms, leading to greater productivity dispersion. However, unproductive entrepreneurs are not essential for the aggregate economy as they are not operative and are merely the suppliers of funds. Therefore, it only leads to a larger number of active and productive entrepreneurs when productivity dispersion is higher. As previously mentioned, the earnings-based borrowing constraint generates an asymmetric net worth growth rate, with unproductive firms experiencing a wealth growth rate equal to the market interest rate, while productive firms experience a convexly increasing

wealth rate in line with their productivity. This feature distinguishes it from the standard collateral-based borrowing constraint and enables productive firms to rapidly build their net worth by utilizing more leverage. As a result, an increase in micro-level uncertainty results in positive outcomes for aggregate productivity and hence the BigTech sector is sensitive to second-moment shocks. We name this mechanism the “(expected) good firms principle” in reference to the “good news principle” proposed by Bernanke (1983). At the same time, as lending is based on expected earnings instead of realized earnings, the default probability also increases alongside the enhanced allocation efficiency.

Figure 6: Expected Good-Firms Principle



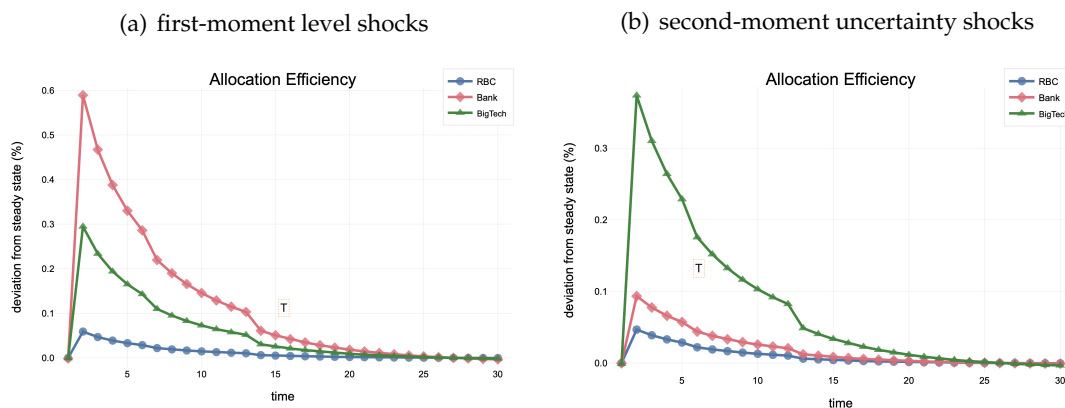
Our theoretical implications in this section have important policy implications, especially for emerging countries with underdeveloped financial markets, where financial frictions can hinder capital and wealth accumulation, ultimately slowing economic growth. In those countries with underdeveloped banking systems, if superstar firms, especially those tech giants, can offer financial services to other firms, they have the potential to narrow the income per capita gap with developed economies as the technological advantage of these superstar firms is better suited to improving aggregate capital alloca-

tive efficiency than the traditional banking sector. However, at the same time, these tech giants cannot replace the traditional banking sector as they also bring higher financial instability. Therefore, as discussed later in Section 3.2, there exists an optimal balance between these two types of financial institutions.

2.4.2 Business cycles: a different financial accelerator

Now we turn to explore the business cycle implications of BigTech. More specifically, we examine its impacts by conducting two experiments and then comparing the corresponding results to a real business cycle (RBC) benchmark model without financial frictions. The two experiments are the first-moment productivity level shocks $\Delta\bar{\mu}$ and the second-moment micro-uncertainty shocks $\Delta\sigma$. For each experiment, we calculate impulse responses from three models: the benchmark RBC model, a model with only the banking sector, and a model with only the BigTech sector. We set the tightness of borrowing constraints to be $\lambda_B = \lambda_F = 0.2$, but our findings do not depend on this specific choice. We report the behaviors of all variables relative to their steady-state values due to differing steady-state values across the models.

Figure 7: Interactions with Different Economic Fundamental Shocks



In the first experiment, we investigate the response of the economy to a one-time shock to the level of productivity. Graph (a) in Figure 7 illustrates the main effects of

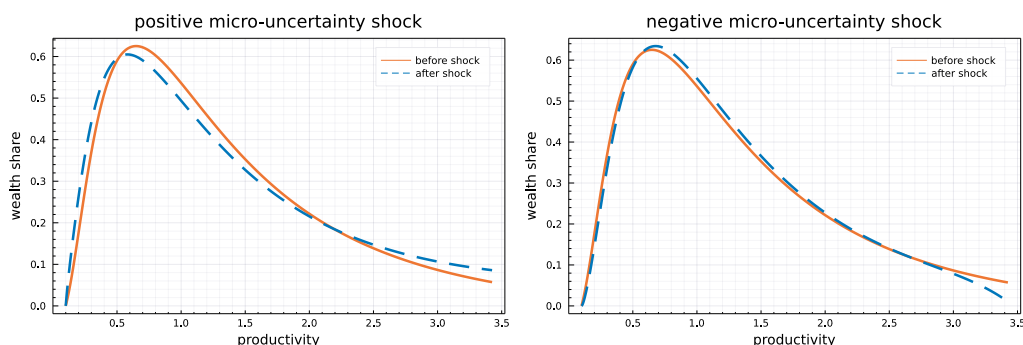
this temporary impact. The benchmark RBC model shows that there are no significant amplified effects on the aggregate economy, and the impact declines quickly over time. However, in a model with a collateral-based borrowing constraint, the increase in the aggregate productivity level is substantial and persistent over time. This result is consistent with the classical financial accelerator literature, which emphasizes the importance of credit-market frictions as drivers of cyclical economic fluctuations in the presence of asymmetric information and agency problems in credit markets. As for the model with BigTech, the patterns are similar to those in the model with banks, but the magnitudes are smaller and the effects are less persistent with the chosen parameter values.

In the second experiment, we examine the impact of a one-time shock to micro-level uncertainty, which affects only the dispersion of individual productivity but has no impact on the mean. Graph (b) in Figure 7 shows that a positive shock to micro-level uncertainty has no effect in the benchmark or the economy with only a banking sector, but has significant impacts and propagation effects when a BigTech sector with earnings-based borrowing constraint is present.

The underlying mechanism can be explained as follows. As demonstrated in Figure 8, a positive (or negative) shock to micro-uncertainty produces significant and prolonged changes in wealth inequality, which drives up (or down) the demand for investment, leading to a positive (or negative) feedback loop. The slow decay of entrepreneurial net worth inequality contributes to the persistence of the effects. Our result here implies that with a new BigTech sector, small changes in micro-level uncertainty could generate significant economic fluctuations. In contrast, these features are negligible in the benchmark RBC economy or in the economy with only a banking sector, as an asymmetric wealth growth rate is a unique characteristic of BigTech.

