Bank Lending and Firm Dynamics in General Equilibrium*

Sabrina Studer† Yingnan Zhao‡

September 5, 2023

Abstract

This paper models a dynamic lending relationship between banks and firms under asymmetric information in general equilibrium with endogenous firm entry. Banks and firms sign long-term lending contracts that specify the optimal levels of bank loans and repayments in each period depending on a firm's productivity history. In equilibrium, interest rates, wage rates, and firm dynamics, such as growth rates and growth volatility over a firm's lifespan are determined. Further, we extend the model to investigate the long-run effects of an exogenous credit contraction on firms. In a calibrated economy, we find that most of the negative effects are mitigated by an endogenous increase in the size of banks' balance sheets. Nevertheless, firm entry contributes disproportionally to the effect on aggregate variables: More than half of the contraction in capital employment and almost all the decline in aggregate output and labor employment come from a decline in the firm entry. Besides, younger firms, which are also smaller on average, are much more adversely affected than older ones. The growth rate of young firms remains high though, as they operate at smaller sizes and banks increase the provision of loans more strongly when firms grow older.

Keywords: Dynamic contracts, asymmetric information, firm dynamics, general equilibrium

JEL: D52, D82, E44, G21

* We are deeply indebted to Josef Falkinger and Fabrizio Zilibotti for their invaluable guidance and patience. We thank Martin Brown, John H. Moore, Asim Khwaja, Kjetil Storesletten, Christoph Winter, Miquel Faig, Gueorgui Kambourov, seminar participants at the University of Toronto, McMaster University, University of Queensland, Shanghai University of Finance and Economics, and conference participants at 2017 Macro Financial Modeling (MFM) Summer Session, 2017 Asian Meeting of Econometric Society, 2017 EEA-ESEM, 2016 Royal Economic Society Conference, 20th Workshop on Dynamic Macroeconomics for insightful comments and suggestions.

† Department of Economics, University of Zurich, sabrinastuder@sunrise.ch.

‡Corresponding author, Department of Economics, University of Toronto, 3359 Mississauga Road, Mississauga, ON, L5L 1C6, Canada, Tel: +1 905-828-3910, yingnan.zhao@utoronto.ca.

1 Introduction

Despite the rapid evolution of financial markets in recent decades, bank credit remains one of the most important external financing sources for firms across many countries (Berger and Udell, 1998; Hackethal and Schmidt, 2004; Aldasoro and Unger, 2017). It determines a firm's market entry decisions, its size, and its dynamics (e.g., growth rates and growth volatility) throughout the lifespan. Meanwhile, the allocation of credit among firms is constrained by banks' aggregate credit availability. Recent evidence suggests that disruptions in bank credit supply often generate uneven effects on firms depending on their age, size, and history with banks, etc. (Gertler and Gilchrist, 1993; Khwaja and Mian, 2008; Duygan-Bump et al., 2015). This paper studies the effects of a contraction in banks' aggregate credit supply on firms and contributes to the literature in three aspects: First, we endogenize, in a unified framework, the effects on both the allocation of credit among firms (the "horizontal" dimension) and the performance of individual firms over their lifespan (the "vertical" dimension). Second, we investigate these effects in the long run, which can be empirically challenging but are important in the discussion of a change in bank regulation or in banks' preference of hoarding liquidity, etc. Third, we allow the aggregate credit supply to be determined endogenously. This channel proves crucial in determining the magnitude of the effects in the long run, as a significant fraction of the contraction in credit availability is compensated by an endogenous increase in the size of the bank balance sheet.

To do this, we construct an infinite-horizon general equilibrium model where firms make endogenous market entry decisions and banks decide the optimal levels of lending to firms over their lifetime via long-term lending contracts. The allocation of credit among firms adds up to the aggregate supply of bank credit, the latter of which is determined by workers' optimal savings in banks and an exogenous policy parameter, the bank reserve ratio. The factor prices – interest rates and wage rates – are pinned down in equilibrium and in turn, affect the optimal decisions and the allocation of credit among firms.

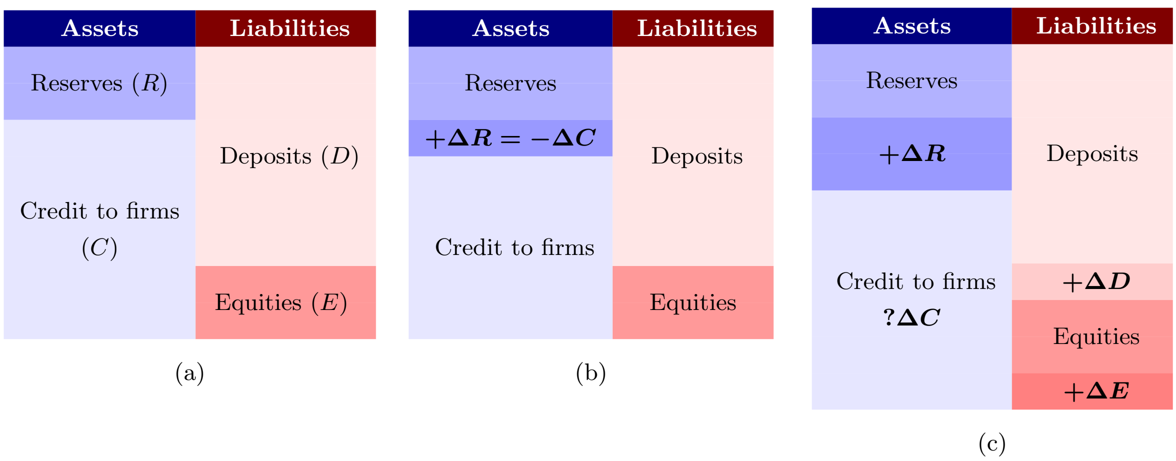
The long-term lending relationship is motivated by the existence of idiosyncratic productivity shocks on firm production and asymmetric information between banks and firms regarding the realization of the shocks. In the model, firms rely on bank loans to finance their production in each period and have limited liability for their debts. To prevent firms from misreporting their true levels of output, banks offer them long-term lending contracts, which specify the optimal levels of bank loans and required repayments in each period according to the entire historical performance of firms. We follow the literature on dynamic contracts to formulate the optimal lending contracts recursively using promised values to firms (i.e., entrepreneurs' expected utility from future cash flows) as a state variable.[1] The variable summarizes all information about a

[1]The recursive formulation of dynamic contracts under asymmetric information was originally developed in

firm's history of productivity realizations, while the contract terms (i.e., the levels of loans and repayments) are functions of the variable. In line with the literature, we show that endogenous borrowing constraints emerge as a characteristic of optimal lending contracts. Moreover, despite the simple assumptions on firm shocks, the model outcome matches most of the empirical regularities on firm dynamics qualitatively. This includes a positive correlation between firm size and firm age, a negative correlation between firm age and firm growth, and firm age and the volatility of firm growth.

Once establish the bank-firm lending relationship at the micro-level, we embed the dynamic contracts in a general equilibrium framework. We model firm entry by the optimal occupational choice of newborn households to become entrepreneurs. Specifically, households decide to become a worker or an entrepreneur immediately after birth by comparing the lifetime utility of the two occupations. Workers supply labor in firms, consume, and deposit savings in banks. Entrepreneurs employ labor and capital to produce output and consume the profits from production. In equilibrium, households are indifferent between the two occupations and the share of entrepreneurs is determined by the labor-market-clearing condition.

Figure 1: Bank balance sheets and general-equilibrium effects of a credit contraction



Note: Subplot (a) represents the initial composition of a bank's balance sheet. In partial equilibrium where the size of the balance sheet is fixed, an increase, $+\Delta R$, in the share of reserves leads to a one-to-one decrease, $-\Delta C$, in the credit supply to firms as illustrated in subplot (b). In general equilibrium, however, the size of deposits and bank equities depends on the equilibrium factor prices, which are endogenous. Hence, the effect on the credit supply to firms can be ambiguous (subplot (c)).

In the model, banks are the financial intermediaries in the economy that take deposits from workers and provide loans to entrepreneurs via optimal lending contracts. Banks hold an exogenous share (i.e., reserve ratio) of the deposits they receive from workers as reserves. As is illustrated in Subplot (a) of Figure 1, the aggregate bank credit to firms is determined by the aggregate deposits from workers, equities accumulated from bank operations, and the reserve

Green (1987), Thomas and Worrall (1990), and Atkeson and Lucas (1992).

ratio of banks. This also constitutes one of the conditions that determine the interest rates and wage rates in equilibrium. A credit contraction is modeled as an exogenous increase in the reserve ratio and our analysis focuses on its long-run effects. This means that we compare the stationary equilibrium of two scenarios where banks maintain different reserve ratios. The equilibrium factor prices and the optimal lending contracts are affected through multiple channels.

First, a credit contraction decreases capital supply to firms and drives up the capital cost - the interest rates - in equilibrium. A higher interest rate increases the costs of production for firms and lowers their profits at a given level of loans. This, in turn, decreases bank profits from the optimal lending contracts for three reasons: First, banks cannot require higher repayments than what a firm earns under the limited-liability constraint. Therefore, as a firm earns less, repayments decrease and so do bank profits. Second, lower firm profits decrease the consumption of entrepreneurs. This implies a higher marginal utility and thus gives the risk-averse entrepreneurs a stronger incentive to report a worse productivity realization to be eligible for a lower debt payment. and increase the marginal. This again lowers banks' profits. Lastly, risk-neutral banks discount future cash flows using the interest rate. As the interest rate increases, bank profits, which equal the present value of the cash flows decrease.

These channels generate important implications for the allocation of credit among firms and firm dynamics in the long run. As bank profits decrease, the zero-profit condition implies that banks lower the initial promised values to firms in equilibrium. According to the optimal contracts, this means that new entrants receive a lower level of loans and thus start at smaller sizes. However, due to the higher interest rate, banks also adjust the intertemporal allocation of the promised values by decreasing the level of loans today and increasing it in the future. Therefore, the *growth rate* of young firms is less affected because of a "base-year" effect (mathematically, a smaller denominator in the expression of the growth rate) and the fact that banks increase the level of loans more rapidly over time. The latter also suggests that firms catch up in size as they stay in a lending relationship for longer, despite a smaller size at entry. Hence, in the long run, the negative effect on the credit supply to old firms is quite limited. In fact, there may even be a reallocation of credit from young firms toward old ones as suggested by our calibrated economy. Lastly, the volatility of firm growth declines across all ages as the distribution of firm size concentrates within a narrower range.

To evaluate the model implications quantitatively, we calibrate the model to the U.S. economy, simulate the life path of entrepreneurs and workers with idiosyncratic shocks, and solve for the stationary equilibrium numerically. In the calibrated economy, a permanent increase in the reserve ratio from zero to 20% decreases average bank loans by 0.07% in the long run. In particular, average loans to firms below 10 years old, which account for 60% of all firms, decrease by 0.14%, twice as much as the average decline, while average loans to older firms (age above 20, accounting for 19% of all firms) even increase by 0.22%. In other words, the contraction of aggregate bank loans is reflected by a tighter financial constraint on younger firms. Empirical evidence shows that banks actively reallocate credit among borrowers after a credit contraction, which leads to substantial heterogeneity in the magnitude of credit constraints that firms experience (De Jonghe et al., 2020). The findings in our paper suggest that the results, despite at a much smaller magnitude, remain to be true in the long run. Moreover, financial supports that target small and young firms are not only more effective than an overall monetary or fiscal expansion over business cycles (Verani, 2018) but can also be beneficial in the longer term if a credit contraction is expected to linger.

The small magnitude of the long-run effect on the economy is mainly due to an endogenous increase in the size of banks' balance sheets. The idea is illustrated in Figure 1. In a partial equilibrium with fixed factor prices, the size of a bank's balance sheet is constant. Thus, an increase in the reserve ratio leads to a one-to-one decline in the credit supply to firms (subplot (b)). However, this is no longer the case in general equilibrium as the sizes of deposits and bank equities depend on the equilibrium factor prices (subplot (c)) and are thus endogenous. In the calibrated economy, the overall size of bank assets rises by 13% in the new stationary equilibrium as the reserve ratio increases from zero to 20%. In particular, the total deposits from workers and bank equities increase by 9.5% and 18.6%, respectively. As a result, the aggregate credit supply to firms decreases only by 0.3%.

Lastly, despite the small change in the credit supply to firms, a decomposition of the general equilibrium effect indicates that firm entry plays a significant role: More than half of the contraction in capital employment and almost all of the decline in aggregate output and labor employment come from a decline in the firm entry. Recent studies argue that the extensive margin, to a large extent, determines the magnitude of the economic consequences of negative shocks over business cycles (Clementi and Palazzo, 2016; Messer et al., 2016).[2] Our findings indicate that similar results emerge when looking at the long-run performance of the economy as well.

The model framework presented in this paper is potentially useful for studying the long-run effects of other regulations imposed on a bank's balance sheet as well. In Appendix G, we provide an example of how the model can be modified to analyze the effects of a tightening in capital requirements.

Our work builds on the theoretical literature that uses dynamic contracts to study bank-firm lending relationships in the presence of asymmetric information regarding firms' productivity

[2]Other papers that study how firm entry and exit amplify the effects of aggregate shocks in general equilibrium include Campbell (1998), Bilbiie et al. (2012), and Chatterjee and Cooper (2014), etc. But firm dynamics is not one of their focuses.

realizations.[3] Clementi and Hopenhayn (2006) investigate the effect of endogenous borrowing constraints on firm dynamics. However, the model is a partial equilibrium. Thus, the endogenous determination of the aggregate credit availability of banks and the allocation of credit among firms are not part of their discussion. Smith and Wang (2006) consider long-term lending contracts in general equilibrium similar to ours. Nevertheless, they assume that firms operate at a fixed scale and thus banks provide a constant amount of loans to all firms. In other words, the allocation of credit among firms is trivial and there are no firm dynamics.

The two studies that are most closely related to ours are Dyrda (2017) and Verani (2018). Both papers consider long-term lending contracts and firm dynamics in a general equilibrium framework. Dyrda analyzes the effects of a change in the volatility of productivity shocks on firm dynamics over the business cycle, and Verani investigates the implications of two types of financial frictions – private information and limited enforcement – for the credit supply to firms and the welfare of the economy. Both papers assume a small open economy with an exogenous interest rate. This is equivalent to allowing financial intermediaries unlimited access to external funding. Hence, their models do not accommodate the analysis of an aggregate credit supply contraction.

Our paper also contributes to the literature on lending relationships between banks and firms. Early theoretical work on the subject goes back to Sharpe (1990) and Rajan (1992), where banks acquire insider information about firms via lending relationships and use the information advantage to extract rents from firms. Their models suggest that as firms stay longer in a relationship, it becomes more costly for them to switch lenders. Hence, firms are "locked" in the relationships. These are in contrast to the role of lending relationships and the model implications in this paper. Specifically, we assume that banks use incentive-compatible lending contracts to overcome the information asymmetry and both banks and firms commit to the lending relationships. As a result, firms are incentivized to report truthfully about their output in equilibrium even though banks do not monitor or obtain information about firms. We show that firms get better terms of contracts (i.e., higher levels of bank loans and lower repayments) as they are in a lending relationship for longer. These features are consistent with findings of the empirical literature (Petersen and Rajan, 1994; Berger and Udell, 1995; Bharath et al., 2011). Moreover, our paper complements the literature by investigating how lending relationships are affected by a contraction of aggregate credit supply. There has been an emerging body of empirical studies on this topic recently (Iyer et al., 2014; Sette and Gobbi, 2015; Cohen et al., 2021), but there seems to be limited theoretical work so far.[4]

[3]The other strand of the dynamic-contracting literature assumes limited enforcement, meaning that lenders are unable to force borrowers to repay their debts. Albuquerque and Hopenhayn (2004) show that this friction may result in endogenous borrowing constraints on firms. Cooley et al. (2004) conclude that limited enforceability amplifies the macroeconomic impact of new technological innovations via its influence on credit to firms.

[4]Using data from Italy, Sette and Gobbi (2015) show that a longer lending relationship translates into higher credit growth and lower costs of credit. The positive effect of a relationship on credit supply increased during

Lastly, at a broader level, we contribute to a large body of literature that studies how financial frictions propagate and amplify macroeconomic fluctuations via the lending channel. The pioneering work in the area goes back to Bernanke and Gertler (1989), Kiyotaki and Moore (1997), Carlstrom and Fuerst (1997), and Bernanke et al. (1999).[5] But as was mentioned in Gertler (1992) and Clementi and Hopenhayn (2006), one general limitation of the literature is that the financial arrangements are not intertemporally optimal, meaning that contracts' provisions are not contingent on all public information. The optimal lending contracts in this paper are incentive-compatible and depend on the entire history of firm performance.

The remainder of the paper is organized as follows: Section 2 describes the model. In particular, we characterize the recursive formulation of the dynamic lending contracts and define the stationary general equilibrium. In Section 3 we calibrate the model and solve for the equilibrium numerically. Section 4 investigates the long-run effects of a credit contraction on the credit allocation among firms and firm dynamics. Section 5 concludes.

2 A Model of Firm Dynamics in General Equilibrium

2.1 Model Set-up

Consider a discrete-time model with an infinite time horizon, $t \in \{0, 1, \ldots\}$. At $t = 0$, a unit mass of *ex-ante* identical households is born. Households survive to the next period with an exogenous probability Δ. Those who do not survive are replaced by newborn households at the beginning of the next period. This implies a constant population and different cohorts, $\tau \in \{0, 1, ...\}$, at any time.[6]

Households are *ex-ante* identical and make irreversible occupational decisions of becoming a

the financial crisis in 2007-2009. Similarly, Iyer et al. (2014) find a stronger reduction in bank credit for firms with weaker lending relationships using Portuguese data during the same period. Based on an analysis of bank lending during the Great Depression, Cohen et al. (2021) show that loan rates in a lending relationship are less responsive to changes in bank funding costs. Recent theoretical contributions include Hachem (2011) and Bolton et al. (2016). They study how lending relationships with asymmetric information are affected by changes in monetary policy and vary over business cycles, respectively. Hachem has a learning model, where the multi-period lending relationships are essentially a sequence of one-period arrangements. Bolton et al. model a two-period lending relationship. In both scenarios, bank lending is a zero-one choice (i.e., lend or not) without deciding the levels of bank loans.

[5] Quadrini (2011) and Brunnermeier and Sannikov (2014) provide a comprehensive review of the literature.

[6] In the stationary equilibrium (defined in Section 2.4), the age distribution of the households is given by $\{(1-\Delta)\Delta^{\tau}\}_{\tau \in \{0,1,\ldots\}}$. Note that even though the firm entry and exit rates are different, the number of firms in the economy is constant in a stationary equilibrium. The entry rate, λ, is the share of *newborn* households that become an entrepreneur whereas the exit rate, $1-\Delta$, is the share of *incumbent* firms that exit the market. In other words, the denominators used to calculate the two variables are different by definition.

worker or an entrepreneur right after birth. Workers work in firms, earn wages, consume, and save. Entrepreneurs operate firms that employ capital and labor to produce one type of output, which can be used as consumption goods or capital. In equilibrium, households are indifferent between the two occupations. The share of entrepreneurs, λ, is determined endogenously by the labor-market-clearing condition, defined in Section 2.4.

In addition to households, there are risk-neutral financial intermediaries (referred to as "banks") that channel savings from workers to entrepreneurs as capital for production. The banking sector is competitive and banks make zero profits in equilibrium.

2.2 Households

Households are endowed with one unit of labor and no wealth. Their instantaneous utility function, $U(c, l)$, increases in consumption, $c \geq 0$, and decreases in labor supply, $l \in [0, 1]$.[7] The function is strictly concave and bounded in c. The future discount factor is β. We describe the lifetime problems of households with each occupation below.

2.2.1 Workers

Workers maximize lifetime utility by optimally deciding their consumption, c, labor supply, l, and savings, A'. Given the interest and wage rates, $\{r, w\}$, the problem can be formulated recursively using savings from the previous period, $A \geq 0$, as the state variable:

$$V^W(A; r, w) = \max_{c,l,A'} \left\{ U(c, l) + \Delta\beta V^W(A'; r', w') \right\}, \tag{1}$$

subject to

$$\begin{aligned} c + p^A A' &= wl + (1 + r)A, \\ c \geq 0,\ l &\in [0, 1],\ A' \geq 0. \end{aligned} \tag{2}$$

$V^W(A; r, w)$ is the value function that represents the maximal lifetime utility of a worker with savings A. Variables of tomorrow are denoted with an apostrophe. The future utility is discounted by both the time factor, β, and the probability, Δ, that a worker survives to the next period. Equation (2) defines the budget constraint. Namely, consumption, c, and savings for the future, $p^A A'$, come from workers' labor income, wl, and the gross return from previous savings, $(1 + r)A$. $p^A = \Delta$ is the actuarially fair price, which implies that workers only need to

[7] We assume that entrepreneurs supply a constant amount of labor. They do not earn wages but receive profits from production as described in Section 2.2.2. The same utility function for workers and entrepreneurs is needed because households make the optimal occupational decision by comparing the lifetime utility of a worker and of an entrepreneur. The values are only comparable under the same instantaneous utility function.

save $p^A A$ to receive a gross return of $(1+r)A$ because banks redistribute the savings of workers who did not survive to those who did.[8]

Denote the policy functions of the optimal savings A' and the labor choice l, respectively, by

$$A' = g(A; r, w), \quad l = h(A; r, w). \tag{3}$$

2.2.2 Entrepreneurs and Firms

Each entrepreneur operates a firm for their entire life. If an entrepreneur dies at the end of a period, their firm also exits the market. This implies an exogenous firm exit rate of $1 - \Delta$. Firms produce one type of output using capital, k, labor from workers, l, and a fixed amount of entrepreneurs' own labor, L^E. Denote the production function as $Y = \theta_s F(k, l, L^E)$. $F(.)$ exhibits decreasing returns to scale and is continuous and strictly concave. Firm productivity, θ_s, is subject to an idiosyncratic shock in each period and may end up in one of the two states, $s \in \{h, l\}$, with probability π_s $(\pi_h + \pi_l = 1)$.[9] Without loss of generality, assume $\theta_h > \theta_l$ and $\mathbb{E}(\theta_s) = 1$. The shocks are independent and identically distributed across entrepreneurs and over time. Hence, firms differ in both age and their history of productivity shocks. We assume that entrepreneurs have *private information* about productivity realizations in each period. Other agents can neither observe the realizations directly nor infer them from other sources of information, such as firms' output levels or entrepreneurs' consumption.[10] The asymmetric information generates important implications for the lending relationship between banks and firms, which are discussed in Section 2.3.1.

Firms borrow from banks to finance capital rents and labor costs incurred before production. The level of credit, b, and required repayments, m_s, are determined by banks according to the optimal financial contracts described in Section 2.3.1. A firm's profit maximization problem can thus be defined as follows: Given the loan terms and the factor prices, $\{r, w\}$, entrepreneurs decide the optimal demand for capital, k, and labor, l, to maximize the expected net revenues from production in each period. For succinct notations, we suppress all time indices of the

[8] Banks collect savings from workers of all cohorts and redistribute the proceeds (savings plus interest) to those who survive to the next period. To show this analytically, we have $(1 + r_{t+1}) \sum_{\tau=0}^{\infty} (1 - \Delta)\Delta^\tau p_t^A A_{\tau,t+1} = \sum_{\tau=0}^{\infty} (1 - \Delta)\Delta^{\tau+1}(1 + r_{t+1})A_{\tau,t+1}$. The left-hand side of the equation represents the proceeds from savings of all workers, where the term inside the summation symbol is the total savings of workers of cohort τ at time t. The right-hand side is the gross returns to workers who have survived (i.e., a share Δ of the original cohort size) to period $t + 1$. The equation implies that $p_t^A = \Delta$.

[9] The model and the properties of the optimal contracts in Section 2.3.3 can be generalized to any finite number of states (e.g., $\mathcal{S} = \{1, 2, \ldots, S\}$) without major modifications.

[10] One interpretation of this assumption is an economy where the costs of monitoring firm activities are so high that banks' optimal strategy is to never incur the action.

variables.

$$\max_{k,l} \mathbb{E}(\theta_s)F(k,l,L^E) - \mathbb{E}(m_s), \tag{4}$$

subject to

$$wl + (r+\delta)k \leq b.$$

$(r+\delta)$ is the effective cost of capital that depreciates at rate $\delta \in (0,1)$. The expectations in the objective function reflect the fact that productivity realization is unknown by the time that firms decide the demand for inputs. Hence, firms' optimal capital, k^*, and labor demand, l^*, are independent of the current productivity realizations. Denote them, respectively, as

$$k^* = k(b;r,w), \quad l^* = l(b;r,w), \tag{5}$$

and the expected maximal revenue as $R(b;r,w) \equiv F(k^*,l^*,L^E)$.

Once firms make the optimal capital and labor demand, production starts under uncertainty. Entrepreneurs repay the bank loans using proceeds from production and consume the residual. To isolate the role of bank credit in determining firm size distribution and firm dynamics, we assume that entrepreneurs are hand-to-mouth and do not accumulate wealth. Therefore, their consumption in state $s \in \{h,l\}$ in any period is given by

$$c_s^E = \theta_s R(b;r,w) - m_s. \tag{6}$$

This also simplifies entrepreneurs' lifetime utility maximization problem to the static profit maximization problem defined in (4). At the end of the period, a share $1-\Delta$ of the incumbent firms exit the market. The rest together with the new entrants go into the next period and repeat the sequence of events. The timing within one period is summarized in Figure 2.[11]

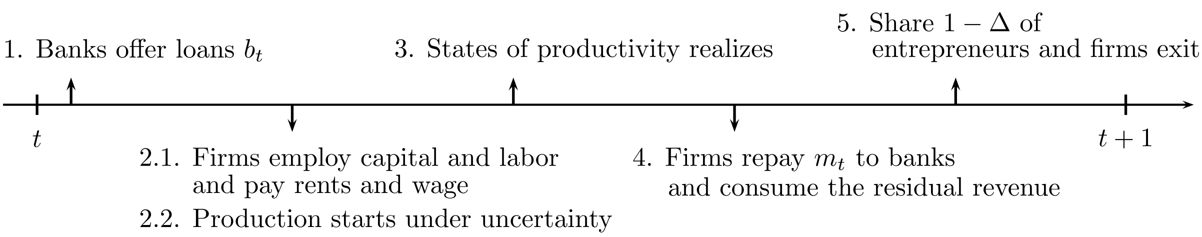


Figure 2: The timing of events within one period

Lastly, since the production technology is decreasing return to scale, there exists an efficient loan size, b^*, that maximizes firms' expected revenues net of the cost of production covered by bank loans:

$$b^* \equiv \arg\max_b \; R(b;r,w) - b. \tag{7}$$

[11]A detailed description of the timing in the model is provided in Appendix A.

We show in Appendix B that in a frictionless economy, b^* is the optimal level of bank loans in any contingencies. This is no longer the case in the current economy due to the existence of asymmetric information. In Proposition 2, we prove that endogenous borrowing constraints on firms emerge as a characteristic of the optimal lending contracts.

2.3 Banks and Lending Contracts

Banks are risk-neutral financial intermediaries that take deposits from workers and lend to entrepreneurs to cover operational costs.[12] We describe the dynamic lending contracts that determine the optimal level of loans and required repayments in Sections 2.3.1 and 2.3.2. We then explore the properties of the optimal lending contracts in Section 2.3.3. These properties generate important implications for firm dynamics.

2.3.1 Dynamic Lending Contracts

Banks offer each newborn entrepreneur a take-it-or-leave-it lifetime lending contract. The contract specifies a level of loans, b_t, and required repayments, m_t, in each period t.[13] Both banks and entrepreneurs are fully committed to the terms. But if a firm exits the market, its lending relationships with banks are terminated automatically.

Banks maximize their profits when deciding the optimal contract terms. The expected profits

[12]The assumption that firms borrow working capital from banks through long-term lending contracts to cover operational costs is partly motivated by the extensive literature on relationship lending. For example, by studying the bank-firm lending relationship leading up to the Great Depression, Cohen et al. (2021) showed that a large proportion of bank loans to non-financial firms was used as working capital. Importantly, these loans were unsecured, in the sense that they were either uncollateralized or collateralized by goods that were "difficult to value, costly to repossess, and, if liquidated, could be sold only with a long delay and/or at a considerable loss". Thus, these loans tended to be relationship-based, in that firms borrowed repeatedly from the same bank. The technical benefits of the assumption are discussed in Footnote 13 where dynamic lending contracts are described.

[13]In the current setup, the total amount of credit that banks provide to firms is $b_t + k_t$, where k_t is defined in (4). In return, firms pay banks m_t and the remaining physical capital, $(1-\delta)k_t$. The dynamic contracts may be formulated in alternative ways. For example, a bank may decide instead the *total* amounts of credit to a firm and of required repayments, and the firm decides the optimal allocation of the credit for capital employment and operational costs. But these alternatives will generally complicate the expressions of entrepreneurs' consumption in (6) and the capital-market-clearing condition (24), which show up in the constraints of the optimal contracts and the equilibrium conditions, respectively, without affecting the model implications. Besides, this may also introduce extra trivial margins in the firm's optimization problem described in (4). In particular, when deciding the optimal capital demand, firms will not only consider the marginal product of capital and labor as in (4), but also internalize the effect that a marginal increase in capital demand decreases the available funds to cover operational costs, which constrains firms' labor employment. Therefore, the optimal capital demand will be lower than the current setup, but the linear relationship between bank credit and capital demand remains.

from a lending contract are given by:

$$\Pi = \sum_{t=0}^{\infty} \left(\frac{\Delta}{1 + r_t} \right)^t \mathbb{E}\left(m_t - b_t\right), \tag{8}$$

where $\mathbb{E}(m_t - b_t)$ is the expected net cash flow in period t. Risk-neutral banks discount future cash flows at the current interest rate, r_t. Δ in the coefficient reflects the fraction of firms that survive to the next period.

In an economy *without* frictions (i.e., under perfect information), banks act like firm owners when deciding the contract terms. Specifically, they offer the efficient level of loans, b^*, defined in (7), and the actual revenues, $\theta_s R(b^*; r, w)$, as repayments in each period. The level of loans guarantees that firm revenues are maximized. The repayments are the maximum amount a bank can ask for under limited liability. This is the first-best scenario, where the level of bank loans is independent of firms' history of productivity realizations and repayments depend only on firms' productivity in the current period.[14]

However, this arrangement is no longer optimal in an economy *with* financial frictions. Under asymmetric information, banks only observe entrepreneurs' *reported* productivity realizations, $\hat{\theta}_t$, at any time t, but not the *true* ones, θ_t. If banks still provide the same level of loans in all circumstances and require state-contingent repayments, entrepreneurs will report whichever state leads to the lowest repayments. This violates the revelation principle which states that any equilibrium outcomes can be achieved by a truth-telling reporting strategy with $\hat{\theta}_t = \theta_t$. Therefore, the optimal contract terms under asymmetric information must be history-dependent and subject to incentive constraints as described in Section 2.3.2. In what follows, we no longer distinguish between the *reported* productivity and the *true* ones.

Let firms' productivity realizations leading up to period t be $h^t = (\theta_0, \theta_1, \ldots, \theta_{t-1})$. According to the timing of events in Figure 2, bank loans, b_t, are provided before production takes place and thus depend on the productivity history, h^t, up to period t. Repayments, m_t, are made after the current-period productivity, θ_t, realizes, and hence, is a function of both h^t and θ_t. Denote the contract terms as follows:

$$b_t = b(h^t), \quad m_t = m(h^t, \theta_t). \tag{9}$$

Lastly, the banking sector is perfectly competitive and allows free entry. Therefore, banks expect zero profits from each contract in equilibrium, and the size and ownership of banks are irrelevant.[15] In the remainder of the paper, we assume that one bank provides financing to all firms in the economy.

[14]The efficient level of bank loans, b^*, is independent of the lending scheme between banks and firms, and banks and workers. It only depends on the production technology of the economy. In particular, the scenario is equivalent to assuming that banks own the technology and decide the operation size accordingly.

[15]Notice that the zero-profit condition does not necessarily imply zero bank equities. Specifically, bank profit

2.3.2 Recursive Formulation of the Optimal Contract

The possible paths of productivity realizations increase exponentially as firms age, even with only two productivity realizations, $\theta_s \in \{\theta_h, \theta_l\}$. This makes it infeasible to derive the optimal contracts denoted in (9) in a sequential problem. Following the literature on dynamic contracts, we formulate the optimal contracting problem recursively using promised values as a state variable.[16] A promised value, V_t^E, at time t is the present value of an entrepreneur's expected utility from the sequence of consumption, $\left\{c_i^E\right\}_{i=\{t,t+1,\dots\}}$, starting from period t:

$$V_t^E \equiv \sum_{i=t}^{\infty} (\beta\Delta)^{i-t} \mathbb{E} U(c_i^E, L^E), \tag{10}$$

where consumption at time i, $c_i^E = \theta_i R(b_i; r_i, w_i) - m_i$, $i \in [t, +\infty)$, is defined in (6). The expectation is taken over firms' potential paths of firms' productivity history that determines both bank loans and repayments. Similar to workers, entrepreneurs discount future utility with the time discount factor, β, and the survival rate, Δ.

According to the literature on dynamic contracts, the promised value, V_t^E, summarizes information about firms' entire history of productivity realizations until time t. We can substitute h^t in (9) with V_t^E and reformulate the contract terms as $b_t = b(V_t^E)$ and $m_t = m(V_t^E, \theta_t)$, respectively. In addition, the promised value tomorrow, V_{t+1}^E, summarizes one more productivity realization than the promised value today, V_t^E. Thus, the transition function of the state variable is given by, $V_{t+1}^E = V(V_t^E, \theta_t)$. Taking away the time index, a lending contract can be formulated recursively as $\left\{b(V^E), m_s(V^E), V_s(V^E)\right\}_{s\in\{h,l\}}$, where the subscript s indicates that the repayments and the future promised values depend on the productivity realization today, $\theta_t \in \{\theta_h, \theta_l\}$.

Further, we can write the sequential form of banks' profits defined in (8) in the equivalent recursive form:

$$\Pi = -b + \sum_{s\in\{h,l\}} \pi_s \left[m_s + \frac{\Delta}{1+r} P(V_s^E; r', w') \right]. \tag{11}$$

In addition, banks are subject to the following constraints when deciding the optimal lending contracts.

First, the promised value V^E must be fulfilled by granting entrepreneurs consumption today, $c_s \equiv \theta_s R(b) - m_s$, and in the future via the future promised values, V_s^E, $s \in \{h, l\}$. This defines

is defined as the present value of the future cash flows from a lending contract (i.e., discounting the cash flows to the beginning of the lending contract), whereas the level of bank equities at any time is the sum of all previous cash flows until that time.

[16] Green (1987) and Spear and Srivastava (1987) are the first to prove the existence of a recursive formulation for dynamic optimization problems, where the principle and the agent interact in a repeated moral hazard context. Thomas and Worrall (1990) and Atkeson and Lucas (1992)

the promise-keeping (PK) constraint that governs the intertemporal allocation of consumption for entrepreneurs:

$$V^E = \sum_{s \in \{h,l\}} \pi_s [U\left(\theta_s R(b) - m_s, L^E\right) + \beta \Delta V_s^E]. \tag{PK}$$

Second, entrepreneurs have limited liability (LL). Thus, entrepreneurs are entitled to non-negative consumption in any contingencies and banks cannot require repayments higher than the realized revenues from firm production.

$$m_s \leq \theta_s R(b), \quad \forall s \in \{h, l\}. \tag{LL}$$

Third, the optimal contracts must be incentive-compatible (IC). This requires that entrepreneurs' expected utility from reporting the true productivity realizations is no less than that from misreporting, and thus eliminates their incentive to misreport.

$$U\left(\theta_i R(b) - m_i, L^E\right) + \beta \Delta V_i^E \geq U\left(\theta_i R(b) - m_j, L^E\right) + \beta \Delta V_j^E, \ \ \forall i, j \in \{h, l\}. \tag{IC}$$

The (IC) constraint also suggests that any gap between the repayments in the two states, $m_i - m_j$, must be compensated by a gap in the future promised value, $V_j^E - V_i^E$. Otherwise, entrepreneurs will always report the state that requires lower repayments. The first-best scenario described in Section 2.2.2 violates the (IC) constraint and is thus no longer optimal. Firms with high productivity shock will misreport their type and benefit from higher consumption today without affecting their future stream of consumption.

Lastly, the promised values, V^E, must be credibly fulfilled with non-negative and finite cash flows. Otherwise, they are either granted by violating the limited-liability constraint in some future periods or are never satisfiable. Denote the lifetime utility of an entrepreneur from a sequence of zero and infinite consumption, respectively, by

$$V_{min}^E \equiv \lim_{c \to 0} \frac{1}{1-\beta\Delta} U(c, L^E) \ \text{ and } \ V_{max}^E \equiv \lim_{c \to \infty} \frac{1}{1-\beta\Delta} U(c, L^E).$$

The credibility constraint indicates that V_{min}^E and V_{max}^E are natural bounds of the promised values.

$$V^E \in [V_{min}^E, V_{max}^E]. \tag{CC}$$

Numerical results in Section 3.2 show that (CC) is essential in determining the analytical properties of the optimal contracts in the region close to V_{min}^E (see Section 3.2 for a detailed discussion).

To sum up, the optimal lending contracts, $\left\{b, m_s, V_s^E\right\}_{s \in \{h,l\}}$ are defined by the following program. Given the factor prices, $\{r, w\}$, banks maximize their expected profits defined in (11) subject to the constraints:

$$\max_{\{b,m_s,V_s^E\}_{s\in\{h,l\}}} \Pi = -b + \sum_{s\in\{h,l\}} \pi_s \left[m_s + \frac{\Delta}{1+r} P(V_s^E; r', w') \right] \tag{12}$$

s.t., (PK), (LL), (IC), and (CC).

2.3.3 Properties of the Dynamic Lending Contracts

In this section, we discuss the theoretical properties of dynamic lending contracts and their implications on firms' credit availability.

The optimal contracts are incentive compatible. Proposition 1 shows that banks induce the truth-telling behavior of entrepreneurs by postponing rewards for reporting a high productivity realization. Specifically, they require higher repayments from high-productivity firms and compensate them with higher future promised values.