Our results here are related to but also different from several papers documenting the real effects of second-moment shocks (e.g., Bloom, 2009). In the standard uncertainty literature, there are *negative* impacts of uncertainty on the real economy, as higher uncertainty causes firms to temporarily pause their investment and hiring, leading to a decline

Figure 8: Underlying Mechanism: net worth inequality

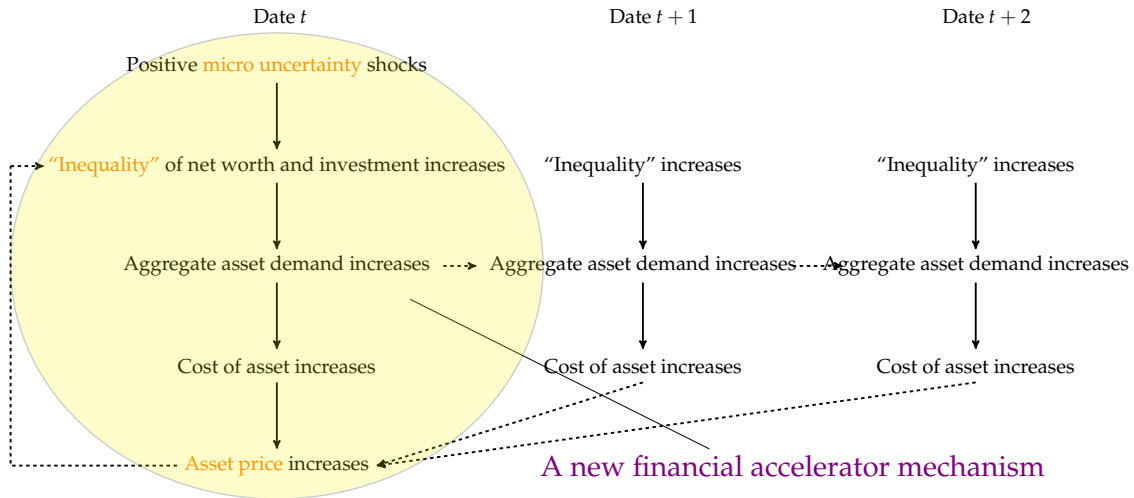


in productivity growth. In contrast, in our framework, as the BigTech sector develops, the economy becomes more sensitive to uncertainty shocks in a *positive* relationship. The uncertainty effects in our paper are more related to the growth option literature and the Oi-Hartman-Abel effects (e.g., Oi, 1961; Hartman, 1972; Abel, 2014). The difference comes from the fact that the underlying mechanism in this paper comes from the feedback loop between earnings-based borrowing constraints and micro-level uncertainty, instead of the existence of (fixed) adjustment costs.

Generally speaking, our main conclusion in this section is that BigTech acts as a new type of financial accelerator, and it is different from the classical one associated with the banking industry (e.g., Bernanke and Gertler, 1989; Kiyotaki and Moore, 1997) in three different ways. First, the primitive shock comes from *micro-uncertainty*, not *aggregate productivity*. Second, the financial friction arises from *earnings*-based borrowing constraints, not *collateral*-based ones. Third, the feedback loops occur between net worth *inequality*, not net worth *level*, and asset prices. We summarize this new financial accelerator mechanism in Figure 9.

Of course, the effectiveness of this new financial accelerator mechanism relies heavily on the stringency of the borrowing constraint and the persistence of the productivity process. Figure A5 and A6 in the appendix demonstrate that the significance of this type of financial accelerator mechanism in propelling economic fluctuations is greater when the borrowing constraint is less restrictive and the productivity process has greater

Figure 9: BigTech as a Different Financial Accelerator



persistence. This conclusion is driven by how these variables affect the wealth share of the most productive firms in the economy.

3. Extensions

3.1 The Role of Algorithm Bias

In this section, we extend our baseline model to allow for extrapolative belief and investigate the role of algorithm bias in both bank and BigTech lending. More specifically, in this extension, the expected future productivity takes the following new functional form:

$$\tilde{\mathbb{E}}[z] = \tilde{\mathbb{E}}[\tilde{z} + dz] = \underbrace{\frac{1}{\theta} [\bar{\mu} + (\theta - 1) \tilde{z}]}_{\text{rational expectation}} + \underbrace{\gamma (\tilde{z} - \bar{\mu})}_{\text{algorithm bias}} \quad (25)$$

The difference is that now the expected future productivity is composed of two different parts: *rational expectation* and *algorithm bias*. The first part is based on the actual productivity diffusion process shown as in (9), meanwhile the latter assumes that BigTech lending has a tendency to overestimate the likelihood of a positive future state when current news is favorable, and vice versa when current news is negative. This degree of

algorithm bias is captured by the parameter γ . More specifically, when $\gamma = 0$, the model reverts to the rational expectations framework, where the machine learning techniques adopted in large technology companies are unbiased. In contrast, when $\gamma > 0$, the expectations incorporate conditional mean shifts that extrapolate based on recent news. For instance, after receiving positive news, the algorithm assigns a higher probability to future positive states and a lower probability to future negative states, compared to the objective distribution. It is important to note that this distortion of the algorithm is only on conditional expectations, and unconditional forecasts remain unbiased because the average innovation $z - \bar{\mu}$ is, by definition, zero. Besides, this modified extrapolative expectation framework is forward-looking and satisfies the law of iterated expectation. Generally speaking, our modeling approach here is a slightly modified version of the standard extrapolative belief literature, as described in some recent studies such as Bordalo et al. (2018).

We introduce this assumption of algorithm bias to explore its potential amplified effects on the macro outcomes. We choose the parameter value of algorithm bias γ to be 0.4 in order to match the observed default probability of BBB-rated corporate bonds in the data (i.e., 1.5%).

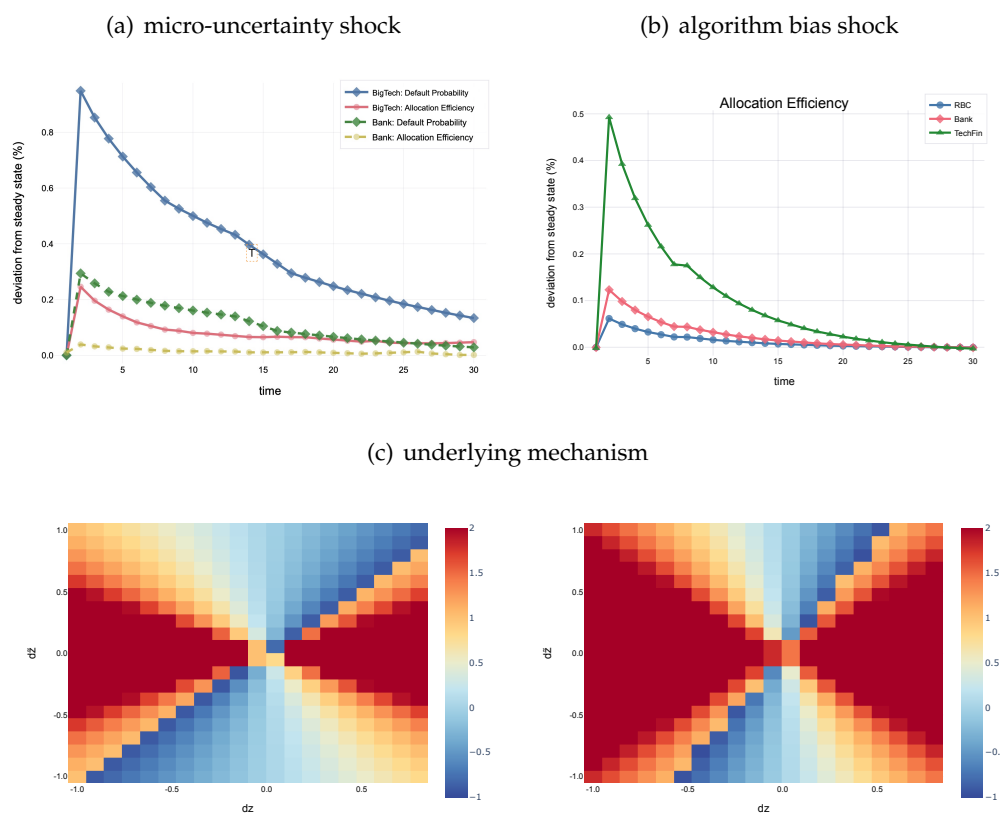
Our main findings in this section are threefold. First, the introduction of algorithm bias amplifies the efficiency-instability tradeoff associated with the BigTech sector. As shown in Graph (a) of Figure 10, after a positive one-time shock to micro-level uncertainty, both banking and BigTech sectors observe a long and persistent boom in aggregate productivity. In addition to the magnitude difference, another crucial difference between these two sectors is the default probability alongside these output booms. For entrepreneurs borrowing from BigTech, a temporary increase in the dispersion of individual productivity can lead to a substantial and persistent increase in the default probability. However, these effects on entrepreneurs borrowing from banks are very limited. Therefore, after a one-time positive shock to micro-level uncertainty, within the BigTech industry, we can observe the fragile boom phenomenon: the whole economy grows at

the cost of increasing default probability. In contrast, such an impact is negligible if we live in an economy with banks only.

Second, we investigate the heterogeneous impacts of a one-time shock to algorithm bias in three different economies (i.e., RBC benchmark, bank-only, and BigTech only), and the impacts are similar to the previous micro-uncertainty shocks. The main results are presented in Graph (b) in Figure 10. As we can see from this graph, the standard RBC model and the model with banks only are both insensitive to shocks to the algorithm bias $\Delta\gamma$. This result is intuitive as the borrowing constraints or the resulting corporate leverages are not closely related to the bias of individual productivity expectation. As mentioned before, the distortion of this algorithm bias is only on conditional expectations. Meanwhile, the unconditional forecasts remain unbiased. Therefore, for collateral-based borrowing constraints, changes in the degree of algorithm bias only generate some negligible impacts, as they affect more on the dispersion of expected earnings, not the average level. At the same time, we can see that a temporary shock to the algorithm bias can lead to persistent and amplified impacts in an economy with a BigTech sector. The only difference is that a micro-uncertainty shock alters the *realized* distribution of individual productivity, but an algorithm bias shock changes the *perceived* distribution.

Third, Graph (c) of Figure 10 illustrates the economic mechanism behind this phenomenon. Both figures present the overlending or underlending situations in the economy. The left and right graphs represent the economy before and after a temporary micro-uncertainty shock, respectively. If the color is more leaning towards red, then it means that we have a severer overlending problem. In contrast, if the color is more leaning towards blue, it means that we have more underlending issues. The definitions of overlending and underlending are the same as before. As we can see from this graph, after a positive temporary shock to micro-level uncertainty, we have more overlending problems in the economy due to the earnings-based borrowing constraint associated with the BigTech companies. As a result, we expect a higher default probability despite that it can generate an increase in aggregate output and resource allocation efficiency.

Figure 10: Fragile Boom in BigTech



Our findings in this section are consistent with the view that financial markets are *less* stable in booms than in recessions. For instance, Minsky (1992)'s financial instability hypothesis states that economic prosperity encourages borrowers and lenders to be reckless. As a result, seeds of recessions are sowed in booms. Besides, Greenspan and Shiller (2016)'s irrational exuberance claims that an overheated economy generates bubbles. Our theoretical implications on financial stability are consistent with these hypotheses. More importantly, our exercise here indicates that the rising BigTech sector brings more financial instability, especially in the presence of algorithm biases, and the regulators should pay serious attention to it.

3.2 Optimal BigTech Development

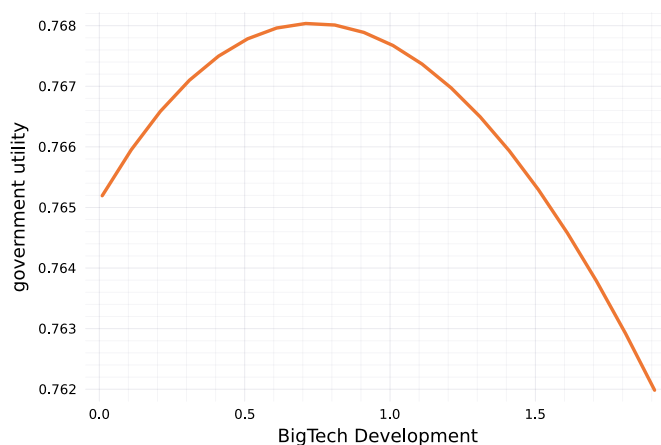
Finally, we endogenous the value of \mathcal{S} to investigate the optimal size of BigTech. As mentioned before, although the existence of a new BigTech sector increases resource allocation efficiency, it also decreases financial stability in the whole economy. In order to explore the optimal size of the BigTech sector, we make the following assumption on the preference of the central government:

$$\mathcal{U}(\mathcal{S}) = \mathcal{Z}^\iota(\mathcal{S}) - \eta\mathcal{P}(\mathcal{S}) \quad (26)$$

where \mathcal{Z} and \mathcal{P} represent the equilibrium aggregate productivity and default probability defined as before. The parameters ι and η represent how much the central government cares about resource allocation efficiency and financial stability, respectively. The functional form \mathcal{U} is similar to the mean-variance analysis widely used in the asset pricing literature. Here we use this utility function to capture the efficiency-risk tradeoff of BigTech and also measure how much financial instability the central government is willing to take on in exchange for allocation efficiency.

Similarly, both \mathcal{Z} and \mathcal{P} depend on the relative size of BigTech \mathcal{S} . Therefore, we can plot the relationship between the central government's utility and the BigTech development in Figure 11. As we can see from this figure, the relationship is not monotone. Instead, for any given values of ι and η , there exists an optimal size \mathcal{S}^* of the BigTech sector in the economy that maximizes the utility of the regulator. This non-monotonicity reflects the efficiency-risk tradeoff in our model. Our exercise here implies that although the technology advantage allows the BigTech companies to perform more efficiently with earnings-based borrowing constraints, it does not mean that the new BigTech should fully replace the traditional banking sector, as the former can bring more instability to the system. Therefore, for a social planner who cares about efficiency and stability both, there exists an optimal size of this new financial intermediation.

Figure 11: Optimal BigTech Development



4. Conclusion

This paper investigates the macro-finance implications of BigTech. We introduce both a banking sector and a BigTech sector into a continuous-time general equilibrium model with incomplete markets, heterogeneous entrepreneurs, and defaultable debt. The fundamental difference between banks and BigTech lies in the types of borrowing constraints. Entrepreneurs borrowing from banks are subject to the standard collateral-based borrowing constraints. In contrast, technology or data advantages allow BigTech to resolve agency costs and perform *expected*-earnings-based lending.

Our baseline model mainly has two macro-finance implications. First, there is an efficiency-instability tradeoff associated with BigTech development. On one hand, as it allows more productive firms to use more leverage and grow faster, BigTech credit is more efficient in resource allocation compared to traditional bank loans. On the other hand, due to the difference between *expected* and *realized* earnings, BigTech credit leads to overlending issues and hence a higher default probability in the steady state. This efficiency-instability tradeoff implies that BigTech cannot fully replace the role of traditional banks. Second, as for its impacts on the business cycles, BigTech can be interpreted as a different financial accelerator. We show that a transitory micro uncertainty shock can

lead to amplified and persistent changes in allocation efficiency and aggregate productivity. This new financial accelerator mechanism, associated with a new type of financial intermediation, differs from the classic one in three aspects: micro uncertainty instead of aggregate productivity is the primitive shock; financial friction comes from earnings-based borrowing constraints instead of collateral-based ones; and the feedback loops happen between net worth inequality, instead of net worth level, and asset prices.

Finally, we use two model extensions to discuss the role of algorithm bias and optimal BigTech development, respectively. In the first extension exercise, we assume that BigTech's predicted future earnings have an extrapolative algorithm bias component: they tend to overestimate the likelihood of a positive future state when the current news is favorable, and vice versa when negative. With the presence of algorithm bias, the financial instability concern of BigTech becomes severer. In our second model extension, we endogenous the optimal size of BigTech. We show that if the government cares about both resource allocation efficiency and financial stability, our theory indicates that there exists an optimal degree of BigTech development in the whole economy.

Generally speaking, our paper develops the first step for investigating the macro-finance implications of BigTech. Our focus is on its different borrowing constraints and the corresponding misallocation and financial instability implications. Potentially, the arrival of this new financial intermediation could also affect the labor market and firm dynamics differently, which we leave for future research.