Proposition 1. *An incentive-compatible contract satisfies* $m_h \geq m_l$ *and* $V_h^E \geq V_l^E$*, where* $m_h = m_l$ *if and only if* $V_h^E = V_l^E$*. Moreover, the optimal contract implies that* $c_h \geq c_l$*.*

Proof. See Appendix C.1.

Proposition 1 also implies that the optimal lending contracts provide entrepreneurs only imperfect insurance over idiosyncratic shocks. Specifically, entrepreneurs who report low productivity repay less (i.e., they are insured), but their consumption is lower than that after a high productivity shock (i.e., the insurance is imperfect). This proposition indicates that the lending contracts are a constrained optimum due to the presence of asymmetric information.

Moreover, the optimal lending contracts imply endogenous financial constraints on firms. The result is summarized in the following proposition.

Proposition 2. *For a strictly concave value function* $P(V^E)$*, the optimal level of bank loans is no larger than the efficient level* b^* *defined in* (7)*.*

Proof. See Appendix C.2.

As is shown in Section 2.2.2, banks provide the efficient level of bank loans, b^*, in a frictionless economy. The tradeoff is between the marginal cost of providing loans and the marginal revenues from firm production using the additional unit of loans. In an economy with asymmetric information, however, the marginal benefits of providing more loans are lower because higher loans tighten the (IC) constraint. Specifically, as firm size increases, the output gap, $(\theta_h - \theta_l)R(b; r, w)$, between a high and a low productivity state widens. This increases the potential utility gain of entrepreneurs with high productivity shocks from misreporting their productivity realizations. To maintain the truth-telling incentives, banks must compensate them by widening the gap, $V_h^E - V_l^E$, in the future promised values between high and low states. This is costly due to the concavity of the profit function, $P(V^E)$. Therefore, the optimal level of bank loans is below the efficient level, b^*.

2.4 Stationary General Equilibrium

So far, we have characterized the optimization problems of all agents in the economy and got (i) workers' optimal paths of consumption, savings, and labor supply from (1); (ii) entrepreneurs' optimal capital and labor demand from (4); and (iii) banks' optimal lending contracts in Section 2.3.2. In this section, we define the stationary equilibrium of the economy and explain the conditions that determine three equilibrium variables: the interest rate r, the wage rate w, and firm entry measured by the share, λ, of newborn households that start a firm.

2.4.1 Equilibrium Conditions

Due to free entry and perfect competition in the banking sector, banks expect zero profit from each newly signed contract in equilibrium. Using the notation of bank profits in (12), we get

$$P(V_0^E; r, w) = 0. \tag{13}$$

V_0^E is the initial promised value to a newborn entrepreneur, and also their lifetime expected utility. In equilibrium, households are indifferent between the two occupations, and thus V_0^E equals the expected lifetime utility of a worker, $V^W(0; r, w)$, defined in (1):

$$V_0^E = V^W(0; r, w). \tag{14}$$

Plugging (14) into (13), we get an equation of the factor prices, r and w:

$$P(V^W(0; r, w); r, w) = 0. \tag{15}$$

In addition, labor, capital, and goods markets clear in equilibrium. We derive the market-clearing conditions for the first two here, and the last one clears automatically according to Walras's Law. The derivation is shown in Appendix C.5.

Equilibrium in the labor market requires that the aggregate labor supply, L^S, from workers equals the aggregate demand, L^D, from firms:

$$L^S = L^D. \tag{16}$$

L^S and L^D are integrals over the optimal decisions of individual workers and entrepreneurs, $h(A; r, w)$ and $l(V^E; r, w)$, defined in (3) and (5), respectively. Denote the distributions of the two state variables, A and V^E, as $\psi^W(A; r, w)$ and $\psi^E(V^E; r, w)$. We get

$$L^S \equiv (1-\lambda) \int h(A; r, w) d\psi^W(A; r, w), \quad L^D \equiv \lambda \int l(V^E; r, w) d\psi^E(V^E; r, w), \tag{17}$$

where $1 - \lambda$ and λ reflect the shares of workers and entrepreneurs in the total population, respectively. Similar techniques are applied when calculating the aggregate variables in the capital and goods markets. In principle, λ, $\psi^W(.)$, and $\psi^E(.)$ may vary over time, as new workers and entrepreneurs enter the market and incumbents exit or see their state variables evolve. In this paper, we focus on the stationary general equilibrium that features a constant (but endogenous) share of entrepreneurs in each cohort, and invariant distributions of the state variables. Proposition 3 shows that such distributions exist. An illustration of the distribution of V^E is given in Appendix H.1 Figure 14b.[17] The definition of a stationary equilibrium is provided in Section 2.4.2.

Proposition 3 (Stationary distributions of A and V^E)**.** *For a given interest rate and wage rate (r, w), there exists a stationary distribution of workers, $\psi^W(A; r, w)$, and of entrepreneurs, $\psi^E(V^E; r, w)$.*

Proof. See Appendix C.4.

The capital-market-clearing condition requires that the aggregate capital demand, K^D, equals the aggregate supply, K^S, in equilibrium:

$$K^D = K^S. \tag{18}$$

K^D is an integral over firms' optimal capital demand, $k(V^E; r, w)$ defined in (5).

$$K^D \equiv \lambda \int k(V^E; r, w) d\psi^E(V^E; r, w).$$

The capital supply comes from banks. In particular, banks use deposits from workers, D, and their equities, E, to finance the provision of bank loans, B, and capital, K^S:

$$K^S + B = D + E. \tag{19}$$

The aggregate deposits and bank loans are given respectively by

$$D \equiv (1-\lambda) \int A d\psi^W(A; r, w), \tag{20}$$

[17] Since workers of the same age save the same amount, the savings take discrete values, $\{A_\tau\}_{\tau \in \{0,1,\ldots\}}$, with A_τ being the level of savings of cohort τ. The probability of A_τ is thus the same as the population size of cohort τ. According to Footnote 6, $\Pr(A = A_\tau) = (1-\Delta)\Delta^\tau$.

and

$$B \equiv \lambda \int b(V^E; r, w) d\psi^E(V^E; r, w), \tag{21}$$

where $b(V^E; r, w)$ is the optimal level of bank loans defined in (12). The levels of bank loans and capital supply are related at the firm level via firms' optimal capital demand.

Banks accumulate equities from retained earnings. This includes net income from loans, $M - B$, capital rents, rK^S, net of the interest payments on deposits, rD. Therefore, banks' equities in the next period, E', come from their equities today, E, and the retained earnings:

$$E' = E + M - B + rK^S - rD, \tag{22}$$

where $M \equiv \lambda \int \sum_{s \in \mathcal{S}} \pi_s m_s(V^E; r, w) d\psi^E(V^E; r, w)$ represents the aggregate loan repayments from entrepreneurs, with $m_s(V^E; r, w)$ being the optimal repayments defined in Section 2.3.2.

In a stationary equilibrium, bank equities are constant over time. Namely, $E' = E$. Substituting K^S in (22) with equation (19), we get the equation that defines the equilibrium level of bank equities as a function of bank loans and repayments:

$$rB = rE + M - B. \tag{23}$$

Intuitively, providing loans to entrepreneurs brings in net earnings of $M - B$, but generates opportunity costs of rB due to the loss of capital rents. Besides, by financing part of the lending to firms with internal funds (i.e., equities), banks save the interest payments, rE, on deposits. In a stationary equilibrium, the opportunity costs of providing loans, rB, must be exactly compensated by the gains from loans and self-financing, $M - B + rE$. This gives the two sides of equation (23).

Using (18), (19), and (23), we get the capital-market-clearing condition as a function of the aggregate variables that are integrals of individuals' optimal decisions:

$$K^D = D + \frac{B - M}{r}. \tag{24}$$

The labor- and capital-market-clearing conditions defined in (16) and (18), together with the zero-profit condition in (15), allow us to solve for the factor prices and share of entrepreneurs, $\{r, w, \lambda\}$, in equilibrium.

2.4.2 Definition of a Stationary General Equilibrium

With the equilibrium conditions and the optimization problems of all agents, we define the stationary general equilibrium in the economy.[18]

[18] We prove the uniqueness and existence of such an equilibrium, using the quantitative properties of the bank-profit and the excess-capital-demand equations. A general proof of the uniqueness and existence of stationary

Definition 1. *A stationary general equilibrium is characterized by a stationary distribution, $\psi^W(A; r, w)$, of savings among workers, and their capital and labor supply, $\{g(A; r, w), h(A; r, w)\}$, a stationary distribution, $\psi^E(V^E; r, w)$, of promised values among entrepreneurs, and their capital and labor demand, $\{k^*(V^E; r, w),\ l^*(V^E; r, w)\}$, the levels of bank loans and repayments, $\left\{b(V^E; r, w), m_s(V^E; r, w)\right\}_{s\in\mathcal{S}}$, the interest rate, the wage rate, and the share of entrepreneurs, $\{r, w, \lambda\}$, such that for given (r, w),*

(1) workers maximize the lifetime utility according to (1)*,*

(2) entrepreneurs maximize the expected output according to (4)*,*

(3) banks maximize expected profits from the dynamic lending contracts according to the program defined in Section 2.3.2.

The factor prices, (r, w), and the share of entrepreneurs, λ, are such that,

(1) banks make zero profit in expectation and households are indifferent between the two occupations according to (15)*.*

(2) labor, capital, and goods markets clear according to (16)*,* (18)*, and* (C.13)*, respectively.*

Analytical solutions and properties of optimal dynamic contracts under asymmetric information are difficult to derive in general due to non-convex constraints.[19] The problem is even more complex when dynamic contracts are embedded in a general equilibrium framework with additional conditions and variables. We solve for the stationary equilibrium numerically and show the results in the next two sections. A sufficient condition for the existence of a *locally* unique stationary equilibrium is derived in Appendix F and the algorithm for solving the model numerically is described in Appendix E.3.

3 Calibration and Numerical Results

In this section, we describe the quantitative analysis of the model. We illustrate the optimal lending contracts and the main results on firm dynamics – firm size, firm growth, and the volatility of growth over time. To calibrate the model, we assume specific functional forms of

equilibrium is still missing from the literature. The quantitative properties we used are straightforward and intuitive and should hold in a more general context. Even if they do not, these properties can be interpreted as sufficient conditions such that a stationary general equilibrium exists and is unique.

[19]This is unlike dynamic contracting under limited commitment where analytical solutions and properties generally exist. See, for example, Albuquerque and Hopenhayn (2004) and the related chapter in Ljungqvist and Sargent (2018).

the utility and the production functions and give exogenous parameter values summarized in Table 1.

One period in the model corresponds to one year in reality. The survival rate $\Delta = 0.92$ is chosen such that the death rate $1 - \Delta = 8\%$ corresponds approximately to the annual exit rate of manufacturing firms in the United States. Households' instantaneous utility function, shared by workers and entrepreneurs, is given by

$$U(c, l) = -exp(-\gamma c) - \eta l^2, \quad \gamma, \eta > 0. \tag{25}$$

It includes a CARA part for consumption with γ being the absolute risk aversion and a parabola part for the disutility of labor supply. This functional form gives us computational simplicity. The values of the preference parameter, $\gamma = 2$ and $\eta = 0.5$, are internally calibrated such that workers' optimal labor supply is about 30% of their labor endowment. Entrepreneurs' fixed labor input, L^E, is set to a third of their labor endowment. Households' future discount rate $\beta = 0.963$ is within the range of standard values in the literature. Following the functional form and the parameter values of the utility function, the boundaries of the promised value are $V^E_{min} = -9.26$ and $V^E_{max} = -0.49$.

Table 1: Parameters Values

Parameters		Value
Survival rate	Δ	0.92
Discount rate	β	0.963
Household preferences	γ	2
	η	0.5
Probability of high state	π_h	0.5
Expectation of firm productivity	θ	1
Standard deviation of firm productivity	σ	0.25
Fixed entrepreneur labor	L^E	1/3
Share of capital	α_k	0.35
Share of labor	α_l	0.6
Productivity scale	$\bar{a}$	1/3
Depreciation rate	δ	10%

Firms' production function is given by

$$Y(k, l) = \theta_s \bar{a} k^{\alpha_k} l^{\alpha_l}, \quad s \in \{h, l\}, \tag{26}$$

where $\bar{a}$ denotes the level of total factor productivity (TFP) and θ_s is the idiosyncratic shock on TFP. Let the two realizations of θ_s be $\theta_h = \theta + \sigma$ and $\theta_l = \theta - \sigma$, $\sigma > 0$. We assume that the two states happen with the same probability and normalize the expectation of θ_s, θ, to 1 and the standard deviation σ to 0.25. $\alpha_k = 0.35$ and $\alpha_l = 0.6$ represent the share of capital and labor employed in firm production, respectively, with the values set to be consistent with data. $\alpha_l + \alpha_k < 1$ implies that the production function exhibits decreasing returns to scale.[20] The depreciation rate $\delta = 10\%$ corresponds to a common number in literature reflecting a quarterly depreciation rate of approximately 2.5%.

3.1 Stationary General Equilibrium

We simulate the history of productivity shocks of ten million entrepreneurs to receive a stationary distribution of the promised values.[21] Using the parameters in Table 1 and the simulated distribution, we solve for the stationary equilibrium of the economy numerically.[22] The results are summarized in Table 2.

Table 2: Equilibrium Variables

Equilibrium factor prices and firm entry		Value
Interest rate	r^*	4.17%
Wage	w^*	0.16
Share of entrepreneurs	λ^*	7.64%
Aggregate variables		Value
Aggregate bank loans	B	0.067
Bank equities	E	0.093
Aggregate labor employment	L^D	0.266
Aggregate capital employment	K^D	0.175
Aggregate output	Y	0.072

Note: The lifetime utility of an entrepreneur, $V_0^E = -8.358$, is identical to that of a worker in equilibrium according to (14). The aggregate variables are calculated with equations (17), (20)-(23), taking into account the shares of workers, $1 - \lambda^*$, and of entrepreneurs, λ^*, respectively.

[20]This is consistent with the estimates of empirical studies on the U.S. industry, which indicate that average firms have constant or slightly decreasing returns to scale (Basu and Fernald, 1997).

[21]The distribution is stationary in the sense that even if we further increase the number of entrepreneurs, the changes in the aggregate variables (e.g., aggregate bank loans, repayments, etc.) calculated from the new distribution of promised values will be within a computationally tolerated range. The technical details are described in Appendix E.3.

[22]The algorithm for solving the stationary equilibrium can be found in Appendix E.4. We provide the theoretical intuitions for the existence and uniqueness of the stationary equilibrium locally in Appendix F.

The interest rate in the calibrated economy is 4.17%. The average labor supply per worker $L^S/(1-\lambda) = 0.29$ corresponds to about 30% of a worker's labor endowment. Both the interest rate and the labor supply are consistent with the standard values from the macroeconomic literature. The equilibrium share of entrepreneurs is 7.64%, which is approximately the share of self-employed labor in the United States over the last few years (data from OECD).

3.2 Optimal Lending Contracts

Given the equilibrium factor prices in Table 2, we solve for the optimal lending contracts and banks' profits (i.e., the value function) defined in Section 2.3.2. The results are illustrated in Figure 3.[23] The worker's problem defined in (1) is a standard lifetime utility maximization problem and the entrepreneurs' optimal demand for capital and labor can be solved analytically. We show the numerical solutions in Appendix D.

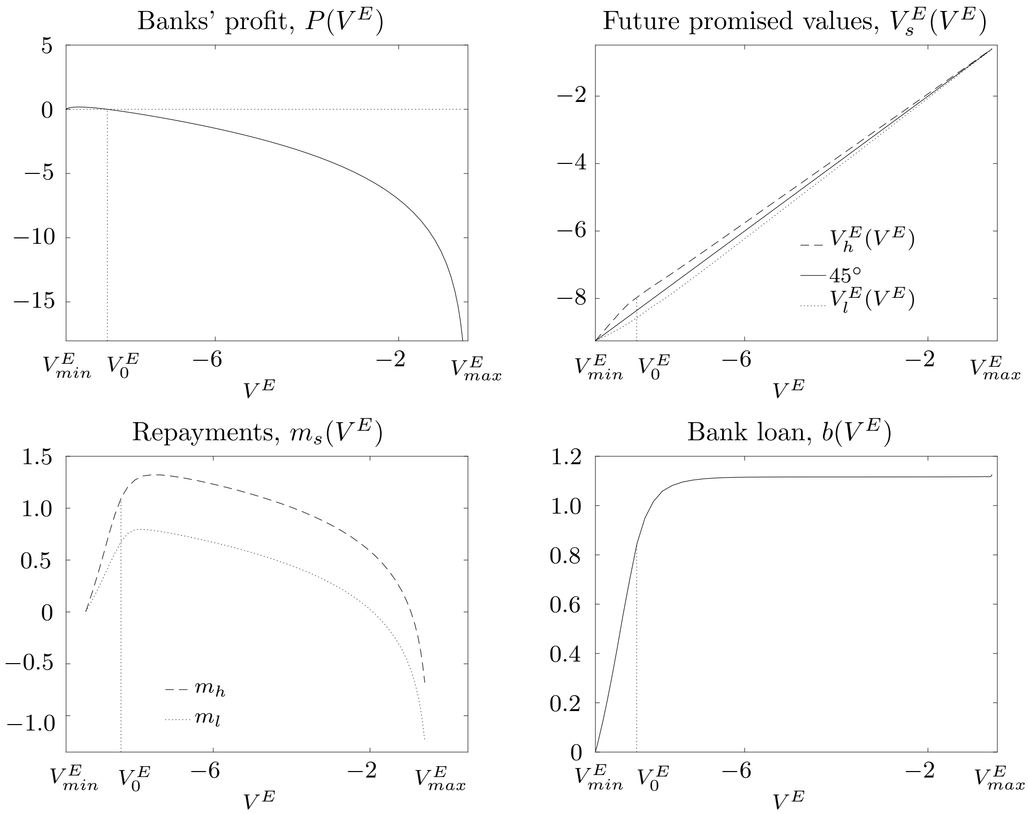


Figure 3: The value function, $P(V^E)$, and the optimal contract, $\{b, m_s, V_s\}_{s\in\{h,l\}}$

The bank profit, $P(V^E)$, is strictly concave and decreases with the promised value, V^E. The concavity is a common property of dynamic contracts and the intuition for the negative rela-

[23]See Appendix E.2 for the numerical procedure to solve the recursively formulated lending contracts.

tionship is as follows.[24] Firms' production revenues in each period are shared between banks and entrepreneurs. As the promised value, V^E, increases, a greater share of the revenues is consumed by entrepreneurs, and thus bank profits decrease. Free entry and perfect competition in the banking sector suggest that banks make zero expected profits from a newly signed contract in equilibrium. Therefore, $P(V^E) = 0$ pins down the initial promised value, V_0^E, of an entrepreneur, which is their lifetime expected utility at birth.

The optimal level of bank loans, $b(V^E)$, in subplot 4 strictly increases in V^E and converges gradually to the efficient level, $b^* = 1.18$, defined in (7).[25] This suggests the presence of endogenous borrowing constraints due to asymmetric information proved in Proposition 1 and the relaxation of the incentive constraint as the promised value increases. Specifically, at a higher promised value, entrepreneurs are granted higher consumption. This implies a lower marginal utility and hence a lower incentive to misreport their productivity realization to receive higher consumption. Therefore, the optimal level of bank loans converges to the efficient level in a frictionless economy.