References

- Abel, Andrew B., "Optimal Investment Under Uncertainty," *American Economic Review*, 2014, 73 (1), 228–233.
- Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," *Review of Economic Studies*, 2022, 89 (1), 45–86.
- Asker, John, Allan Collard-Wexler, and Jan De Loecker, "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 2014, 122 (5), 1013–1063.
- Bach, Francis, "Breaking the Curse of Dimensionality with Convex Neural Networks," *Journal of Machine Learning Research*, 2017, 18, 1–53.
- Beck, Thorsten, Robin Döttling, Thomas Lambert, and Mathijs A. Van Dijk, "Liquidity Creation, Investment, and Growth," *Journal of Economic Growth*, 2022. <https://doi.org/10.1007/s10887-022-09217-1>.
- Bernanke, Ben and Mark Gertler, "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review*, 1989, 79 (1), 14–31.
- Bernanke, Ben S., "Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression," *American Economic Review*, 1983, 73 (3), 257–276.
- , Mark Gertler, and Simon Gilchrist, "The Financial Accelerator in a Quantitative Business Cycle Framework," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, Vol. 1, Elsevier, 1999, chapter 21, pp. 1341–1393.
- Bloom, Nicholas, "The Impact of Uncertainty Shocks," *Econometrica*, 2009, 77 (3), 623–685.
- Boot, Arnoud, Peter Hoffmann, Luc Laeven, and Lev Ratnovski, "Fintech: What's Old, What's New?," *Journal of Financial Stability*, 2021, 53, 100836.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Diagnostic Expectations and Credit Cycles," *Journal of Finance*, 2018, 73, 199–227.
- , —, and —, "Overreaction and Diagnostic Expectations in Macroeconomics," *Journal of Economic Perspectives*, 2022, 36 (3), 223–244.
- , —, —, and Stephen J. Terry, "Real Credit Cycles," 2021. NBER Working Paper No. 28416.
- , —, Yueran Ma, and Andrei Shleifer, "Overreaction in Macroeconomic Expectations," *American Economic Review*, 2020, 110 (9), 2748–2782.
- Brunnermeier, Markus K. and Yuliy Sannikov, "A Macroeconomic Model with a Financial Sector," *American Economic Review*, 2014, 104 (2), 379–421.
- and —, "Macro, Money and Finance: A Continuous-Time Approach," 2017. Handbook of Macroeconomics.
- Brunnermeier, Markus, Thomas Eisenbach, and Yuliy Sannikov, "Macroeconomics with financial frictions: A survey," 2013. Advances in Economics and Econometrics.
- Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin, "Finance and Development: A Tale of Two Sectors," *American Economic Review*, 2011, 101 (5), 1964–2002.

- Carlstrom, Charles T. and Timothy S. Fuerst, "Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis," *American Economic Review*, 1997, 87 (5), 893–910.
- Caselli, Francesco and Nicola Gennaioli, "Dynastic Management," *Economic Inquiry*, 2013, 51 (1), 971–996.
- Cornaggia, Jess, Brian Wolfe, and Woongsun Yoo, "Crowding Out Banks: Credit Substitution by Peer-To-Peer Lending," 2019. Unpublished working paper.
- Cornelli, Giulio, Jon Frost, Leonardo Gambacorta, Raghavendra Rau, Robert Wardrop, and Tania Ziegler, "Fintech and Big Tech Credit: Drivers of the Growth of Digital Lending," *Journal of Banking and Finance*, 2023, 148, 106742.
- Drechsel, Thomas, "Earnings-Based Borrowing Constraints and Macroeconomic Fluctuations," *American Economic Journal: Macroeconomics*, 2023, 15, 1–34.
- Fernandez-Villaverde, Jesus, Samuel Hurtado, and Galo Nuno, "Financial Frictions and the Wealth Distribution," 2019. NBER Working Paper No. 26302.
- Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery, "The Role of Technology in Mortgage Lending," *Review of Financial Studies*, 2019, 32, 1854–1899.
- Gambacorta, Leonardo, Yiping Huang, Han Qiu, and Jingyi Wang, "How do Machine Learning and Non-traditional Data Affect Credit Scoring? New Evidence from a Chinese Fintech Firm," 2019. BIS Working Paper No. 834.
- , —, Zhenhua Li, Han Qiu, and Shu Chen, "Data versus Collateral," *Review of Finance*, 2023, 27 (2), 369–398.
- Greenwald, Daniel, "Firm Debt Covenants and the Macroeconomy: The Interest Coverage Channel," 2019. MIT Sloan Research Paper No. 5909-19.
- Hart, Oliver and John Moore, "A Theory of Debt Based on the Inalienability of Human Capital," *Quarterly Journal of Economics*, 1994, 109 (4), 841–879.
- Hartman, Richard, "The effects of price and cost uncertainty on investment," *Journal of Economic Theory*, 1972, 5 (2), 258–266.
- Haskel, Jonathan and Stian Westlake, *Capitalism without Capital: The Rise of the Intangible Economy*, Princeton University Press, 2018.
- Hau, Harald, Yi Huang, Hongzhe Shan, and Zixia Sheng, "FinTech Credit and Entrepreneurial Growth," 2018. Unpublished Working Paper.
- He, Zhiguo and Arvind Krishnamurthy, "Intermediary Asset Pricing," *American Economic Review*, 2013, 103 (2), 732–770.
- Holmstrom, Bengt and Jean Tirole, "Financial Intermediation, Loanable Funds, and The Real Sector," *Quarterly Journal of Economics*, 1997, 112 (3), 663–691.
- Huang, Jing, "Fintech Expansion," 2022. Unpublished working paper.
- Huang, Yiping, Longmei Zhang, Zhenhua Li, Han Qiu, Tao Sun, and Xue Wang, "Fintech Credit Risk Assessment for SMEs: Evidence from China," 2020. IMF Working Paper No. 20-193.

- Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante, "Monetary Policy According to HANK," *American Economic Review*, 2018, 108 (3), 697–743.
- Kiyotaki, Nobuhiro, "Credit and Business Cycles," *Japanese Economic Review*, 1998, 49, 18–35.
- and John Moore, "Credit Cycles," *Journal of Political Economy*, 1997, 105 (2), 211–248.
- Krusell, Per and Anthony A. Smith, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 1998, 106 (5), 867–896.
- L’Huillier, Jean-Paul, Sanjay R. Singh, and Donghoon Yoo, "Incorporating Diagnostic Expectations into the New Keynesian Framework," 2023. Unpublished working paper.
- Li, Huiyu, "Leverage and Productivity," *Journal of Development Economics*, 2022, 154, 102752.
- Lian, Chen and Yueran Ma, "Anatomy of Corporate Borrowing Constraints," *Quarterly Journal of Economics*, 2021, 136, 229–291.
- Liu, Lei, Guangli Lu, and Wei Xiong, "The Big Tech Lending Model," 2022. NBER Working Paper No. 30160.
- Manea, Cristina, Fiorella De Fiore, and Leonardo Gambarcorta, "Big Techs and the Credit Channel of Monetary Policy Transmission," 2023. BIS Working Papers No. 1088.
- Minsky, Hyman P., "The Financial Instability Hypothesis," 1992. The Jerome Levy Economics Institute Working Paper No. 74.
- Moll, Benjamin, "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?," *American Economic Review*, 2014, 104 (10), 3186–3221.
- Oi, Walter Y., "The Desirability of Price Instability Under Perfect Competition," *Econometrica*, 1961, 29 (1), 58–64.
- Philippon, Thomas, "The FinTech Opportunity," 2016. NBER Working Paper 22476.
- Raissi, M., P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, 2019, 378 (1), 686–707.
- Shiller, Robert, *Irrational Exuberance*, Princeton University Press, 2016.
- Stiglitz, Joseph E. and Andrew Weiss, "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 1981, 71 (3), 393–410.
- Tang, Huan, "Peer-to-Peer Lenders versus Banks: Substitutes or Complements?," *Review of Financial Studies*, 2019, 32 (5), 1900–1938.
- Tella, Sebastian Di, "Uncertainty Shocks and Balance Sheet Recessions," *Journal of Political Economy*, 2017, 125 (6), 2038–2081.
- Thakor, Anjan V., "Fintech and Banking: What do We Know?," *Journal of Financial Intermediation*, 2020, 41, 1–13.
- Thakor, Richard and Robert Merton, "Trust in Lending," 2019. MIT Sloan Research Paper No. 5524-18.

APPENDIX

A. Proof

Proof of Lemma 1 and 2

To begin with, it can be easily shown that lenders will not verify if an entrepreneur announces the good state as entrepreneurs have no incentives to announce the good state when the actual is a bad one.

There are several conditions on this optimal contract. First, the lenders must be break-even. As shown in Equation (3), the funding cost for investors is $(1 + r)b$, and the expected return should be a weighted average of his payoff in the good state and his payoff in the bad state. Since there is no verification in the good state, the investor's payoff is $z_G k - c_G$. However, in the bad state, there is a probability of q that investors will verify entrepreneurs' earnings, and the cost of verification is assumed to be f . Therefore, his expected payoff in the bad state is $z_B k - q(c_B + f) - (1 - q)\tilde{c}_B$.

The second condition that an optimal contract needs to satisfy is that the entrepreneurs have no incentives to lie about the realized outcomes. Of course, entrepreneurs do not have any incentives to lie when it is a bad state. Entrepreneur's consumption in the good state is always c_G . If he lies about the state and says it is a bad state. Then his expected consumption should be $(1 - q)[(z_G - z_B)k + c_B]$.

The third condition that an optimal contract needs to meet is that the executions of contracts are feasible, which means all of c_G , c_B , and \tilde{c}_B should be at least higher than or equal to zero. More importantly, p is a probability so it must lie between 0 and 1.

In this incomplete-collateralization case, the optimal verification probability is

$$q = \frac{(1 + r)b - z_B k}{p(z_G - z_B)k - (1 - p)f} \quad (\text{A1})$$

when f is too large, q will be higher than 1 or even negative, which makes this earnings-based borrowing constraint infeasible. Therefore, in this paper, the micro-foundation for

these two types of borrowing constraints is from the different earnings verification costs for banks and BigTech firms. As a result, they choose different ways of lending contracts, and BigTech leads to specialization-induced fragmentation in the financial services industry. In the following part of the paper, we will take these two types of borrowings as given, and investigate their macroeconomic implications.

Therefore, the expected profits of using these two types of lending can be shown as follows:

$$\pi = \begin{cases} (1+r)[pz_G k + (1-p)z_B k - (1-p)f] & \text{if cash flow-based lending} \\ (1+r)lk & \text{if asset-based lending} \end{cases} \quad (\text{A2})$$

Therefore, if $f < f^* = \frac{pz_G + (1-p)z_B - l}{1-p}k$ or $l < l^* = pz_G + (1-p)z_B - (1-p)\frac{f}{k}$, then the lenders will strictly prefer cash flow-based lending over asset-based lending.

Proof of Lemma 3

Based on expected productivity $\tilde{\mathbb{E}}[z]$, the first order conditions on the optimal capital-to-labor ratio satisfies the following condition

$$\frac{l}{k} = \frac{(r+\delta)(1-\alpha)}{\alpha w} \quad (\text{A3})$$

Therefore, the firm's expected profits can be written as

$$\tilde{\mathbb{E}}[\pi] = \left(\frac{(1-\alpha)(r+\delta)}{\alpha w} \right)^{1-\alpha} \tilde{\mathbb{E}}[z] - \frac{r+\delta}{\alpha} \quad (\text{A4})$$

As the profit is a linear function of $\tilde{\mathbb{E}}[z]$, the firm's optimal choice of capital stock in the banking sector is a corner solution and meets the following condition

$$k^B(a, z, \tilde{z}) = \begin{cases} \frac{a}{1-\lambda_B} & \tilde{\mathbb{E}}[z] \geq \tilde{z} \\ 0 & \tilde{\mathbb{E}}[z] < \tilde{z} \end{cases}$$

where $\tilde{z} = \left(\frac{r+\delta}{\alpha}\right)^\alpha \left(\frac{w}{1-\alpha}\right)^{1-\alpha}$.

At the same time, the firm's realized profits can be written as

$$\pi = \left(\frac{(1-\alpha)(r+\delta)}{\alpha w}\right)^{1-\alpha} z - \frac{r+\delta}{\alpha} \quad (\text{A5})$$

With the optimal capital holding equation shown as above, the wealth growth for entrepreneurs in the banking sector as

$$da_B = \left\{ 1_{\tilde{\mathbb{E}}[z] \geq \tilde{z}} \times \left[\frac{\zeta z - \frac{r+\delta}{\alpha}}{1-\lambda_B} + r - \rho \right] + 1_{\tilde{\mathbb{E}}[z] < \tilde{z}} \times (r - \rho) \right\} a_B dt \equiv \Gamma^B(z, \tilde{z}) a_B dt \quad (\text{A6})$$

Similarly, for BigTech entrepreneurs, we can derive the entrepreneur's optimal choice on capital stock as follows:

$$k^F(a, z, \tilde{z}) = \begin{cases} \frac{1}{1+\lambda_F \left(\frac{r+\delta}{\alpha} - \zeta \tilde{\mathbb{E}}[z]\right)} a & \tilde{\mathbb{E}}[z] \geq \tilde{z} \\ 0 & \tilde{\mathbb{E}}[z] < \tilde{z} \end{cases} \quad (\text{A7})$$

The only difference is the borrowing constraint for highly productive entrepreneurs. Hence, the wealth growth rate for entrepreneurs in the BigTech sector can be shown below:

$$da_F = \left\{ 1_{\tilde{\mathbb{E}}[z] \geq \tilde{z}} \times \left[\frac{\zeta z - \frac{r+\delta}{\alpha}}{1+\lambda_F \left(\frac{r+\delta}{\alpha} - \zeta \tilde{\mathbb{E}}[z]\right)} + r - \rho \right] + 1_{\tilde{\mathbb{E}}[z] < \tilde{z}} \times (r - \rho) \right\} a_F dt \quad (\text{A8})$$

As $\tilde{\mathbb{E}}[z] = \tilde{\mathbb{E}}[\tilde{z} + dz] = \frac{1}{\theta} [\bar{\mu} + (\theta - 1)\tilde{z}] + \gamma(\tilde{z} - \bar{\mu})$, we can rewrite it as follows:

$$da_{\mathcal{F}} = \left\{ \mathbb{1}_{\tilde{\mathbb{E}}[z] \geq \tilde{z}} \times \left[\frac{\zeta z - \frac{r+\delta}{\alpha}}{1 + \lambda_{\mathcal{F}} \left[\frac{r+\delta}{\alpha} - \zeta \left(\frac{1}{\theta} - \gamma \right) \bar{\mu} - \frac{\zeta(\theta-1+\theta\gamma)}{\theta} \tilde{z} \right]} + r - \rho \right] + \mathbb{1}_{\tilde{\mathbb{E}}[z] < \tilde{z}} \times (r - \rho) \right\} a_{\mathcal{F}} dt \quad (\text{A9})$$

Proof of Lemma 4

As shown in the previous lemma, the wealth follows a process of $da_j = [\Gamma_j(t, z, \tilde{z}) a_j - c_j] dt$ for the entrepreneurs in sector j . Therefore, the Bellman equation \mathcal{V}_j should satisfy the following equation

$$\rho \mathcal{V}_j(t, a, z, \tilde{z}) = \max_{c_j} \left\{ \log c_j + \frac{1}{dt} E [d\mathcal{V}_j(t, a, z, \tilde{z})] \right\} \quad (\text{A10})$$

subject to the condition that $da_j = [\Gamma_j(t, z, \tilde{z}) a_j - c_j] dt$.