Repayments, $m_s(V^E)$, decrease in the promised value, except for a small region close to $V^E = V^E_{min}$ to be discussed at the end of this section. This implies that banks fulfill a higher promised value by lowering the required repayments, which increases entrepreneurs' consumption.[26] In fact, as illustrated in Figure 4, banks make net transfers to firms with high promised values and extract positive earnings from those with low promised values. Since firms' promised values increase with firm age on average in our numerical specification (see Section 3.3), older firms receive more consumption from banks. In other words, the optimal contracts are back-loaded with respect to entrepreneurs' consumption.

The future promised values, $\left\{V_s^E(V^E)\right\}_{s\in\{h,l\}}$, in subplot 2 govern the evolution of V^E. In line with Proposition 1, firms with higher promised values today, V^E, will receive higher promised values in the future. Starting from the initial promised value, V_0^E, firms that experience more high productivity shocks end up with high promised values and may grow out of the financial constraint, whereas the financing status of firms with more low productivity shocks will dete-

[24]Ljungqvist and Sargent (2018) provide detailed discussions about the properties of the value function in a dynamic contracting problem.

[25]The specific shape of $b(V^E)$ is the result of the functional forms of the utility and the production function and their relative curvature compared to each other (see (C.11) in Appendix C).

[26]The promise-keeping constraint (PK) requires that $V^E = \sum_{s\in\{h,l\}} \pi_s \left[u(c_s, L^E) + \beta \Delta V_s\right]$, where $c_s = \theta_s R(b) - m_s$ denotes entrepreneurs' consumption in state $s \in \{h, l\}$, and $R(b)$, defined in (4), is a firm's output given bank loans b. As V^E increases, banks may fulfill the higher promised value by increasing the future promised values V_s, increasing the level of bank loans b, and/or decreasing the repayments m_s, the latter two of which increase an entrepreneur's consumption today c_s. However, the costs of increasing V_s and b rise significantly at higher V^E, due to the steep decrease in the value function $P(V^E)$ and the decreasing marginal product of bank loans, respectively. Therefore, the optimal choice of banks is to lower the required repayments m_s.

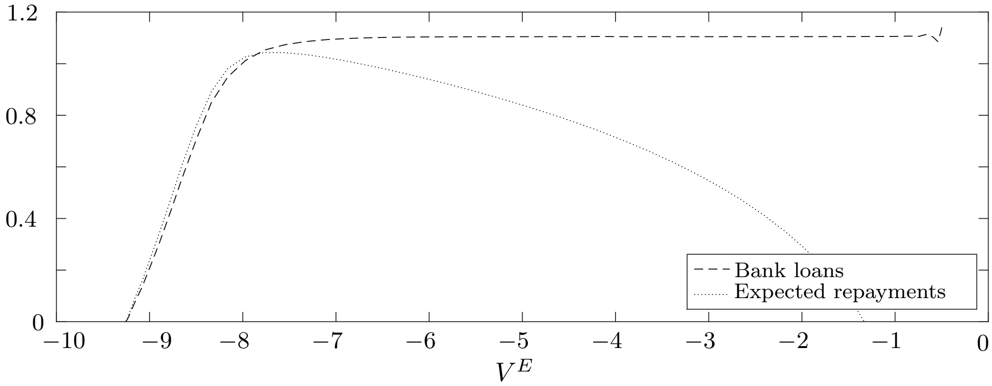


Figure 4: Bank loans, $b(V^E)$, and expected repayments, $\sum_{s\in\{h,l\}} \pi_s m_s(V^E)$

riorate.[27] The combination of requiring higher repayments and granting high future promised values after a high productivity shock ensures the truth-telling behaviors of entrepreneurs.

Lastly, we discuss the properties of the optimal contracts in the small neighborhood of the lower bound. By definition, $V^E = V^E_{min}$ corresponds to a lifetime stream of zero consumption. According to the credibility constraint (CC) and the limited liability constraint (LL), banks can neither promise a utility below V^E_{min}, nor are they allowed to require a repayment higher than the firms' actual level of output. Therefore, the only contract that satisfies all the constraints is $b = 0, m_s = 0, V^E_s = V^E_{min}, s \in \{h, l\}$. In other words, $V^E = V^E_{min}$ is an absorbing state for firms, where banks offer no loans and make zero profits from them. For any firms in the neighborhood of V^E_{min}, the set of feasible contracts expands. The high marginal product of loans close to $b = 0$ incentivizes banks to increase the provision of bank loans sharply so that they can extract positive repayments from firms. This explains the increasing segments of $b(V^E)$ and $m_s(V^E)$ in Figure 4. Banks' revenues from higher repayments dominate the cost of providing higher promised values to entrepreneurs. In the small neighborhood of V^E_{min}, $P(V^E)$ increases with V^E.

3.3 Firm Dynamics

In this section, we show the firm dynamics derived from the simulated paths of productivity shocks and the optimal contract in Section 3.2. Figure 5 summarizes the main results, which include the age-size relationship, the growth rate, and the variance of growth of firms at different ages τ (referred to as "cohort τ"). A details procedure for getting these variables is given in Appendix H.2.

[27] In Appendix H.3, we illustrate three examples of firm productivity shocks over 50 years and the corresponding dynamics of promised value, the level of bank loans, and repayments.

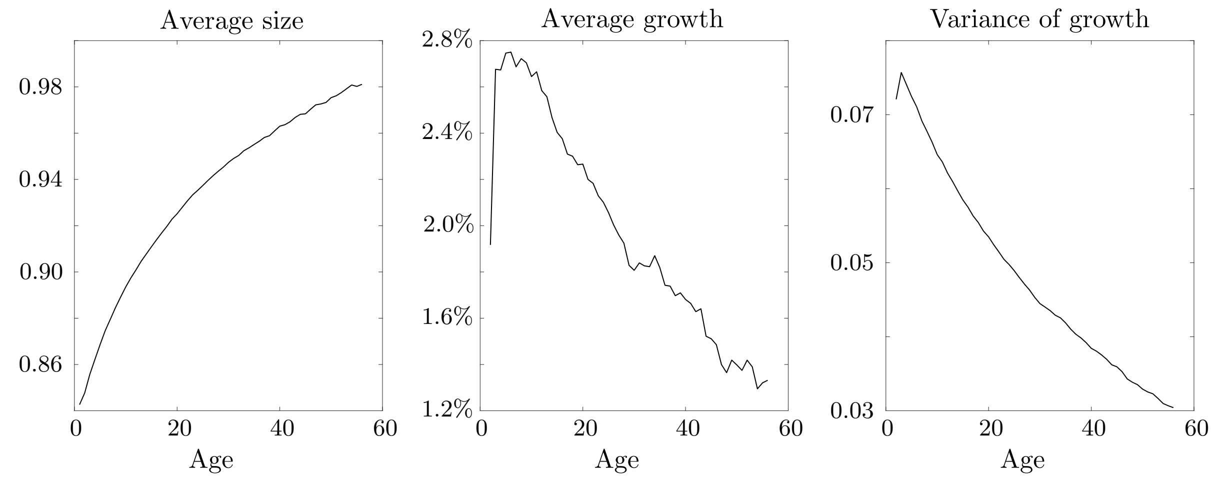


Figure 5: Firm dynamics

Note: The average size is the average level of bank loans that firms of each cohort receive. Firm growth is the percentage increase in the level of bank loans from that in the previous period. The average and the variance of growth rates are the mean and variance of firm growth in each cohort. All three variables illustrated are the 5-year moving average (e.g., the value of the average size at age 10 is the weighted average size of firms with age 10-14, with the weights being the population size of each cohort).

Subplot 1 illustrates a positive relationship between firm size and firm age. The firm size at age τ is measured by the average level of bank loans firms of age τ receive.[28] Initially, all newborn firms receive the same amount of loans, $b(V_0^E)$, determined by the initial promised value V_0^E. The level is well below the efficient level of bank loans defined in (7), which implies that all firms are financially constrained at birth. The situation improves for an average firm over time due to an increase in the expected promised value. According to the sufficient conditions in Proposition 4, which are satisfied under the numerical specification, it is optimal for banks to postpone the rewards to the future by promising entrepreneurs higher future cash flows.[29]

Firms exhibit positive growth in size over time. But subplots 2 and 3 suggest that the average and the variance of the growth rates decrease as firms get older.[30] The growth rate, g_τ, of a firm at age τ is defined as the percentage increase in the level of bank loans from the previous period, $g_\tau \equiv \frac{b_\tau - b_{\tau-1}}{b_{\tau-1}}$. The average and the variance of the growth rates are the mean and the variance of g_τ among firms of cohort τ, respectively. The results are mainly driven by a diminishing gap between V_h^E and V_l^E as the promised value V^E rises. Increasing the gap between V_h^E and V_l^E is

[28] Firm size can be equivalently defined as the optimal level of capital k^* or labor employment l^* for a given level of bank loans, b, defined in (D.1).

[29] The sufficient conditions for $\mathbb{E}(V_s^E) \geq V$ are $\beta(1+r) \geq 1$ and that the first-order derivative of the value function, $P'(.)$, is non-concave. Both conditions are satisfied in the numerical analysis.

[30] The observation discussed in Footnote 28 explains the outlier of the average growth (and also of the variance) in the first year. Note that the less smooth pattern for young firms is because they have less diverse paths of productivity shocks. In this sense, there is only a small number of observations on firm size, despite the large number of young firms. The less smooth pattern for older firms arises since firms are dying and not many observations are left.

costly due to the concavity of banks' profit function, $P(V^E)$. But a wider spread allows banks to extract more repayments from firms without violating the incentive-compatibility constraint. As V^E increases, entrepreneurs' marginal utility from today's consumption decreases, and hence their incentive to misreport the productivity realization is lower. This diminishes banks' need to increase the spread between V_h^E and V_l^E. At a smaller gap between V_h^E and V_l^E, the change in V^E over time and the variance of the change are also smaller. Therefore, older firms grow less on average, but in a more stable way.

These patterns of firm dynamics are consistent with long-established empirical facts (e.g., Evans, 1987; Hall, 1987) and have been found in the literature on dynamic contracts. As was mentioned in Clementi and Hopenhayn (2006), being able to derive realistic firm dynamics with only i.i.d. idiosyncratic shock is an important advantage of using dynamic contracts to characterize the lending relationship between banks and firms. Generally, one would need specific assumptions on the productivity shocks on firms or the screening and learning process in a bank-firm relationship.[31]

4 Long-Run Effects of A Credit Contraction

In this section, we extend the benchmark model to study the long-run effects of a permanent credit contraction on the allocation of credit among firms, firm dynamics, and the macroeconomy. The credit contraction emerges due to an exogenous shock on banks' liquidity demand: A share, μ, of deposits from workers must be held as reserves. This changes the composition of bank assets illustrated in Table 3.

Table 3: Banks' balance sheets with reserves

Assets	**Liabilities**
- Reserves, μD	- Deposits from workers, D
- Credit to firms, $K^S + B$	**Equities:** Retained earnings, E

Even though the liquidity shock, μ, is exogenous, the decline in the credit provided to firms is determined endogenously. Specifically, a greater μ implies a change in the equilibrium factor prices, which in turn affects the size of banks' balance sheets via workers' savings decisions and

[31] Clementi and Hopenhayn (2006) have risk-neutral entrepreneurs. To generate a positive relationship between firm size and firm age, their model requires that the interest rates be large enough that banks allocate firm consumption into the future, which increases the future promised value and thus firm size. Our model avoids such restriction on the interest rates by introducing a non-trivial lower bound, V_{min}^E, for the promised value. As illustrated in Figure 3, the lower bound plays a crucial role in determining the property of the optimal contracts for V^E close to the region of the lower bound.

the optimal lending contracts to firms. We evaluate the long-run effects on the economy by solving and comparing the stationary equilibrium at various levels of reserve ratios in the range of $[0, 40\%]$ and a reserve rate of $r^e = 2\%$.[32] The numerical procedure, the functional forms, and the values of the other parameters are the same as in Section 3.

In Section 4.1, we show analytically how a credit contraction affects the equilibrium conditions of the economy. Section 4.2-4.4 illustrates the numerical results of the effects on the optimal contracts, the allocation of credit and firm dynamics, and the macroeconomic variables.

4.1 Equilibrium Condition in the Capital Market

According to the modified balance sheet in Table 3, we have

$$K^S + B = (1-\mu)D + E, \tag{27}$$

where B and D are defined in (21) and (20), respectively. The evolution of banks' equities, E, is now given by

$$E' - E = M - B + rK^S + r_e\mu D - rD, \tag{28}$$

where M is defined in (22) and $r_e\mu D$ represents the interests banks receive from reserves. In other words, the change in bank equities comes from the revenues on the assets minus the costs. In a stationary equilibrium, where $E' = E$, the costs must be exactly compensated by the revenues. Thus, we get the capital-market-clearing condition:

$$K^D = K^S = D + \frac{B-M}{r} - \frac{r_e}{r}\mu D, \tag{29}$$

The equation suggests that holding other variables constant, an increase in the reserve ratio, μ, decreases aggregate capital supply, and thus affects the economy via changes in the factor prices in general equilibrium. Besides, it reveals an interesting relationship between the interest rate, r_e, paid on reserves and the effect of the liquidity demand shock on the economy in stationary equilibrium. At a lower reserve rate, the negative impact of the liquidity demand shock on capital supply is smaller, implying a weaker overall effect on the economy. When the reserve rate is higher, the effects are stronger. The key to seeing this is the negative relationship between r_e and bank equities, E. In particular, by canceling the K^S from (27) and (29), we get

$$M - B + rE = rB + (r - r_e)\mu D. \tag{30}$$

[32]According to the H.8 release from the Federal Reserve on the assets and liabilities of commercial banks in the United States, the reserve ratio, defined as the ratio of cash assets (including vault cash, cash items in process of collection, balances due from depository institutions, and balances due from Federal Reserve Banks) to deposits, was within the range of 2% to 13% between 1973 and 2008. The ratio peaked at approximately 30% in the aftermath of the financial crisis. Papers on liquidity hoarding (e.g., Berrospide, 2021) suggest that banks hold not only more cash assets, but also other liquid assets. As an upper bound for banks' reserve ratio, we included in the liquid assets the sum of cash assets, securities in bank credit, and federal funds including reverse repos, and divide the sum by total deposits. The ratio peaked at 40%. Therefore, we set the upper bound of μ to 40%.

Intuitively, in an economy with a smaller r_e, the costs of holding reserves are greater because of the larger interest rate gap between deposits and reserves. To offset its negative impact on equities, banks must finance a greater share of their lending with internal funds (i.e., equities) instead of external debts (i.e., deposits). In a stationary equilibrium, this leads to a larger size of bank equities and an increase in the overall size of bank assets. This partly compensates for the decrease in capital supply.

In the extreme case that banks earn no interest on the reserves (i.e., $r_e = 0$), bank equities increase by the exact amount of the reserve holding, μD, in a stationary equilibrium, leaving the capital-market-clearing condition in (29) unaffected by the exogenous shock on liquidity demand. Put differently, the equilibrium outcomes will be independent of μ if the reserves are paid no interest.[33] On the other hand, if bank reserves earn the same return as the capital supplied to firms (i.e., $r^e = r$), bank equities are unaffected by μ directly. Intuitively, as long as the returns of the two assets are the same, it does not affect banks' cash flows or their equity accumulation whether the assets are held as reserves or provided as capital. Thus, in this case, capital supply to firms will decrease by the amount of banks' reserve holding.

4.2 Optimal Dynamic Contracts

In this section, we analyze the effect of the exogenous liquidity shock on the optimal contracts that banks provide to firms. An increase in the holding of liquid assets decreases the capital supply to firms according to the capital-market-clearing condition, (29). This drives up the interest rates and implies a lower wage rate and firm entry in the new stationary equilibrium, as are illustrated in Figure 6. A detailed analysis of the mechanism that drives the changes is provided in Section 4.4.

Banks take the change in the factor prices as given and decide the terms of the optimal contracts accordingly. The results are illustrated in Figure 7. The solid and the dashed lines correspond to the optimal contracts at $\mu = 0$ and $\mu = 40\%$, respectively. The arrows indicate the directions of the movements of the curves. Our quantitative analysis suggests a dominating effect from a higher interest rate, which we will focus on in the analysis below.

Three factors drive the rotation of banks' profit at different levels of promised values. First, firms' costs of production rise with the interest rate. Due to the limited-liability constraint, banks have to lower the required repayments from firms. This generates a negative impact on banks' profits from contracts at any promised values and thus a downward shift on the

[33] This endogenous relationship between r_e and bank equities is missing in Smith and Wang (2006). Even though $r_e = 0$ in their model, they assume a one-to-one decrease in capital supply K^S as banks hold reserves.

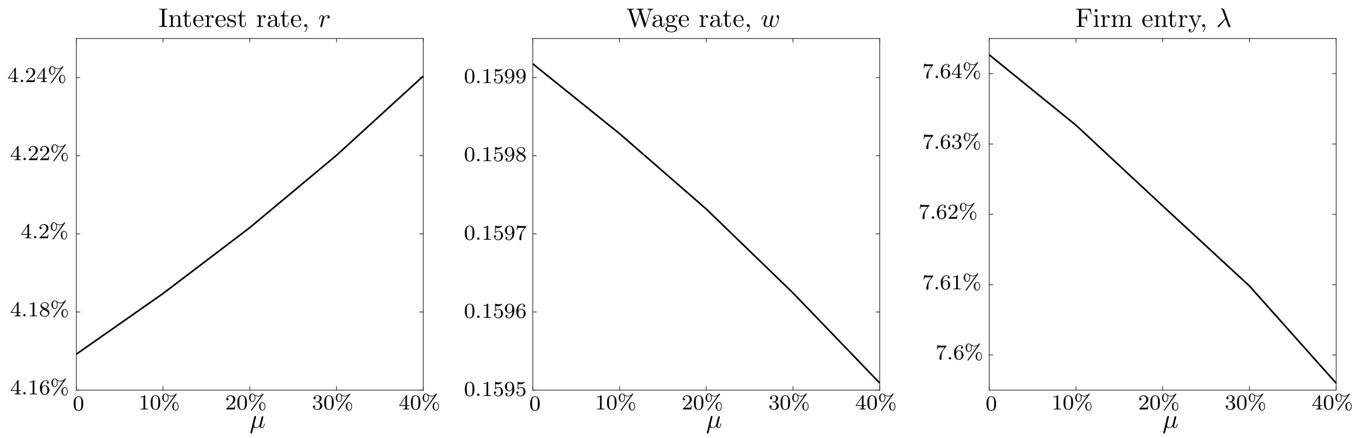


Figure 6: Effects of a credit contraction on equilibrium factor prices and firm entry

$P(V^E)$ curve. Second, when choosing the optimal contract terms at a higher interest rate, banks have more incentive to postpone granting the promised utility into the future and decrease entrepreneurs' consumption today (i.e., extract more repayments from entrepreneurs). This increases entrepreneurs' marginal utility of consumption today and incentivizes them to misreport their productivity realization. Therefore, inducing a truth-telling incentive defined in (IC) becomes more costly, which also decreases bank profits. Lastly, risk-neutral banks discount future cash flows with the interest rate. As was discussed in Section 3.2, banks' optimal intertemporal allocation of consumption to entrepreneurs is to receive positive cash flows from firms with low promised values and make net transfers to those with high promised values. An increase in the interest rate implies a decrease in profits from firms with low promised values and an increase (i.e., less negative profits) from firms with high promised values. The effect via the channel of the discount factor on bank profits dominates the first two negative effects in the domain of high promised values. Thus, we get a counterclockwise rotation of $P(V^E)$.