With the guess and verify approach, we can show that the optimal consumption choice is $c_j = \rho a_j$ for all entrepreneurs in the economy. Assume that the value function takes the form of $\mathcal{V}_j(t, a, z, \tilde{z}) = \mathcal{B}_j v_j(t, z, \tilde{z}) + \mathcal{B}_j \log a_j$. Then we have

$$E [d\mathcal{V}_j(t, a, z, \tilde{z})] = \frac{\mathcal{B}_j}{a_j} da + \mathcal{B}_j E [dv_j(t, z, \tilde{z})] \quad (\text{A11})$$

Combining Equations (A10) and (A11) gives us the following equation:

$$\rho \mathcal{B}_j v_j(t, z, \tilde{z}) + \rho \mathcal{B}_j \log a_j = \max_{c_j} \left\{ \log c_j + \frac{\mathcal{B}_j}{a_j} [\Gamma_j(t, z, \tilde{z}) a_j - c_j] + \mathcal{B}_j \frac{1}{dt} E [dv(t, z, \tilde{z})] \right\} \quad (\text{A12})$$

The first-order condition gives us $c_j = \frac{a_j}{\mathcal{B}_j}$. Substituting back in, we have

$$\rho \mathcal{B}_j v_j(t, z, \tilde{z}) + \rho \mathcal{B}_j \log a_j = \log a_j - \log \mathcal{B}_j + \mathcal{B}_j \Gamma_j(t, z, \tilde{z}) - 1 + \mathcal{B}_j \frac{1}{dt} E [dv_j(t, z, \tilde{z})]$$

which is

$$(\rho \mathcal{B}_j - 1) \log a_j = -\rho \mathcal{B}_j v_j(t, z, \tilde{z}) - \log \mathcal{B}_j + \mathcal{B}_j \Gamma_j(t, z, \tilde{z}) - 1 + \mathcal{B}_j \frac{1}{dt} E [dv_j(t, z, \tilde{z})] \quad (\text{A13})$$

Therefore, we can conclude that $\mathcal{B}_j = \frac{1}{\rho}$ for both sectors, and we have

$$c_j = \rho a_j \quad (\text{A14})$$

$$da_j = [\Gamma_j(z, \tilde{z}) a_j - \rho] dt \quad (\text{A15})$$

Finally, the value function is

$$\mathcal{V}_j(t, a, z, \tilde{z}) = \frac{1}{\rho} [v_j(t, z, \tilde{z}) + \log a_j] \quad (\text{A16})$$

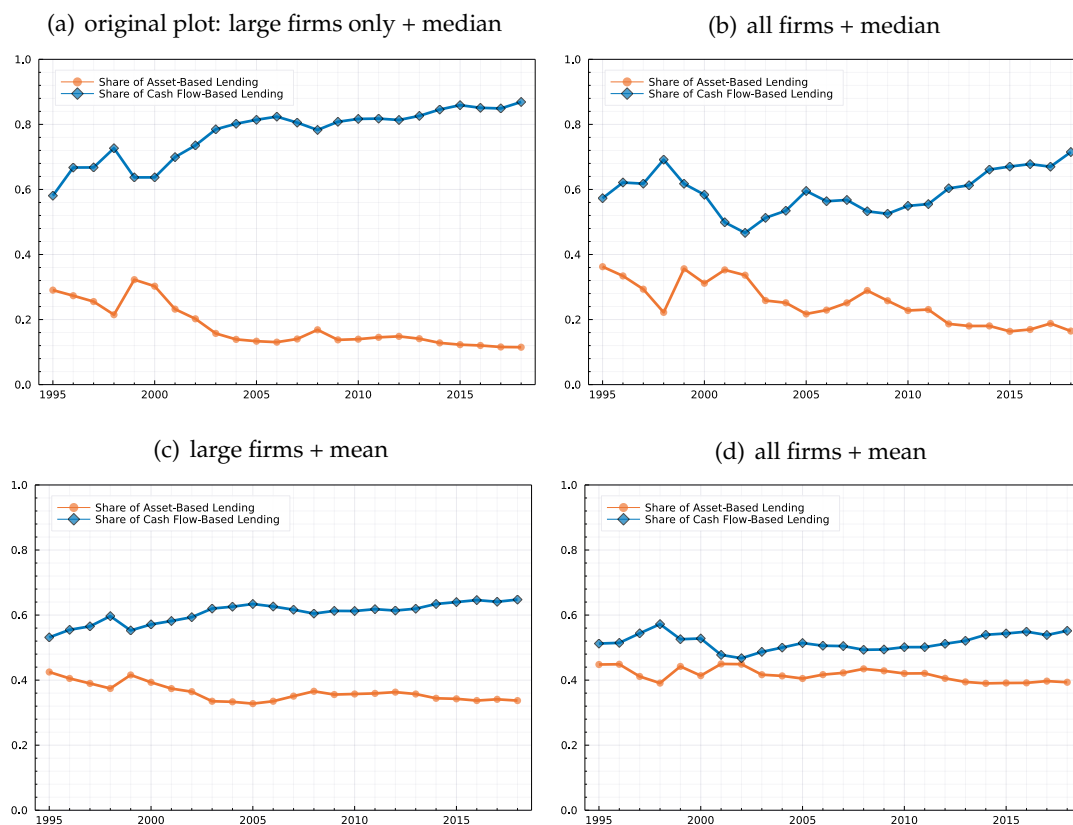
and $v_j(t, z)$ satisfies the following condition:

$$\rho v_j(t, z, \tilde{z}) = \rho \log \rho + \Gamma_j(t, z, \tilde{z}) - \rho + \mathcal{B}_j \frac{1}{dt} E [dv_j(t, z, \tilde{z})] \quad (\text{A17})$$

B. A Note on Lian and Ma (2021)

Lian and Ma (2021) document the prevalence of cash flow-based lending. They argue that 20% of debt by value is based on tangible assets, whereas 80% is based predominantly on cash flows from corporate operations. Here we want to argue that their main conclusion, especially the dominating use of cash flow-based lending, is not robust. A better and less controversial way of interpreting the empirical result is the co-existence of earnings-based and collateral-based borrowing constraints.

Figure A1: Anatomy of Corporate Borrowing Constraints



Notes: This figure presents the anatomy of corporate borrowing constraints with different choices for summarizing the data. The main data source for this figure is obtained directly from the replication package for Lian and Ma (2021).

Graph (A) in Figure A1 replicates one of their main results in the paper. To be clear, all the raw data used in this section are directly obtained from Quarterly Journal of Economics Dataverse.¹ In their paper, Lian and Ma (2021) argue that “Figure I, Panel A, shows that the median share of asset-based and cash flow–based lending among large nonfinancial firms is generally less than 20% and slightly over 80%, respectively, in recent years.” The keywords in their original statement are **large** and **median**. More specifically, when they prepare the data for this graph, first they classify all the firms in Compustat dataset into five different groups according to their total asset levels. Then they drop the bottom 20% firms out of the sample. Finally, they compute the *median* share of asset-based and cash flow–based lending. As we can see from the replicated result in Graph (A), the median share of asset-based lending on average is 17.8%, while that of cash flow–based lending is 77.2%.

To begin with, we want to point out that these numbers are sensitive to the choice of sub-samples and the use of a median. In Graph (B) of Figure A1, we plot the same results but without dropping the smallest firms out of the sample. In Graph (C) of Figure A1, we drop all the firms in the lowest quintile but use mean instead of median. In Graph (D), we include all the firms and use the mean to calculate the average value. As we can see from these graphs, whether cash flow-based lending is really prevalent depends on the specific choice of our empirical measure. For example, in Graph (D), the average use of cash flow-based lending is 51.7% while the average use of asset-based lending is 41.6%. In this way, both types of lending are important financial frictions in the real economy.