To understand the downward shifts of the curves for bank loans and repayments, notice that a higher interest rate lowers firm revenues and thus the efficient firm size of the economy defined in (7). As a result, banks provide lower levels of loans to firms at any promised values and the optimal loans converge to a lower efficient level over time. A lower level of bank loans decreases firm production in any contingencies. This drives down the amount of repayments that banks can require without violating the limited liability constraint. Therefore, the optimal $m_s(V^E)$ curves shift down.

Lastly, $V_s(V^E)$ increases for both states, $s \in \{h, l\}$. According to banks' objective function and the (PK) constraint in Section 2.3.2, the marginal costs of increasing future promised values, $V_s(V^E)$, decrease with the interest rate. Therefore, at a higher interest rate, it is more cost-efficient for banks to postpone fulfilling the promised values into the future, which drives up the future promised values in all contingencies. The intertemporal reallocation of promised values

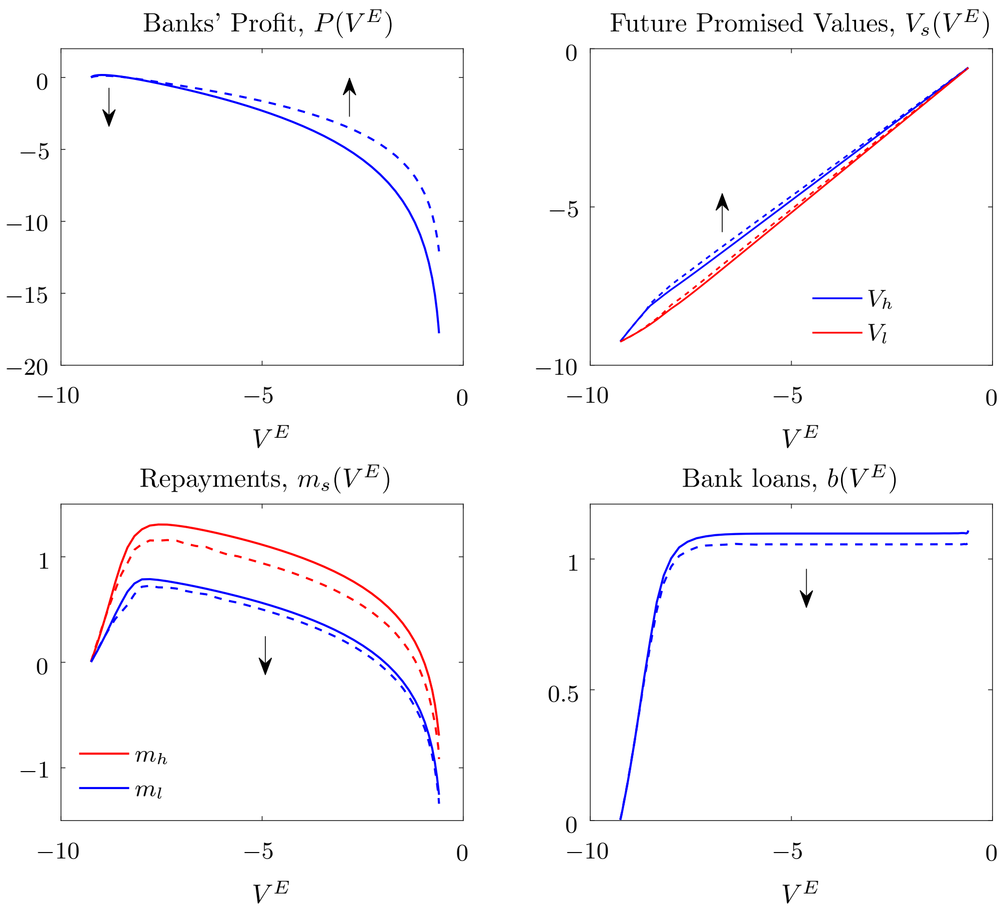


Figure 7: Effects of a credit contraction on optimal contracts

Note: The solid and the dashed lines correspond to the optimal contracts at $\mu = 0$ and $\mu = 40\%$, respectively. The movements of the curves, indicated by the arrows, are a combined result of a higher interest rate and a lower wage rate. Our numerical results suggest that the effect via the interest rate on the optimal contract dominates.

as banks hold more liquid assets generates important implications on firm dynamics. We will discuss this in detail in Section 4.3.

4.3 Reallocation of Credit Among Firms and Firm Dynamics

In this section, we show how credit contraction affects the allocation of credit among firms and firm dynamics in equilibrium. We compare the average firm size, the growth rate of firm size, and the variance of firm growth for $\mu = 0$ and $\mu = 40\%$ in Figure 8. The three variables are as defined in Section 3.3.

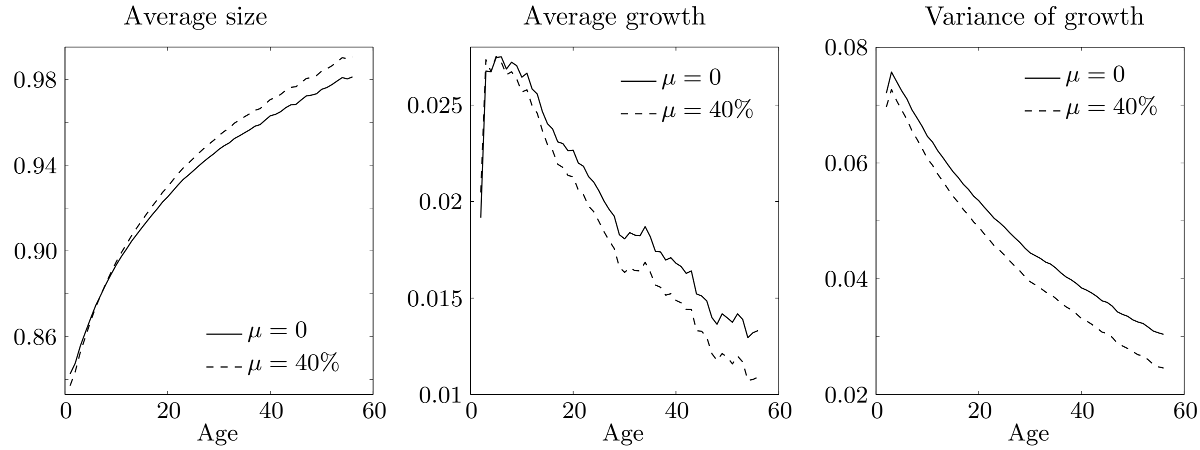


Figure 8: Effects of a credit contraction on firm dynamics

Note: The definitions of the variables and the solid curves for $\mu = 0$ are the same as in Figure 5. The dashed curves correspond to an economy where banks hold 40% of the deposits as reserves. All three variables illustrated are the 5-year moving average.

Our quantitative results suggest a reallocation of bank credit from young firms towards old ones as banks hoard more liquidity. Despite the tightening of aggregate credit, lending to old firms has increased, as is illustrated by the counterclockwise rotation of the average-size plot.

Intuitively, there are two counteracting forces at work. On the one hand, a change in factor prices decreases bank profits due to the higher costs of firm production. This lowers the initial promised value (determined by the zero-profit condition) to newborn firms and the optimal levels of bank loans to all firms. As a result, firms shrink in size, and the curve shifts down.[34] On the other hand, as the interest rate increases, banks discount future cash flows more heavily and thus postpone fulfilling the promised values in the future. Therefore, even though firms start with a lower promised value and are offered lower levels of loans, they catch up in size over time as banks adjust the intertemporal allocation of loans to firms through the changes in the promised values. The latter mechanism generates the steeper slope of the curve.

Noticeably, the decrease in the level of bank loans to young firms is less than the increase in loans to old firms on average. Nevertheless, the aggregate bank loans decrease by 0.7% as μ increases from 0 to 40% (Figure 9). This is mainly due to a larger number of young firms than the old ones in the economy. Specifically, firms below 10 years old, where the largest decline in bank loans occurs, account for 60% of all firms. While those above 20 years old, which see a

[34] According to the lending contracts, the optimal level of bank loans, $b = b(V^E; r, w)$, is a function of the promised value, V^E, and the factor prices, r and w. New entrants are affected through two margins, $b(.; r, w)$ and $V_0^E(r, w)$, whereas incumbent firms are affected through $b(.; r, w)$. Specifically, for incumbent firms, whose promised values are determined, liquidity hoarding increases the interest rate and lowers the optimal bank loans they receive. However, for new entrants, their initial promised value, $V_0^E(r, w)$ is also negatively affected by the higher interest rate. Since the level of bank loans is positively related to the promised values, this generates an additional negative effect on the size of new entrants.

stronger increase in bank loans, account for less than 20%.

Firm's growth rate, g_τ, at age τ is defined as the percentage increase in the level of bank loans from the last period, $g_\tau \equiv \frac{b_\tau - b_{\tau-1}}{b_{\tau-1}}$. As the aggregate credit contracts, young firms grow slightly faster and old firms slower on average as is illustrated in the average-growth plot in Figure 8. This is primarily driven by a base-year effect through a change in $b_{\tau-1}$. Specifically, due to the reallocation of credit, young firms are smaller and old firms are larger in an economy with less aggregate credit. This automatically generates opposite effects on the growth rates of the two groups of firms holding other factors constant. In addition, banks postpone fulfilling the promised values to firms into the future due to an increase in the interest rate. This positively affects the growth rate of all firms. Our numerical results suggest that the magnitude of the positive effect is smaller than the base-year effect and hence creates the illustrated pattern.[35]

Lastly, the variance of firm growth decreases uniformly for all firms as μ increases. This is mainly due to a decrease in the efficient level of bank loans, which makes the distribution of firm size less dispersed.

To sum up, as banks increase the holding of liquidity, the allocation of credit between young and old firms in the economy becomes more uneven. Young firms are smaller but grow faster and the opposite happens for old firms due to a strong base-year effect. Firm growth is more stable (i.e., the variance of firm growth is smaller) because the firm size distribution is concentrated on a smaller range.

4.4 Decomposing the General Equilibrium Effects

Lastly, we study the long-run effects of credit contraction on the aggregate variables in this section. As banks hold more liquid assets, aggregate bank loans decrease, which constrains firms' demand for capital and labor. The decline in the employment of input factors drives down the aggregate output. The results are summarized in Figure 9 by the black solid lines.

Note that the aggregate variables are determined by the number of firms (i.e., the extensive margin) and the optimal decisions of individual agents in the economy (i.e., the intensive mar-

[35] To understand intuitively why this is the case, consider a firm that has a promised value of $V_{\tau-1}$ and bank loans of $b_{\tau-1} = b^*(V_{\tau-1};\mu)$ at age $\tau-1$, where $b^*(V^E;\mu)$ is the optimal level of bank loans to a firm with promise value V^E in an economy with a reserve ratio μ. The level of bank loans, b_τ, that the firm receives in the next period depends on the transition function of the state variable, $V_s(V^E;\mu)$, and $b^*(V^E;\mu)$. V_s increases with μ because banks are incentivized to increase future promised values when facing a higher interest rate. This positively affects firm growth as mentioned in the text. However, the effect is largely dampened by the contraction in the optimal levels of bank loans, $b^*(V^E;\mu)$, which decreases with μ. Therefore, the overall effect of bank liquidity hoarding on the transition from $b_{\tau-1}$ to b_τ is relatively small and the resulting pattern in Figure 8 is mainly due to the base-year effect.

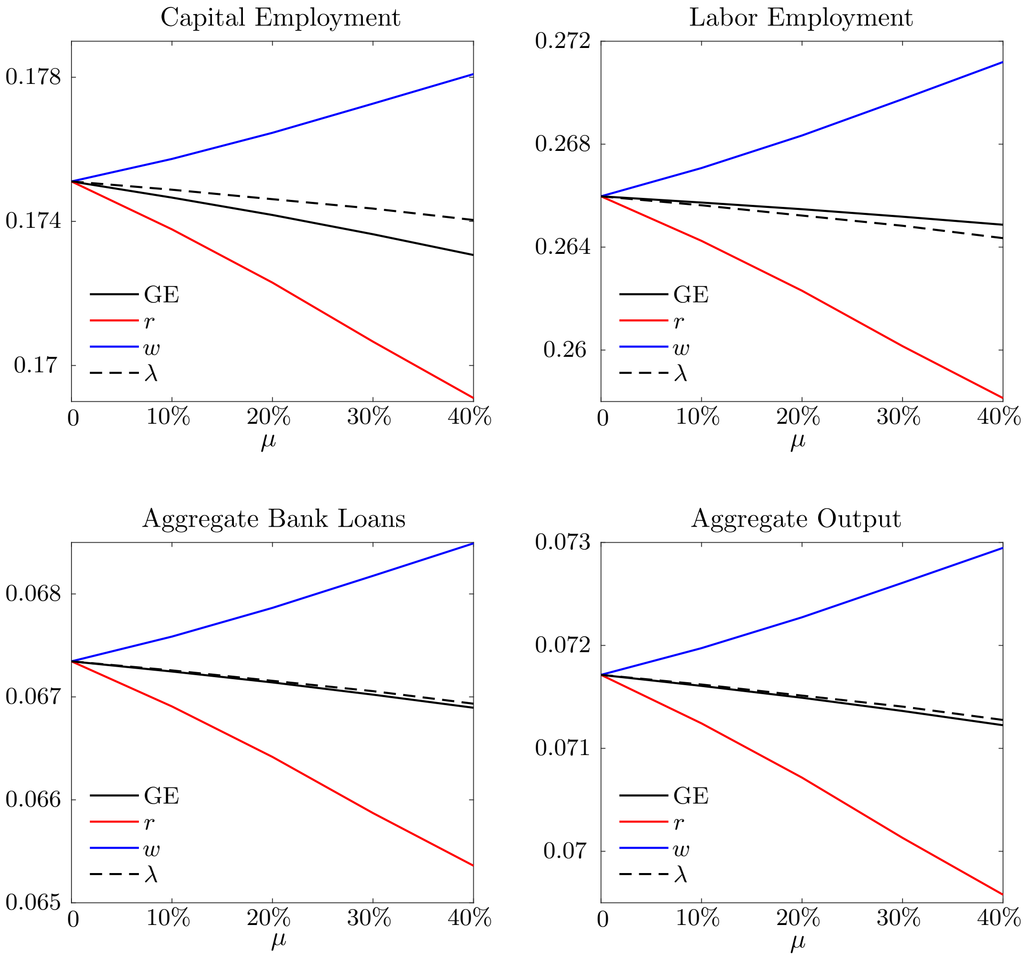


Figure 9: Decomposing the general equilibrium effects on aggregate variables

Note: The subplots illustrate the values of the aggregate variables in stationary general equilibrium (black lines) for $\mu \in [0, 40\%]$. The other three curves are derived by letting only one of the variables, the interest rate (red), the wage rate (blue), and firm entry (black dash) take its value in the equilibrium at different μ, and keeping the other two at their equilibrium values for $\mu = 0$.

gin), whose behavior depends on the equilibrium factor prices. Therefore, to better understand the mechanism that drives the changes in the aggregate variables, we further decompose the general equilibrium effects into individual margins via the interest rate, the wage rate, and the firm entry.

We solve for the equilibrium factor prices, $\{r^*(\mu), w^*(\mu)\}$, and the firm entry, $\lambda^*(\mu)$, at any reserve ratio $\mu \in [0, 40\%]$. Then, we calculate the values of the aggregate variables by allowing only one of the three variables to change with μ. For example, to isolate the effect of the credit contraction on the economy via the change in the interest rate (i.e., the interest-rate margin), we let the interest rate take the equilibrium value, $r^*(\mu)$, at any μ, and hold the wage rate and firm entry constant at their equilibrium values, $\{w^*(0), \lambda^*(0)\}$, at $\mu = 0$.[36] The results are shown in Figure 9. The dashed line represents the effect through the firm-entry margin. The gap between the black and the dashed lines comes from the factor-price margin, which is the combined effect of the wage margin (blue line) and the interest-rate margin (red line).

The decline in aggregate bank loans comes from a combined effect of a decline in firm entry and a lower level of loans to firms on average. Noticeably, the firm-entry channel (dashed line) contributes to almost all of the decline in aggregate loans (black line), with the rest coming from the net effect of the factor prices. Specifically, a higher interest rate decreases the amount of loans that banks provide to firms in equilibrium. But the effect is largely mitigated by a lower labor cost. Overall, the net effect of the two contributes to a small fraction of the decline in aggregate loans. A similar pattern emerges for the aggregate output. These results also highlight the importance of endogenizing firm entry and both factor prices in determining the magnitude of the effects of a credit contraction.

The decline in capital and labor employment is primarily driven by the lower level of bank loans. However, the rise in the relative capital costs shifts firms' demand for input factors towards more labor. This can be seen in the decomposition in Figure 9 as well. For capital employment, the dashed curve lies above the black line and the gap between the two is wider than that in the bank-loan plot. This indicates that in addition to the lower number of firms and lower average loans, aggregate capital demand decreases because firms move away from the use of capital at the higher interest rate. In contrast, for the labor-employment plot, the dashed line lies below the black line, implying that without considering the decline in firm entry, labor employment will even increase in equilibrium as firms choose a lower capital-labor ratio.

We notice from the numerical results that the long-run effects of credit contraction on the macroeconomy are very small. Specifically, an increase in the reserve ratio from zero to 40% decreases firm entry by 0.6%, capital and labor employment by 1.1% and 0.4%, respectively, and the aggregate output by 0.7%. The small effects are mainly due to an endogenous increase in the size of bank balance sheets. Following the initial increase in the interest rate, more households become workers and each worker saves more because of the substitution effect of the higher

[36] Alternatively, we could solve for *new equilibrium* factor prices and λ when shutting down one of the channels. However, the only proper analysis we can do is to shut down the endogenous firm entry channel. That is, we assume that the share of entrepreneurs, λ, is fixed at the initial equilibrium at $\mu = 0$ and solve for new combinations of $\{r, w\}$ that clear the capital and the labor markets. Banks still make zero profits from the financial contracts, but the lifetime utility of entrepreneurs and workers is no longer equalized. Unfortunately, we cannot fix only one factor price. Banks' zero-profit condition and households' indifferent occupational decisions will pin down the other factor price according to the equilibrium condition defined in (15). Therefore, we need to fix both r and w, and let λ be determined by one of the market-clearing conditions. But this is a trivial exercise as fixing r and w means that the optimal contracts and workers' optimization problems are unaffected. λ affects the aggregate variables only through the extensive margin.

interest rate on their saving decisions, both of which increase aggregate deposits.[37] Besides, bank equities converge to a higher level in the new stationary equilibrium. In the calibrated economy, the aggregate deposits and the equities increase by 21% and 41.4%, respectively. Therefore, even though a fraction of bank assets are held as reserves, the increase in the size of the assets compensates for the potential decline in credit supply to firms and thus relaxes the tightening of the capital markets. Overall, the higher interest rate together with the lower amount of capital in equilibrium implies a stronger contraction in capital supply relative to capital demand. Whereas the lower wage rate and the lower labor employment imply a stronger contraction in labor demand by firms.

5 Concluding Remarks

This paper studies the long-run effects of a credit contraction on firm dynamics and the macroeconomy. We construct a general-equilibrium model with endogenous firm entry and dynamic lending relationships between banks and firms under asymmetric information. Due to the informational friction, endogenous borrowing constraints emerge as a feature of the optimal lending contracts. The credit contraction is modeled as an exogenous liquidity shock on banks.

We show that the liquidity shock does not affect the optimal contracts directly, but indirectly through changes in the equilibrium factor prices. On the one hand, banks shrink lending to firms due to a deterioration of firm profitability. On the other hand, changes in the factor prices induce banks to provide better contract terms to firms that have been in the lending relationship for a longer period. This suggests that older firms are less affected in general, and when the second channel is strong enough as in our calibrated economy, there may even be an increase in the credit towards old firms, despite a decline in aggregate lending. This result implies that financial support that targets young firms and new entrants might be more effective than a beneficial-for-all policy during financial recessions.

Moreover, our model suggests that when implementing a general-equilibrium framework, the effects of credit contraction on the macroeconomy can be limited in the long run. The key mechanism is an endogenous increase in the size of bank balance sheets. Due to a rise in the interest rate, both deposits and bank equities increase, which compensates for the potential decline in the credit supply to firms. Meanwhile, the lower firm entry decreases the aggregate capital demand via the extensive margin and mitigates the credit constraints. As a result, there are only small changes in the factor prices and the aggregate variables in the new equilibrium.

[37] An increase in the interest rate has both an income and substitution effect on workers' saving decisions. Our numerical results show that the substitution effect dominates the income effect for workers with any levels of previous savings. Hence, aggregate deposits increase with the interest rate.

Lastly, by decomposing the effects on the aggregate variables, we show that a decline in firm entry contributes to a significant share of the negative effects on the economy. Consistent with existing studies on the role of firm entry over business cycles, this finding highlights the importance of endogenizing firm entry when analyzing the macroeconomic consequences of credit shocks in the long run.

Potentially, our model can be applied to study the macroeconomic effects of other shocks or regulations imposed on banks' balance sheets as well. We provide an example of how the model can be modified to analyze the effects of capital requirements on firm dynamics and the macroeconomic variables in Appendix G. Besides, we assume that banks hold an exogenous reserve ratio. It would be interesting to let banks optimally decide the share of liquid assets given external shocks and compare the model predictions with those in this paper.

A Timing

This section describes the timing of the events within one period in the economy.

1. At the beginning of a period, a mass $1-\Delta$ of households is born. Each newborn household makes a lifetime irreversible occupational decision to become an entrepreneur or a worker. The new entrepreneurs sign a lifetime lending contract with banks.
2. The banks provide loans to all entrepreneurs in the economy. The amount is determined by the respective terms of the contract (defined in Section 2.3.1). Banks pay workers gross return of deposits from the last period.
3. Firms produce output under uncertainty, using capital from banks and labor from workers as inputs. Capital rents and wages are paid with the loans.
4. After production, entrepreneurs observe their productivity realization and report it to the bank. Then, entrepreneurs make state-contingent repayments to the banks and fully consume the residual revenues from production. Workers consume and save for the next period with their labor income and capital returns.
5. A share $1-\Delta$ of the workers and entrepreneurs dies and the associated firms exit the market.

B An Economy without Financial Frictions

In an economy without financial frictions, the worker's problem is identical to that in Section 2.2.1. The optimal contracts are now defined by the following program:

$$P(V^E; r, w) = \max_{\{b, c_s, V_s^E\}_{s\in\{h,l\}}} -b + \sum_{s\in\{h,l\}} \pi_s \left[\theta_s R(b; r, w) - c_s + \frac{\Delta}{1+r} P(V_s^E; r, w)\right],$$

s.t.,

$$V^E = \sum_{s\in\{h,l\}} \pi_s [U\left(c_s, L^E\right) + \beta\Delta V_s^E], \quad c_s \geq 0. \tag{PK}$$

Let φ be the Lagrangian multiplier for (PK). We get[38]

$$\begin{aligned} \mathcal{L} = & \max_{\{b, c_s, V_s^E\}_{s\in\mathcal{S}}} -b + \sum_{s=\{l,h\}} \pi_s \left[\theta_s R(b) - c_s + \frac{\Delta}{1+r} P(V_s^E)\right] \\ & + \varphi \left\{ \sum_{s=\{l,h\}} \pi_s [U(c_s) + \beta\Delta V_s^E] - V^E \right\} \end{aligned}$$

[38]For simplicity of notation, we ignore $\{r, w\}$ from the output, $R(.)$, and bank profits, $P(.)$, and L^E from the utility function, $U(.)$.

The first-order conditions are thus

$$\frac{\partial \mathcal{L}}{\partial b} = -1 + \mathbb{E}(\theta_s)R'(b) = 0, \tag{B.1}$$

$$\frac{\partial \mathcal{L}}{\partial c_s} = -\pi_s + \varphi \pi_s U'(c_s) = 0, \tag{B.2}$$

$$\frac{\partial \mathcal{L}}{\partial V_s} = \frac{\Delta \pi_s}{1+r} P'(V_s) + \varphi \pi_s \beta \Delta = 0. \tag{B.3}$$

The Envelope theorem gives

$$P'(V) = -\varphi. \tag{B.4}$$

Equation (B.1) suggests that the optimal level of bank loans in an economy without financial frictions is the efficient level defined in (7). Besides, the optimal contracts provide complete insurance for entrepreneurs (i.e., c_s and V_s are constant across states), according to equations (B.2) and (B.3). In particular,

$$\varphi U'(c_s) = 1, \tag{B.5}$$

$$P'(V_s) = -\varphi \beta (1+r). \tag{B.6}$$

Denote c_s and V_s as c and V', respectively, and consumption tomorrow as c'. According to (B.4), (B.5), and (B.6), we get

$$\frac{U'(c)}{U'(c')} = \beta(1+r). \tag{B.7}$$

Given the explicit utility function in Section 3, we get from (B.7), $c' = c + \frac{1}{\gamma} \ln \beta(1+r)$.[39] This implies that the dynamics of c_t is governed by

$$c_t = c_0 + \frac{\ln(\beta(1+r))}{\gamma} t. \tag{B.8}$$

The equilibrium conditions are identical to those in Section 2.4.1. In particular, the lifetime utility of an entrepreneur equals to that of a worker

$$V_0^W(r,w) = \sum_{t=0}^{\infty} (\beta \Delta)^t U(c_t, L^E), \tag{B.9}$$

Plugging this into (B.9) we determine the initial consumption of an entrepreneur:

$$c_0 = -\frac{1}{\gamma} \ln \left[\left(-V_0^W(r,w) - \frac{\eta L_E^2}{1-\beta\Delta} \right) \left(1 - \frac{\Delta}{1+r} \right) \right] \tag{B.10}$$

All firms get the same level of bank loans. Hence, the aggregate bank loans and output are simply $B = \lambda b^*$ and $Y = \mathbb{E}(\theta_s)R(b^*)$, respectively. The aggregate capital and labor demand are now given by:

$$K^D = \frac{\lambda}{r+\delta} \frac{\alpha_k}{\alpha_k + \alpha_l} b^* \quad \text{and} \quad L^D = \frac{\lambda}{w} \frac{\alpha_l}{\alpha_k + \alpha_l} b^*.$$

[39]Since $U(c,l) = -\exp(-\gamma c) - \eta l^2$, $U'(c) = \gamma \exp(-\gamma c)$. This suggests that $\exp(\gamma(c'-c)) = \beta(1+r)$.

Repayments in state s of period t is $m_t^s = \theta_s R(b^*) - c_t$. The total repayments are thus

$$M = \lambda \sum_{t=0}^{\infty} (1-\Delta)\Delta^t \left[\mathbb{E}(\theta_s) R(b^*) - c_t\right] = \lambda \left[\mathbb{E}(\theta_s) R(b^*) - (1-\Delta) \sum_{t=0}^{\infty} \Delta^t c_t\right]. \tag{B.11}$$

Plugging (B.8) in (B.11), we get an explicit expression for M:

$$M = \lambda \left[\mathbb{E}(\theta_s) R(b^*) - c_0 - \frac{\ln(\beta(1+r))}{\gamma} \frac{\Delta}{1-\Delta}\right]. \tag{B.12}$$

The aggregate variables from the workers' side are the same as before. We can thus have the capital- and labor-market-clearing conditions.

C Proofs of the Properties of the Optimal Contract

Following the extensive literature on dynamic contracts under asymmetric information (e.g., Thomas and Worrall, 1990), the incentive constraints (IC) can be simplified to include only the high-productivity specification. In other words, low-productivity firms have no incentive to report a high productivity realization, and hence their incentive constraint never binds. This property is summarized in Lemma 1 and is applied to prove the propositions in Section 2.3.3.

Define the incentive constraints for all $i, j \in \mathcal{S}$ as:

$$C_{i,j} \equiv U(\theta_i R(b) - m_i, L^E) + \beta \Delta V_i^E - U(\theta_i R(b) - m_j, L^E) - \beta \Delta V_j^E \geq 0, \tag{C.1}$$

where i is the actual state and j is the reported state.

Lemma 1. *For strictly concave $P(V^E)$, for all states $s \in \mathcal{S}$, the optimal contract implies that the local downward constraints $C_{s,s-1} \geq 0$ always bind, whereas the local upward constraints $C_{s-1,s} \geq 0$ never bind for $m_s > m_{s-1}$.*[40]

Proof of Lemma 1. First, we prove by contradiction that the local downward constraints must bind. Suppose that there exists an optimal contract $\left\{b, m_s, V_s^E\right\}_{s\in\mathcal{S}}$ such that for some $i \in \mathcal{S}$ the downward constraint does not bind (i.e., $C_{i,i-1} > 0$). Then, the general procedure is as follows: We prove that there exists a mean-preserving contraction transformation on $\left\{V_j^E\right\}_{j=i,\ldots,S}$ such that the new contract $\left\{b, m_s, \hat{V}_s^E\right\}_{s\in\mathcal{S}}$, where $\hat{V}_j^E = V_j^E$, for $j = 1, 2, \ldots, i-1$, fulfills all constraints. In particular, we make a transformation with $\sum_{s\in\mathcal{S}} \pi_s \hat{V}_s^E = \sum_{s\in\mathcal{S}} \pi_s V_s^E$, and $\hat{V}_j^E - \hat{V}_l^E \leq V_j^E - V_l^E$, $\forall j, l \in \mathcal{S}$, with at least one pair of $\{j, l\}$ giving strict inequality. In this case, under the assumption that $P(V^E)$ is strictly concave, the banks' profit increases strictly with the new contract. This contradicts the fact that $\left\{b, m_s, V_s^E\right\}_{s\in\mathcal{S}}$ is an optimal contract.

[40] The lemma generally holds for any finite number of states. The proof follows the standard formulation of the optimal social insurance in Ljungqvist and Sargent (2018), which is based on Thomas and Worrall (1990).

Now we describe explicitly the procedure of performing a mean-preserving contraction transformation on the contract: Keeping $\{m_{i-1}, m_i, V_{i-1}^E\}$ as before, we decrease V_i^E until $C_{i,i-1} = 0$. Since changing V_i^E will influence the local downward incentive constraints for $s = i + 1$ and sequentially $s = i + 2, \ldots, S$, we decrease for each $s = i + 1, \ldots, S$, V_s^E such that $C_{s,s-1} = 0$. As a result we have a new sequence of future promised value $\left\{V_s^{E\prime}\right\}_{s\in\mathcal{S}} = \left\{V_1^E, \ldots, V_{i-1}^E, V_i^{E\prime}, {V_{i+1}^E}', \ldots, V_S^{E\prime}\right\}$. Now we add a positive constant, $\bar{v}$, to the sequence of future promised value, such that the promise keeping constraint is regained. Let $\hat{V}_s^E = V_s^{E\prime} + \bar{v}$. We have a new contract $\left\{b, m_s, \hat{V}_s^E\right\}_{s\in\mathcal{S}}$.

Note that the new contract fulfills the local upward constraints automatically given the strict concavity of the utility function and the fact that $C_{s,s-1} = 0\ \forall s \in \mathcal{S}$ (see argumentation in the last part of this proof). In addition, the promise keeping constraint is still fulfilled due to the mean-preserving transformation, and the limited liability constraints are uninfluenced since b and $\{m_s\}_{s\in\mathcal{S}}$ are unchanged. Finally, for any $j = i, \ldots, S$, V_{j+1}^E must decrease at least as much as V_j^E to guarantee that $C_{j+1,j} = 0$. Therefore, for any $j = i, \ldots, S$, $\bar{v} \leq V_j^E - V_j^{E\prime}$, indicating that $\hat{V}_j^E \leq V_j^E$ and remember that for $j = 1, 2, \ldots, i-1$ $\hat{V}_j^E = V_j^E$,. Since $\{V_s^E\}_{s\in\mathcal{S}}$ fulfills the credibility constraints, so does the new contract.

Further, notice from the procedure that the gap of the promised values between two successive states, s and $s-1$, is either unchanged or decreased, with a definite decrease in $\hat{V}_i^E - \hat{V}_{i-1}^E$. Following this we know $\forall j, l \in \mathcal{S}$, $V_j^E - V_l^E$ is non-increasing. Thus, the new contract is a mean-preserving contraction. This contradicts that $\{b, m_s, V_s^E\}_{s\in\mathcal{S}}$ is an optimal contract. We know that the local downward constraints always bind.

Given that $C_{s,s-1} = 0, \forall s \in \mathcal{S}$, rewriting the constraint we have

$$\beta\Delta(V_s^E - V_{s-1}^E) = U(\theta_s R(b) - m_{s-1}, L^E) - U(\theta_s R(b) - m_s, L^E)$$

Since $\theta_{s-1} < \theta_s, m_{s-1} \leq m_s$ and the utility function is strictly concave, we have

$$\begin{aligned}U(\theta_{s-1}R(b) - m_{s-1}, L^E) - U(\theta_{s-1}R(b) - m_s, L^E) \geq \\ U(\theta_s R(b) - m_{s-1}, L^E) - U(\theta_s R(b) - m_s, L^E) =& \beta\Delta(V_s^E - V_{s-1}^E),\end{aligned}$$

where strict inequality holds for $m_{s-1} < m_s$. Therefore, we have directly from this that the local upward constraint is never binding. Namely, $C_{s-1,s} > 0, \forall s \in \mathcal{S}$.

To prove Proposition 1 - 4, we set up the Lagrangian of the banks' problem defined in Section 2.3.2. Let λ_1 and λ_2 be the Lagrangian multiplier for (PK) and for (IC), respectively, under a high productivity realization. We ignore (IC) under the low state, because it never binds according to Lemma 1. Further, we substitute the repayments using $m_s = \theta_s R(b; r, w) - c_s$,

$s \in \{h, l\}$. This gives

$$\mathcal{L} = \max_{\{b, c_s, V_s^E\}_{s\in\mathcal{S}}} -b + \sum_{s=\{l,h\}} \pi_s \left[\theta_s R(b; r, w) - c_s + \frac{\Delta}{1+r} P(V_s^E)\right]$$
$$+ \lambda_1 \left\{ \sum_{s=\{l,h\}} \pi_s [U(c_s, L^E) + \beta \Delta V_s^E] - V^E \right\}$$
$$+ \lambda_2 \left\{ U(c_h, L^E) + \beta \Delta V_h^E - U((\theta_h - \theta_l) R(b; r, w) + c_l, L^E) - \beta \Delta V_l^E \right\}.$$

The first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial c_h} = -\pi + (\lambda_1 \pi + \lambda_2) U'(c_h) = 0; \tag{C.2}$$
$$\frac{\partial \mathcal{L}}{\partial c_l} = -(1-\pi) + \lambda_1 (1-\pi) U'(c_l) - \lambda_2 U'(c_{hl}) = 0; \tag{C.3}$$
$$\frac{\partial \mathcal{L}}{\partial V_h^E} = \frac{\pi}{1+r} P'(V_h^E) + (\lambda_1 \pi + \lambda_2)\beta = 0; \tag{C.4}$$
$$\frac{\partial \mathcal{L}}{\partial V_l^E} = \frac{1-\pi}{1+r} P'(V_l^E) + (\lambda_1 (1-\pi) - \lambda_2)\beta = 0; \tag{C.5}$$
$$\frac{\partial \mathcal{L}}{\partial b} = -1 + \mathbb{E}(\theta) R'(b; r, w) - \lambda_2 U'(c_{hl})(\theta_h - \theta_l) R'(b; r, w) = 0. \tag{C.6}$$

The Envelope theorem:

$$P'(V^E) = -\lambda_1. \tag{C.7}$$

Equation (C.2) - (C.7) are conditions that the optimal dynamic contract must fulfill. Now we use these conditions to prove the properties of the optimal contracts. Throughout the section, we assume that the value function $P(V^E)$ decreases in the promised value and is strictly concave. This implies that $P'(V^E) < 0$ and $\lambda_1 > 0$.[41]

C.1 Proof of Proposition 1

Using the definition of $C_{i,j}$ in (C.1), we conclude: $C_{s,s-1} + C_{s-1,s} \geq 0$, which is equivalent to

$$U(\theta_s R(b) - m_s, L^E) - U(\theta_s R(b) - m_{s-1}, L^E) \geq U(\theta_{s-1} R(b) - m_s, L^E) - U(\theta_{s-1} R(b) - m_{s-1}, L^E) \tag{C.8}$$

[41] Within a very small range of V_{min}, the value function $P(V^E)$ increases in the promised value (see Figure 3). To see the reason, notice that at $V^E = V_{min}$, the optimal level of bank loans is zero. According to the limited liability constraint (LL), this implies that the maximal level of repayments banks can ask for is zero. So, bank profits are zero. As V^E increases, the size of bank loans increases and so does the level of repayments. Therefore, banks start to make positive profits from the lending contract. This is the region where the value function increases in V^E. As V^E further rises, we enter the standard region, where banks' profits decrease because they promise entrepreneurs higher utility, and the force described above no longer dominates.

Since $\theta_s > \theta_{s-1}$ and given the strict concavity of the utility function in consumption, (C.8) is satisfied only if $m_s \geq m_{s-1}$. It then follows from $C_{s,s-1} \geq 0$ that $V_s^E \geq V_{s-1}^E$.

Further, we prove that $c_h(V^E) > c_l(V^E)$. Multiplying (C.3) and (C.4) with $1-\pi$ and π, respectively, we have

$$-\pi(1-\pi) + (1-\pi)(\lambda_1\pi + \lambda_2)U'(c_h) = 0, \tag{C.9}$$

$$-\pi(1-\pi) + \lambda_1\pi(1-\pi)U'(c_l) - \lambda_2\pi U'(c_{hl}) = 0 \tag{C.10}$$

Deducting (C.10) from (C.9) gives us

$$\lambda_1\pi(1-\pi)\left[U'(c_h) - U'(c_l)\right] = -\lambda_2\left[(1-\pi)U'(c_h) + \pi U'(c_{hl})\right].$$

Since $\lambda_1 \geq 0$, $\lambda_2 \geq 0$, and $U'(.) > 0$, we have $U'(c_h) \leq U'(c_l)$. Given a concave utility function, this implies that $c_h \geq c_l$, where the inequality holds when (IC) is binding (in which case $m_h = m_l$ and $V_h = V_l$).