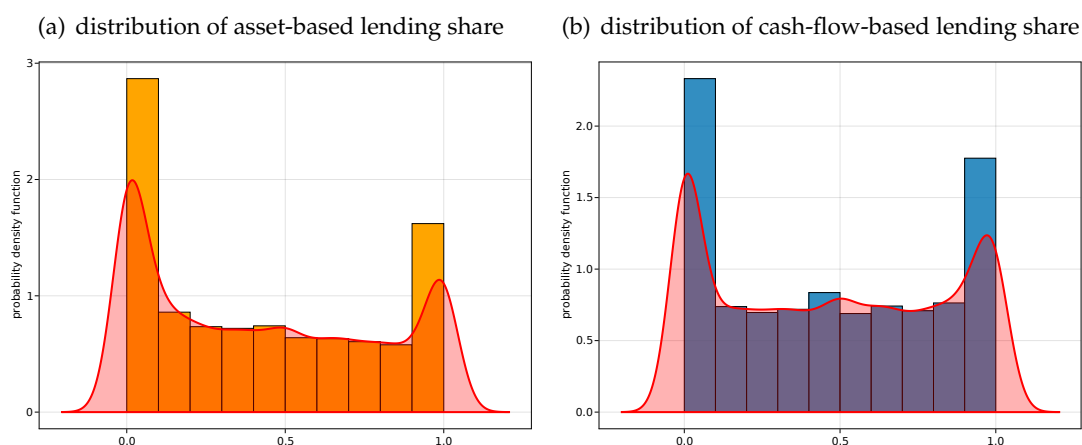
The subsample selection is not the most problematic issue in their work. In fact, Lian and Ma (2021) do mention this point. They find that for small firms, asset-based lending is more common and the median value of asset-based lending among these small firms is roughly 54%.

The real problem comes from the use of the median because the actual distribution of the borrowing constraints is a **bimodal** one. It is true that both the median and mean can

¹Replication data and codes for Lian and Ma (2021) can be downloaded from [here](#).

be interpreted as the “representative” value for the data sample, and sometimes the median is used as an alternative to the mean. However, if the underlying distribution is a bimodal one, both indicators can be misleading, as there is no such representative borrower in the data. Graphs (A) and (B) in Figure A2 present the distribution of individual firms’ use of asset-based and cash flow-based lending, respectively. As we can see from these two graphs, when we attempt to describe the use of borrowing constraints by the individual firm, there is no such representative firm in this economy because some firms rely heavily on cash-flow-based lending while other firms use more collateral-based lending. The detailed breakdown for each year throughout the data sample period can be shown in Figure A3 and A4. Generally speaking, the less controversial way of describing reality is the co-existence of two types of borrowing constraints.

Figure A2: Distributions on the types of borrowing constraints



Notes: This figure presents the distributions of individual firms’ use of two types of lending. The main data source for this figure is obtained directly from the replication package for Lian and Ma (2021). Orange and blue rectangles represent histogram distributions with normalized probability density. Red lines are the Kernel smoothing function fits.

Figure A3: Asset-Based Lending Distribution in each year

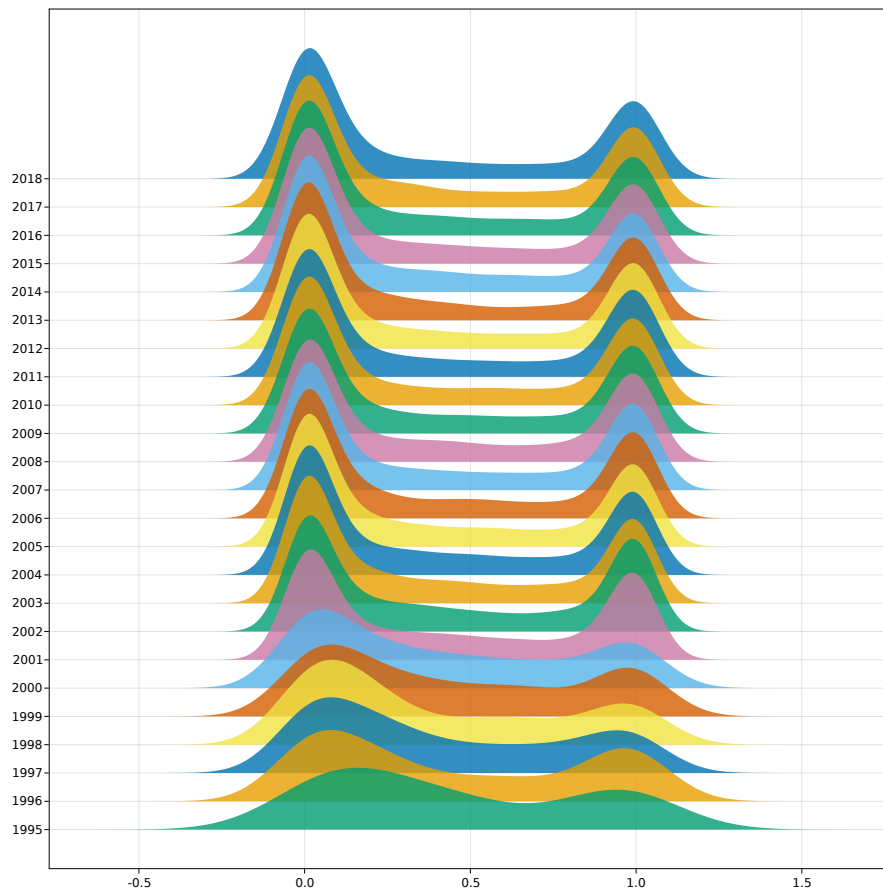
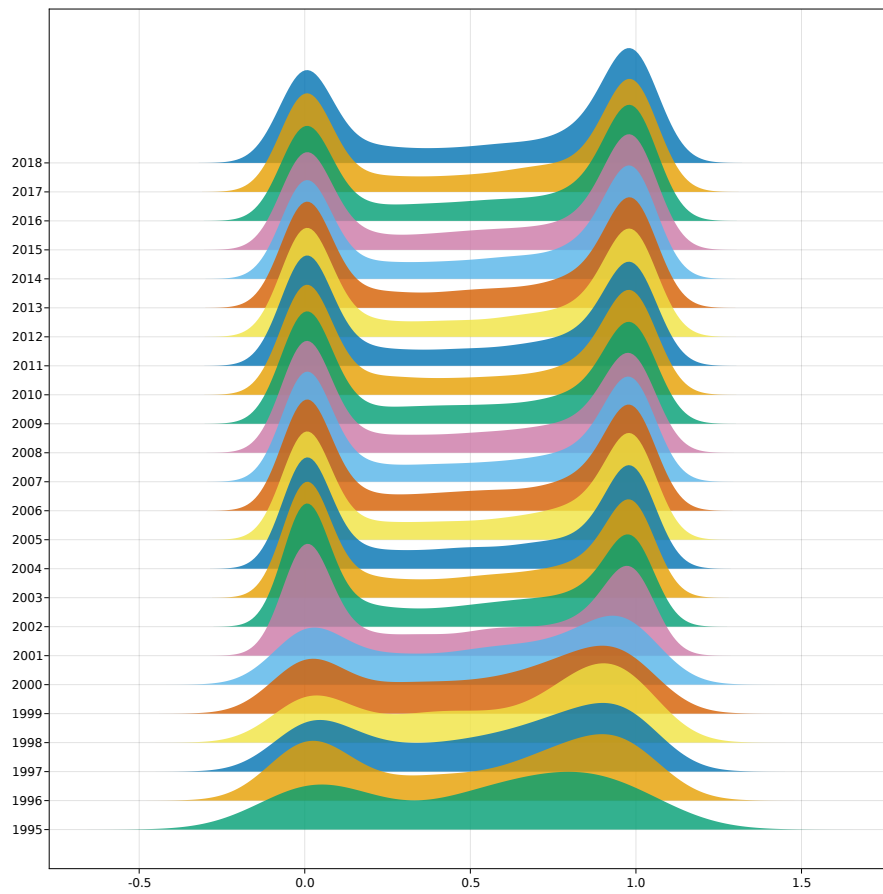