C.2 Proof of Proposition 2

According to (C.6),

$$\mathbb{E}(\theta)R'(b;r,w) - \lambda_2 U'(c_{hl})(\theta_h - \theta_l)R'(b;r,w) = 1. \tag{C.11}$$

Since $\lambda_2 \geq 0$ and $U'(c) \geq 0$, $\mathbb{E}(\theta)R'(b) \geq 1$. Since the efficient level of bank loans defined in (7) satisfies $\mathbb{E}(\theta)R'(b^*) = 1$ and $R(.)$ is a concave function, we have $b \leq b^*$.

C.3 Proposition 4 and Proof

The interest rate is an important factor that determines the optimal contracts and firm dynamics.

Proposition 4. *If the first-order derivative of the value function, $P'(.)$ is non-concave, firm size grows on average (i.e., $V \leq \mathbb{E}(V_s^E)$) when $\beta(1+r) \geq 1$. If $P'(.)$ is concave, firm size shrinks on average (i.e., $V \geq \mathbb{E}(V_s^E)$) when $\beta(1+r) \leq 1$. Moreover, the promised utility grows faster overtime.*

Proof. According to (C.4), (C.5), and the Envelope theorem, we have

$$\frac{1}{1+r}\sum_{s=\{l,h\}}\pi_s P'(V_s^E) = \beta P'(V).$$

Namely,

$$\mathbb{E}P'(V_s^E) = \beta(1+r)P'(V).$$

If $P'(.)$ is non-concave (i.e., linear or convex), from Jensen's Inequality, we know that $\mathbb{E}P'(V_s^E) \geq P'(\mathbb{E}(V_s^E))$. Therefore,

$$\beta(1+r)P'(V) \geq P'(\mathbb{E}(V_s^E)).$$

When $\beta(1+r) \geq 1$, $P'(V) \geq P'(\mathbb{E}(V_s^E))$. Since $P(.)$ is concave, $V \leq \mathbb{E}(V_s^E)$.

Similarly, if $P'(.)$ is concave, $\mathbb{E}P'(V_s^E) \leq P'(\mathbb{E}(V_s^E))$. Therefore,

$$\beta(1+r)P'(V) \leq P'(\mathbb{E}(V_s^E)).$$

When $\beta(1+r) \leq 1$, $P'(V) \leq P'(\mathbb{E}(V_s^E))$. Since $P(.)$ is concave, $V \geq \mathbb{E}(V_s^E)$.

In addition, as r increases, for the same V^E, $\mathbb{E}P'(V_s^E)$ decreases. Therefore, V_h^E and V_l^E are larger. Namely, the growth rate in promised utility is larger.

Intuitively, given the promised values today, banks optimally allocate entrepreneurs' consumption today and in the future via future promised values according to the (PK) constraint. When the interest rate is higher, banks discount future cash flows more heavily. This makes it is more cost-efficient for banks to postpone fulfilling the promised value with higher future promised values. Therefore, higher interest rates imply an increase in the promised values and firm size over time.

C.4 Proof of the Existence of a Stationary Distribution

In this section, we prove that there exists a stationary distribution of households, $\psi^W(A)$, with respect to their today's deposit level, A, and a stationary distribution of firms, $\psi^E(V^E)$, with respect to their today's promised value, V^E. We start with households.

Given the households' saving decision (i.e., the policy function $A' = g(A; r, w)$) and the exogenous shock on survival, the transition function is given by $Q^H(A, B) : [0, \bar{A}] \times \mathcal{B}([0, \bar{A}]) \to [0, 1]$, where $\mathcal{B}([0, \bar{A}])$ denotes the Borel set on $[0, \bar{A}]$, and $Q^H(A, B)$ indicates the probability of a household with deposit A to save $A' \in B$ for tomorrow. Formally, we write the transition function as

$$Q^H(A, B) = \begin{cases} \Delta, & B = \{A' = g(A; r, w)\}, \\ 1 - \Delta, & B = \{A' = 0\}, \\ 0, & \text{otherwise.} \end{cases} \tag{C.12}$$

We next prove for any given initial distribution of households' deposit, ψ_0, the transition function leads to a unique distribution of households' deposit, ψ^*. Notice that the transition function corresponds to an operator, T, that maps the set of probability distribution on $[0, \bar{A}]$, denoted by $\Lambda([0, \bar{A}], \mathcal{B}([0, \bar{A}]))$, into itself. Specifically, $(Tf)(A) = \int f(x)Q_H(A, dx)$, where $f(x) \in \Lambda([0, \bar{A}], \mathcal{B}([0, \bar{A}]))$. Therefore, formally we need to prove that there exists a unique distribution ψ^*, s.t. $\lim_{t \to +\infty} T^t \psi_0 = \psi^*$. We use Theorem 12.12 in Stokey et al. (1989), which states that if the transition function is monotone, has the Feller property, and satisfies a mixing condition, a unique invariant probability distribution exists. We show next that $Q^H(A, B)$ defined above satisfies these conditions.

First, we prove the monotonicity and the Feller property. A transition function is monotone if for any bounded increasing function f, Tf is also bounded increasing. A transition function satisfies the Feller property if for any bounded continuous function f, Tf is also bounded continuous. Notice that

$$(Tf)(A) = \int f(x) Q^H(A, dx) = \Delta f(g(A; r, w)) + (1 - \Delta) f(0).$$

Since $g(A; r, w) : [0, \bar{A}] \to [0, \bar{A}]$, is increasing and continuous in today's deposit A, $Tf(A)$ is increasing (continuous), if $f(x)$ is increasing (continuous).[42]

Now we are left to prove the transition function satisfies the mixing condition. Specifically, there exists $A^* \in [0, \bar{A}]$, $\epsilon > 0$, and $N \geq 1$ such that $(Q^H)^N(A, [A^*, \bar{A}]) \geq \epsilon$ and $(Q^H)^N(\bar{A}, [0, A^*]) \geq \epsilon$. Take $A^* = 0$ and $N = 1$. It is clear from (C.12) that $Q^H(A, [0, \bar{A}]) \geq 1 - \Delta$, and $Q^H(\bar{A}, \{0\}) = 1 - \Delta$. Therefore, the condition is satisfied if we let $\epsilon = 1 - \Delta$.

In the end, we characterize the stationary distribution, ψ^*. Since ψ^* is independent of ψ_0, we may start with an arbitrary ψ_0, defined by $\psi_0(A = 0) = 1$, $\psi_0(A \neq 0) = 0$. $\psi_t = T^t \psi_0$ is thus given by $\psi_t(A = g^i(A; r, w)) = (1 - \Delta)\Delta^i, i = 0, 1, \ldots, t - 1$, with $g^0(A; r, w) = 0$, $\psi^*(A = g^t(A; r, w)) = \Delta^t$. Therefore, $\psi^W(A; r, w) = \psi^* = \lim_{t \to +\infty} T^t \psi_0$.

So far we have proved that the distribution of households' savings converges to a unique stationary distribution regardless of the initial distribution. We prove the result of firms in an analogous way.

Given the future promised value of firms (i.e., the policy function $V'(V^E, \theta; r, w)$) and the exogenous shock on survival and productivity, the transition function is given by $T(V^E, B) : [V^E_{min}, V^E_{max}] \times \mathcal{B}([V^E_{min}, V^E_{max}]) \to [0, 1]$, where $T(V^E, B)$ indicates the probability of a firm with today's promised value V to get a promised value $V' \in B$ for tomorrow. Formally, we write the transition function as

$$T(V^E, B) = \Delta \sum_{s \in \mathcal{S}} \pi_s \mathbb{1}(V'(V^E, \theta_s; r, w) \in B) + (1 - \Delta) \mathbb{1}(V^E_0(r, w) \in B).$$

We need to prove that the transition function $T(V^E, B)$ is monotone, satisfies the Feller property and the mixing condition. Suppose L is the operator associated with $T(V^E, B)$. Given a function $f(x)$ (measurable on $([V^E_{min}, V^E_{max}], \mathcal{B}([V^E_{min}, V^E_{max}]))$, either bounded increasing or

[42] $g(A; r, w)$ increases only if $\Delta R > \beta$. We checked ex-ante that our numerical solution satisfies this condition.

bounded continuous),

$$(Lf)(V^E) = \int f(x)T^E(V^E, dx) = \Delta \sum_{s\in\mathcal{S}} \pi_s f(V'(V^E, \theta_s; r, w)) + (1-\Delta) f(V_0^E(r, w)).$$

Since the policy function $V'(V^E, \theta; r, w)$ is increasing and continuous in V^E, $(Lf)(V^E)$ is increasing (continuous), if $f(x)$ is increasing (continuous). Take $V^* = V_0^E(r, w)$, $N = 1$, and $\epsilon = 1 - \Delta$. Similarly to the argument for households, we have $T(V^E, B)$ satisfies the mixing condition.

Given the idiosyncratic shock on firm productivity each period and the exponential increment of the history of such shocks, an explicit characterization goes very messy very fast. Therefore, we simply denote the stationary distribution as $\psi^E(V^E; r, w)$.

C.5 Derivation of the Good-Market-Clearing Condition

In this section, we derive the goods-market-clearing condition

$$C^E + C^W + \delta K^D = Y + r_e \mu D, \tag{C.13}$$

and prove that it holds automatically when the labor and the capital markets clear.

The aggregate consumption of entrepreneurs is given by

$$C^E = \lambda \sum_{s\in\mathcal{S}} \pi_s \int \left[\theta_s R(b(V^E; r, w)) - m_s(V^E; r, w)\right] d\psi^E(V^E; r, w), \tag{C.14}$$

where the term in the bracket is the consumption of an entrepreneur with promised value V^E and current productivity shock, θ_s. Denote the aggregate output as

$$Y \equiv \lambda \sum_{s\in\mathcal{S}} \pi_s \int \theta_s R(b(V^E; r, w)) d\psi^E(V^E; r, w).$$

We reformulate (C.14) as

$$C^E = Y - M, \tag{C.15}$$

where the definition of aggregate repayments, M, in (22) is used.

To derive an expression for the aggregate consumption of workers, C^W, we start from the budget constraint in (2). Using the optimal saving, $g(A; r, w)$, and labor choice, $h(A; r, w)$, defined in (3), we have

$$C^W = (1-\lambda) \int \left[wh(A; r, w) + (1+r)A - \Delta g(A; r, w)\right] d\psi^W(A; r, w). \tag{C.16}$$

According to the demographic structure of workers, $\Delta\psi^W(A; r, w)$ represents the distribution of A among all workers excluding the newborns. Moreover, since the newborn workers have zero initial wealth, the aggregate savings are independent of whether we include the cohort in the calculation. Hence, $\int A d\psi^W(A; r, w) = \Delta \int g(A; r, w) d\psi^W(A; r, w)$ and (C.16) becomes

$$C^W = (1-\lambda) \int \left[wh(A; r, w) + rA\right] d\psi^W(A; r, w).$$

Using the definition of the aggregate labor supply, L^S, in (17), $L^S = L^D$, and the aggregate deposits, D, in (20), we have

$$C^W = wL^D + rD, \tag{C.17}$$

Plugging (C.15) and (C.17) in (C.13), we get

$$C^E + C^W + \delta K^D = Y - M + wL^D + rD + \delta K^D \tag{C.18}$$

Entrepreneurs use the bank loans as working capital to finance production costs. Thus, the constraint in entrepreneurs' decision problem in (4) is binding: $b(V^E) = wl^*(V^E) + (r+\delta)k^*(V^E)$. In aggregate, this means

$$B = wL^D + (r+\delta)K^D. \tag{C.19}$$

Substituting $wL^D + \delta K^D$ in (C.18) with (C.19), we get

$$C^E + C^W + \delta K^D = Y - M + B + r(D - K^D). \tag{C.20}$$

In a stationary equilibrium, bank equities remain constant with $E = E'$. According to (22), this gives

$$r_e \mu D = B - M + r(D - K^S).$$

Substituting this and $K^D = K^S$ in (C.20), we have the goods-market-clearing condition, (C.13). The equation implies that the goods market clears if the aggregate output, Y, and the interest payments on reserves equal the sum of consumption, $C^E + C^W$, by entrepreneurs and workers and aggregate investments, δK^D, which is the capital depreciation in stationary equilibrium. Without affecting the results of the paper, interest payments on reserves are assumed to be outside the model. However, we can keep it in the model by simply assuming that a lump-sum tax is imposed on workers and the proceeds are used to pay interests on the reserves.[43]

[43] This will change a worker's budget constraint to $c + \Delta A' = wl + (1+r)A - T$, where $T = r_e \mu D$ is the lump-sum tax on workers. Workers' optimization problem is unaffected as T is taken as given. In this case, (C.17) becomes $C^W = wL^D + rD - T$ and the goods-market-clearing condition, (C.13), becomes $C^E + C^W + \delta K^D = Y$.

D Optimization Problem: Workers and Entrepreneurs

D.1 Workers' Optimal Decisions

Figure 10 depicts, as a function of the current period deposit wealth A, the workers' optimal consumption $c(A)$, the labor supply $l(A)$ and the saving decision $A'(A)$ corresponding to the policy functions given in (3) and the lifetime expected utility $V^W(A)$ for given w and r.[44]

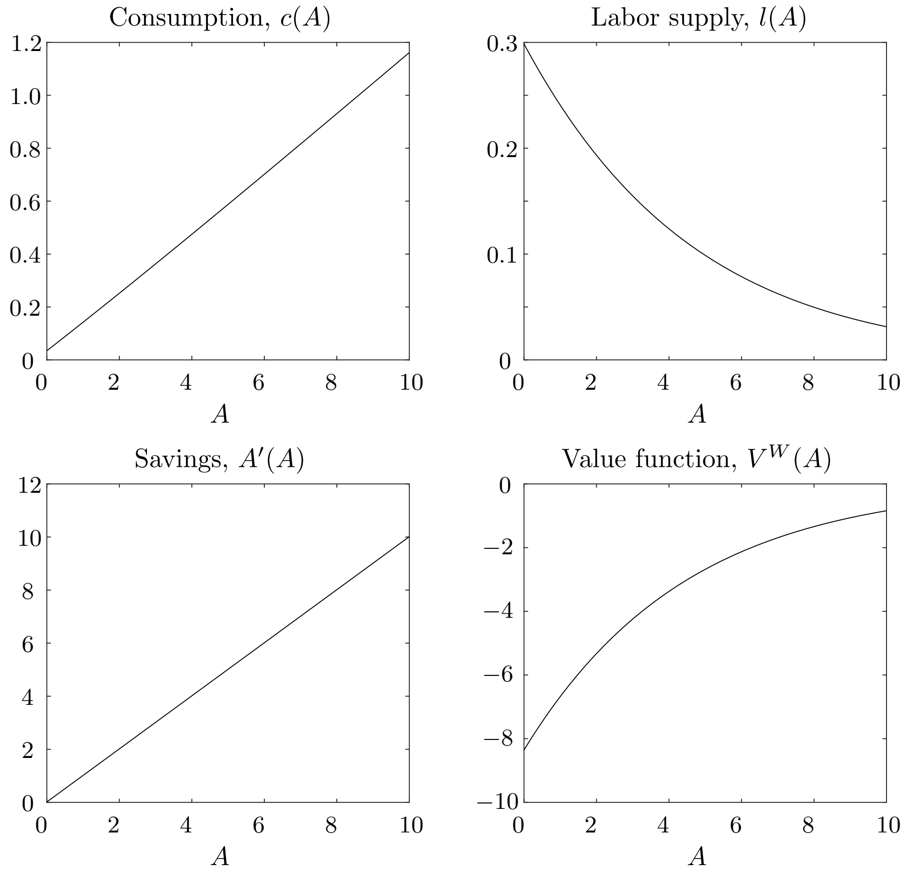


Figure 10: Solution to worker's problem

Because of the income effect, workers with higher deposits consume more today, save more for tomorrow, and supply less labor. Their lifetime expected utility, captured by the value function $V^W(A)$, is an increasing function in A, indicating that workers are better off if endowed with more wealth A. In addition, the envelop theorem implies that $A' > A$ if $\beta(1 + r) > 1$, which is the case in our numerical analysis. In other words, workers postpone their consumption and increase saving over time with a small future discount rate. Nevertheless, this does not mean that the aggregate deposit in the economy will go to infinity in equilibrium. Remember that in

[44]See Appendix E.1 for the procedure to solve the recursive workers' problem numerically.

a stationary equilibrium, there is a constant inflow of newborn workers and exit of the old ones, with the aggregate deposit defined in (20). As long as the growth rate of deposit is lower than the exit rate of the workers – as is the case in the quantitative analysis, the aggregate deposit is finite.[45]

D.2 Entrepreneurs' Optimal Capital and Labor Employment

For a given level of bank loans b and factor prices w and r, the entrepreneur's optimal capital input and labor employment based on the decision problem in (4) are linear in b (see plots in Figure 11):

$$k^* = \frac{1}{r+\delta}\frac{\alpha_k}{\alpha_k+\alpha_l}b \quad \text{and} \quad l^* = \frac{1}{w}\frac{\alpha_l}{\alpha_k+\alpha_l}b. \tag{D.1}$$

The corresponding firm output is thus

$$R(b) = \left(\frac{\alpha_k}{r+\delta}\right)^{\alpha_k}\left(\frac{\alpha_l}{w}\right)^{\alpha_l}\left(\frac{b}{\alpha_k+\alpha_l}\right)^{\alpha_k+\alpha_l}\bar{a}. \tag{D.2}$$

Following from equation (7), the efficient level of bank loans is thus given by

$$b^* = (\alpha_k+\alpha_l)\left[\bar{a}\left(\frac{\alpha_k}{r+\delta}\right)^{\alpha_k}\left(\frac{\alpha_l}{w}\right)^{\alpha_l}\right]^{\frac{1}{1-\alpha_k-\alpha_l}}. \tag{D.3}$$

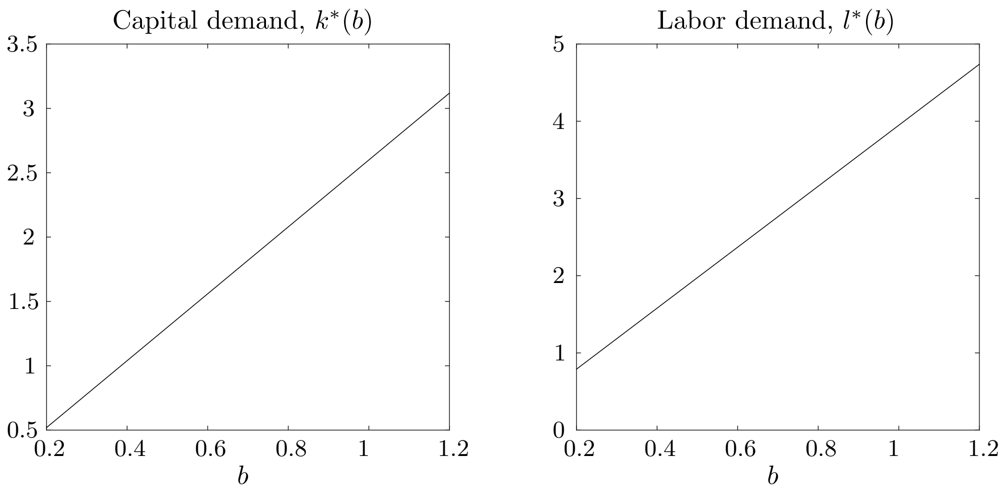


Figure 11: Solution to entrepreneur's problem

E Numerical Procedure

In this section, we describe the dynamic programming algorithm for solving the partial equilibrium (i.e., workers' decision problem in Section E.1 and banks' optimal contract in Section

[45]Technically, this is the sufficient condition for the convergence of the infinite sequence, $\left\{(1-\Delta)\Delta^\tau p^A g(A_\tau, r, w)\right\}_{\tau=0}^{\infty}$.