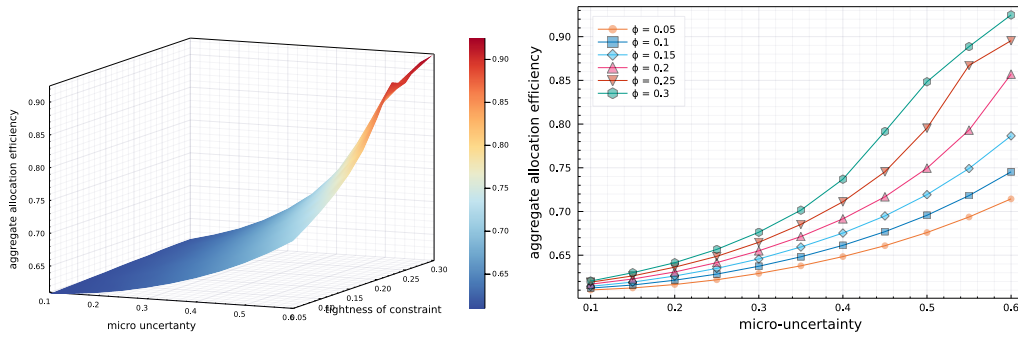
Figure A4: Cash Flow-Based Lending Distribution in each year



C. Additional Figures

Figure A5: Determinants of aggregate allocation efficiency

(a) micro-uncertainty and cash flow-based borrowing constraints



(b) micro-uncertainty and autocorrelation

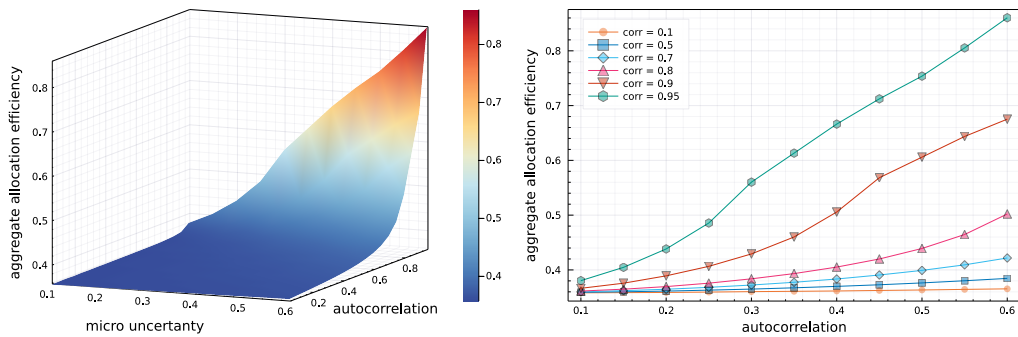
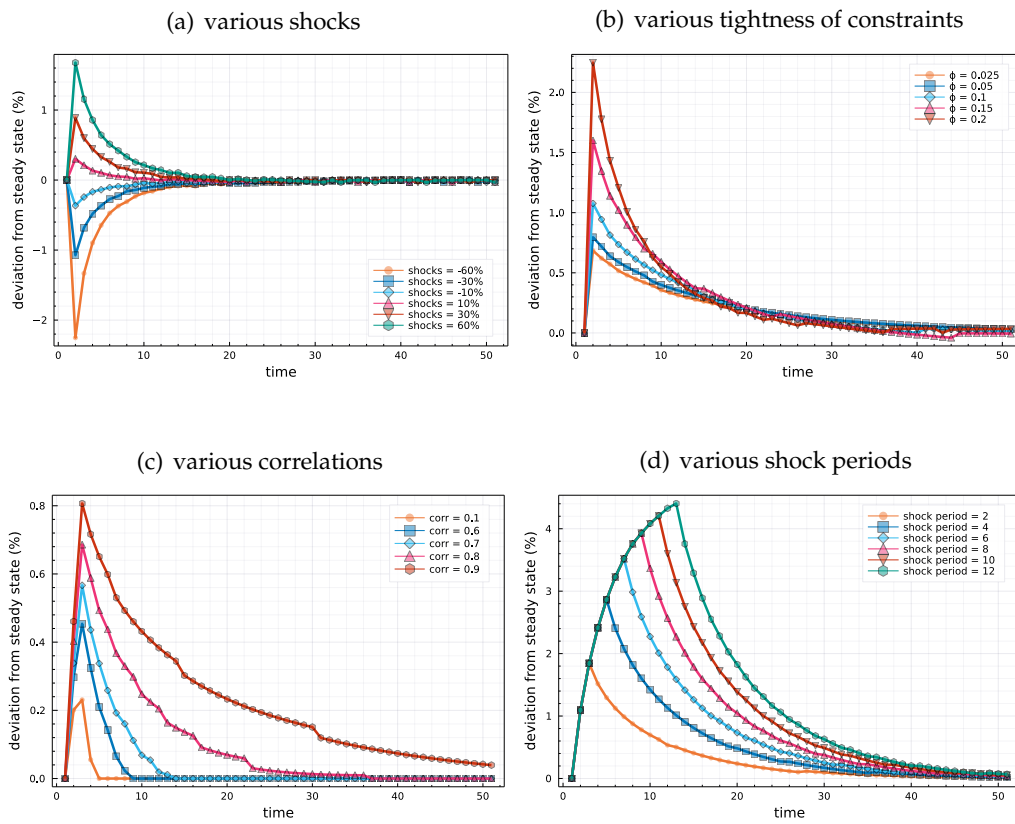
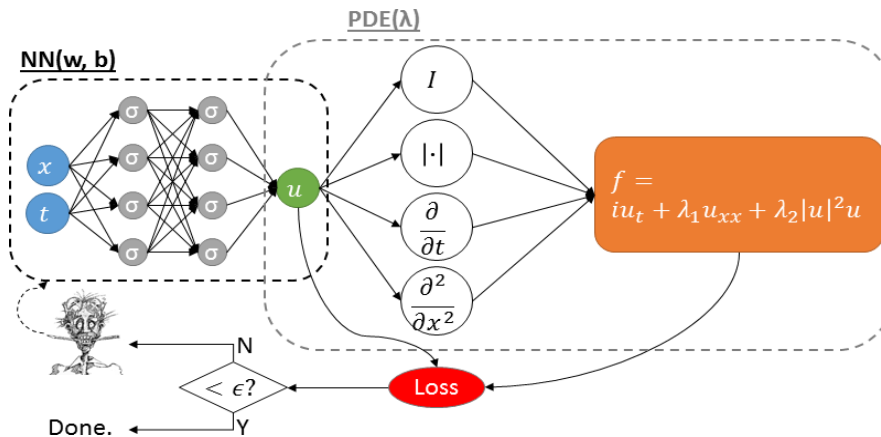


Figure A6: Determinants of business cycle fluctuations



D. Physics-Informed Neural Networks Algorithm

The physics-Informed Neural Networks (PINN) algorithm is proposed by Raissi et al. (2019) and represents a new deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. The basic idea of this algorithm can be summarized in the following graph:



Generally speaking, the idea of PINN is to employ two or more neural networks that share the same parameters. In addition, the objection function is to minimize the sum of mean squared errors of the original neural network and those of partial derivatives. In this way, we can make full use of the synergy between machine learning and classical computational physics to solve some high-dimensional partial differential equations without encountering the curse of dimensionality. More importantly, this approach is feasible because the PINN approximation theorem guarantees that feed-forward neural nets with a sufficiently large enough number of neural nodes can simultaneously and uniquely approximate any partial differential equations and their partial derivatives.