E.2), the procedure to simulate the entrepreneurs' life path and calculate the aggregate variables (Section E.3), and the algorithm to calculate the unique stationary general equilibrium (Section E.4). All computations are done with Matlab.

E.1 Workers

1. Use constant r and w.

2. Set a grid for the state variable A. $Agrid$ denotes the grid points of A. We set $A = [0, 10]$ and generate $nA = 50$ Chebyshev grid points on the interval. We manually replace the lowest Chebyshev point with the lower bound of $A = 0$.

3. Give an initial guess for the functional form of the value function, $V^W(A; r, w)^0$, of the policy functions $l(A)^0$ and $A'(A)^0$ and of $c(A)^0$.[46] We use $V^W(A_i)^0 = -\exp(-A_i) - 0.1$, $c(A_i)^0 = 0.1$, $l(A_i)^0 = 0.6$ and $A'(A_i)^0 = 0$ for each $A_i \in Agrid$, $i \in \{1, ..., nA\}$.

4. Solve on each grid point $A_i \in Agrid$, $i \in \{1, ..., nA\}$, the worker's problem in (1) subject to (2), $c(A_i) \geq 0$, $l(A_i) \in [0, 1]$ and $A'(A_i) \geq 0$. This gives us the optimal solution of the system, $\{c(A_i)^1, l(A_i)^1, A'(A_i)^1\}$ and the corresponding updated value function $V^W(A_i; r, w)^1$ at $A_i \in Agrid$, $i \in \{1, ..., nA\}$.[47] To calculate the updated value function, we interpolate on $V^W(A; r, w)^0$ to get values for $A'(A)$ which lie between two *Agrid*-points.[48]

5. Compare the two successive iterations of value functions, $V^W(A; r, w)^1$ with $V^W(A; r, w)^0$, by defining a distance measure d_{V^W}, such that

$$d_{V^W} \equiv \max_{i \in \{1,...,nA\}} \left| V^W(A_i; r, w)^1 - V^W(A_i; r, w)^0 \right|.$$

   Set the tolerated distance for ending the iterations, $\epsilon_P = 1e^{-4}$. If $d_{V^W} \leq \epsilon_P$, the optimal solution from the current iteration solves the workers' problem and go to Step 5. If $d_{V^W} > \epsilon_P$ go to Step 3 by updating $V^W(A; r, w)^0 = V^W(A; r, w)^1$, $c(A)^0 = c(A)^1$, $l(A)^0 = l(A)^1$ and $A'(A)^0 = A'(A)^1$ as the new starting values.

6. Save the value function $V^W(A_i; r, w) = V^W(A_i; r, w)^1$, the $A'(A_i) = A'(A_i)^1$ and $l(A_i) = l(A_i)^1$ and $c(A_i) = c(A_i)^1$ for each $A_i \in Agrid$, $i \in \{1, ..., nA\}$.

[46] Even though the functions (e.g., value functions, $V^W(A)$ and $P(V^E)$) are continuous per se, they can only be evaluated on the discrete grid points in the numerical exercise. Namely, it is a mapping of each grid point into a number. This applies in all algorithms.

[47] We apply the *fmincon*-command, which finds the minimum of a constrained nonlinear multivariable function using the interior point algorithm.

[48] We use spline interpolation, which is a cubic interpolation of the values of neighbor-points.

E.2 Financial Contract

1. Use constant r and w.

2. Set a grid for the state variable V^E on the interval $\left[V^E_{min}, V^E_{max}\right]$. $V^E grid$ denotes the $nV^E = 50$ Chebyshev grid points of V^E on the interval. We manually replace the lowest Chebyshev point with the lower bound V^E_{min}.

3. Make an initial guess of the functional form of the value function, $P(V^E; r, w)^0$, and of the policy functions, $\left\{b(V^E)^0, m_s(V^E)^0, V^E_s(V^E)^0\right\}_{s\in\{h,l\}}$. We use $P(V^E_i; r, w)^0 = \log(-V^E_i)$, $b(V^E_i)^0 = 1$, $m_h(V^E_i)^0 = 3$, $m_l(V^E_i)^0 = 1$, $V^E_h(V^E_i)^0 = V^E_l(V^E_i)^0 = V^E_i$ for each $V^E_i \in V^E grid$, $i \in \{1, ..., nV^E\}$.

4. Solve for each $V^E_i \in V^E grid$, $i \in \{1, ..., nV^E\}$ the optimal contract in Section 2.3.2 subject to (PK), (IC), (LL) and (CC).[49] This gives the optimal contract at each $V^E_i \in V^E grid$, $i \in \{1, ..., nV^E\}$, $\left\{b(V^E)^1, m_s(V^E)^1, V^E_s(V^E)^1\right\}_{s\in\{h,l\}}$, and thus the updated value function, $P(V^E; r, w)^1$.[50]

5. Compare the two successive iterations of value functions, $P(V^E_i; r, w)^1$ with $P(V^E_i; r, w)^0$, by defining a distance measure d_P, such that
$$d_P \equiv \max_{i\in\{1,...,nV^E\}} \left|P(V^E_i; r, w)^1 - P(V^E_i; r, w)^0\right|.$$
If $d_P \leq \epsilon_P$, then take the current iteration of value function and policy functions as the solution and go to Step 6. If $d_P > \epsilon_P$, start over with Step 3 by updating $P(V^E; r, w)^0 = P(V^E; r, w)^1$ and $b(V^E)^0 = b(V^E)^1$, $m_h(V^E)^0 = m_h(V^E)^1, m_l(V^E)^0 = m_l(V^E)^1$, $V^E_h(V^E)^0 = V^E_h(V^E)^1$, and $V^E_l(V^E)^0 = V^E_l(V^E)^1$ as the new starting value for the next iteration.

6. Save the value function $P(V^E; r, w) = P(V^E; r, w)^1$ and the optimal contract $b(V^E) = b(V^E)^1$, $m_h(V^E) = m_h(V^E)^1, m_l(V^E) = m_l(V^E)^1$, $V^E_h(V^E) = V^E_h(V^E)^1$, and $V^E_l(V^E) = V^E_l(V^E)^1$.

E.3 Life-Path Simulation and Equilibrium Variables[51]

In this appendix, we simulate entrepreneurs' life paths and calculate the aggregate variables related to entrepreneurs by combining the optimal solution from Section E.2 and the entrepreneurs' decisions. Moreover, we calculate workers' aggregate deposits D and the labor

[49] For, (IC) we put in the constraint only the binding local downward constraint, since by the result of Lemma 1 the local upward constraint is never binding for the optimal contract.

[50] As in the algorithm for solving the workers' problem, we apply the *fmincon*-command to solve for the optimal contract at each grid point and use spline interpolation to calculate the value function for the next iteration.

[51] The numerical procedure described in this section is performed for given interest rate and wage rate, $\{r, w\}$.

supply L^S according to (20) and (17), respectively.[52] Using the aggregate variables, we calculate equilibrium values for the entrepreneurial share.

1. Simulate for $N^E = 10,000,000$ entrepreneurs' age and life paths with their history of productivity realizations. We use two random numbers, u_t^i and o_t^i, to denote entrepreneur i's productivity realization and death / survival at time t, respectively. The procedure is as follows:
   (a) Start from entrepreneur $i = 1$, period $t = 1$.
   (b) Use a random number generator to generate two numbers o_1^i and u_1^i, which are uniformly distributed on the interval [0, 1].
   (c) If $u_1^i < \pi_l$, save the entrepreneur's productivity realization in this period as low, $\theta_t^i = \theta_l$, and otherwise as high $\theta_t^i = \theta_h$.[53]
   (d) If $o_t^i < \Delta$, the entrepreneur survives to the next period, increase t by one and go back to Step 1b. If $o_t^i > \Delta$, the life path stops. Save her year of life, $A^i = t$, go to Step 1b and simulate for the next entrepreneur $i + 1$.
   (e) Save all entrepreneurs' years of life, $\{A^i\}_{i=1}^{N^E}$, and the sequence of productivity realizations, $\{\theta_t^i\}_{t=1}^{A^i}, i = 1, \ldots, N^E$.

2. Using the simulation of the life paths from Step 1 and the policy function from Section E.2, we determine the corresponding promised value V^i, bank loans $b(V^i)$ and the repayments $m(V^i, \theta_{A^i}^i)$ for each entrepreneur i in their last period in life $t = A^i$. Notice that the promised utility relevant for calculating the bank loans and repayments is the value at the beginning of the last period. This means that the productivity realization in $t = A^i$, $\theta_{A^i}^i$, is only used for calculating the repayments (see Step 2d). Specifically, the procedure is as follows:
   (a) Set $V_0^i = V^W(0; r, w)$ using the workers' value function from Section E.1.
   (b) Start from entrepreneur i, period $t = 1$.
   (c) If $t \leq A^i - 1$, the promised utility of entrepreneur i at the beginning of period t is $V_t^i = V_s^E(V_{t-1}^i)$, where $V_s^E(V)$ is entrepreneurs' transition function solved in Section E.2. Repeat the step until the condition is no longer satisfied.

[52] Calculating the aggregate deposits and the aggregate labor supply according to analytical expressions instead of applying life path simulation is simply to save computational time. Notice that to guarantee that the equilibrium factor prices approximate the true values under acceptable computational error, ϵ_{GE}, we only need to make sure that the computed aggregate values approximate the true values at higher precision, irrespective of the way they are calculated. And this is guaranteed in Step 4 for the aggregate deposits and the aggregate labor supply (i.e., $\epsilon_L = 1e^{-4} < \epsilon_{GE}$). For the aggregate capital and labor demand, on the other hand, we have checked that by increasing N^E to ten times the current number, the changes of these two aggregate values are below $\epsilon_L = 1e^{-4}$ as well.

[53] We set π_l=1/2 and $\Delta = 0.92$.

(d) Calculate the optimal banks loans and repayments, $\left\{b^i, m^i\right\}_{i=1}^{N^E}$, using the policy functions solved in Section E.2, the productivity realization in the last period of life, $\{\theta^i_{A^i}\}_{i=1}^{N^E}$ from Step 1, and the promised utility from Step 2c, $\{V^i_{A^i-1}\}_{i=1}^{N^E}$. Specifically, $b^i = b(V^i_{A^i-1})$ and $m^i = (V^i_{A^i-1}, \theta^i_{A^i})$.

3. Calculate the aggregate bank loans B and repayments M. Specifically, we aggregate banks loans b^i and repayments m^i over all entrepreneurs $i = 1, \ldots, N^E$ and divide the two sums by N^E to normalize the mass of the population to 1. Further, we calculate the aggregate capital demand K^D and aggregate labor demand L^D according to (D.1).

4. Calculate workers' aggregate deposits D and the labor supply L^S according to (20) and (17), respectively, and the optimal decisions solved in E.1. Specifically, start from the deposits and labor supply of workers of age $t = 1$, $Sum_A = p_A A'(A_{t-1})$ and $Sum_L = l(A_{t-1})$, weighted by their population size, $(1-\Delta)\Delta^{t-1}$. Constantly add to Sum_A and Sum_L the weighted deposits and labor supply of the older generation, until the differences between the sums in two successive iterations is below $\epsilon_L = 1e^{-4}$, respectively.[54]

5. Determine the share of entrepreneurs from the labor market condition, $\lambda = \frac{L^D}{L^D+L^S}$, the zero-profit condition, $P(V^w(0; r, w); r, w)$, and the excess capital demand, $X(r, w) \equiv \lambda K^D - (1-\lambda)D - \lambda\frac{B-M}{r}$. Notice that all endogenous variables, $\left\{\lambda, K^D, D, B, M\right\}$ are functions of the factor prices, (r, w).

E.4 Numerical Procedure for Solving the Stationary General Equilibrium

In this section, we characterize the procedure for solving the stationary general equilibrium. Specifically, we search for the values of factor prices, (r, w), and of firm entry, λ, so that the zero-profit condition is satisfied and the capital market clears. Detailed discussions of the theoretical ground and the intuitions for the algorithm are given in Section F. Denote banks' profit when granting entrepreneurs the lifetime utility of a worker (i.e., $V^E = V^W(0; r, w)$) by

$$\Pi(r, w) \equiv P(V^W(0; r, w); r, w), \tag{E.1}$$

Notice that $\Pi(r, w) = 0$ is the zero-profit condition in the stationary general equilibrium. Furthermore, denote the excess demand for capital in the economy by

$$X(r, w) \equiv \lambda(r, w)K^D(r, w) - (1-\lambda(r, w))D(r, w) - \lambda(r, w)E(r, w). \tag{E.2}$$

The procedure is as follows:

[54] We set $\epsilon_L << \epsilon_{GE}$ so that the equilibrium is not susceptible to calculation error in the workers' aggregate variables.

1. Start with an initial guess for the equilibrium factor prices, (r_0, w_0). Set the tolerance parameter $\epsilon_{GE} = 1e^{-3}$.[55] Denote the upper and the lower bound of r and w as r_U, r_L, w_U, and w_L, respectively. Initialize the value of these variables to zero. Denote the partial derivatives of the banks' profit, Π, and of the excess capital demand, X, with respective to r and w at (r_0, w_0) as S_r^Π, S_w^Π, S_r^X, and S_w^X, respectively.[56]

2. Set $(r', w') = (r_0, w_0)$.

3. Calculate workers' lifetime utility $V^W(0; r', w')$ according to Algorithm E.1, the value function of the optimal lending contract $P(V^E; r', w')$ according to Algorithm E.2, and the value of $\Pi(r', w')$ defined in (E.1).

4. If $|\Pi(r', w')| < \epsilon_{GE}$, the zero-profit condition is satisfied under the pre-defined tolerance parameter. Go to Step 6. Otherwise, update the lower bound of the equilibrium interest rate $r_L = r'$ if $\Pi(r', w') > 0$. Save $\Pi_L = \Pi(r', w')$. If $\Pi(r', w') > 0$, update $r_U = r'$ and denote $\Pi_U = \Pi(r', w')$.

5. If both the upper and the lower bound are found, let $r_0 \equiv \frac{r_U \Pi_L - r_L \Pi_U}{\Pi_L - \Pi_U}$. Otherwise, $r_0 \equiv r' - \Pi(r', w')/S_r^P$.[57] Go to Step 2.

6. Calculate the excess capital demand $X(r', w')$ and share of entrepreneurs $\lambda(r', w')$ according to Section E.3.

7. If $|X(r', w') - \Pi(r', w')| < \epsilon_{GE}$, then the (r', w') and the corresponding $\lambda(r', w')$ in the current iteration are the equilibrium.[58] Otherwise, update the lower bound of the equi-

[55] Since zero cannot be achieved numerically, we set a sufficiently small number as the tolerance parameter. Specifically, we use $|\Pi(r, w)| < \epsilon_{GE}$ as the criterion for determining whether the zero-profit condition is satisfied. Similarly, if $|X(r, w)| < \epsilon_{GE}$, we regard the capital-market-clearing condition to be satisfied.

[56] To approximate the partial derivatives, we calculate the value of $\Pi(r, w)$ and $X(r, w)$ at (r', w'), $(r' + \varepsilon, w')$ and $(r', w' + \varepsilon)$, respectively. According to the definition of partial derivatives, we have $S_r^\Pi \approx \frac{\Pi(r'+\varepsilon, w') - \Pi(r', w')}{\varepsilon}$, $S_w^\Pi \approx \frac{\Pi(r', w'+\varepsilon) - \Pi(r', w')}{\varepsilon}$, $S_r^X \approx \frac{X(r'+\varepsilon, w') - X(r', w')}{\varepsilon}$, and $S_w^X \approx \frac{X(r', w'+\varepsilon) - X(r', w')}{\varepsilon}$. Notice that this procedure is time consuming, because we need to calculate through the entire model at each combination of (r, w). Since the searching region for the general equilibrium is relatively small (within interval of magnitude 0.01), the change in partial derivatives is small. Therefore, we use these values as an approximation of the partial derivatives in all iterations to save computational time.

[57] Due to the inaccurate approximation of the partial derivatives of Π, it is impossible to calculate the exact increase in r'. This means that there may be back and forth in the adjustment of r'. To guarantee that the r_0^* such that $|\Pi(r_0^*, w')| < \epsilon_{GE}$ can be found in finite iterations, we record the upper r_U and the lower bound r_L of region where r_0^* lies in each iteration, and use binary search as is described in Step 5 and 8.

[58] Notice that by setting the criterion as $|X(r, w) - \Pi(r, w)| < \epsilon_{GE}$ instead of $|X(r, w)| < \epsilon_{GE}$, we decreases the computational error of the equilibrium factor prices. Essentially, we want to avoid the case when both $X(r, w)$ and $\Pi(r, w)$ are marginally below ϵ_{GE} but of the opposite sign. By analysis similar as illustrated in Figure 12, this deviates the numerical solution from the true values much more than if both $X(r, w)$ and $\Pi(r, w)$ are marginally below ϵ_{GE} but of the same sign.

librium wage $w_L = w'$ if $X(r', w') > 0$. Save $X_L = X(r', w')$.[59] If $X(r', w') > 0$, update the upper bound of the equilibrium wage $w_U = w'$ and set $X_U = X(r', w')$.

8. If both the upper and the lower bound are found, let $w_0 \equiv (w_L + w_U)/2$, and $r_0 \equiv r' - \left(\Pi(r', w') + S_w^P(w_0 - w')\right)/S_r^P$. Otherwise, we set

$$\begin{pmatrix} r_0 \\ w_0 \end{pmatrix} \equiv \begin{pmatrix} r' \\ w' \end{pmatrix} - A^{-1}b, \tag{E.3}$$

where $A = \begin{pmatrix} S_r^P & S_w^P \\ S_r^X & S_w^X \end{pmatrix}$, and $b = \begin{pmatrix} \Pi(r', w') \\ X(r', w') \end{pmatrix}$.[60] Go to Step 2.

F Theoretical Ground and Intuition of Algorithm E.4

The theoretical ground of the stationary general equilibrium search in Algorithm E.4 is based on the continuity of the aggregate variables and the values functions with respect to (r, w), as well as numerical properties of the zero-profit condition, $\Pi(r, w)$, and the excess capital demand function, $X(r, w)$, defined in (E.1) and (E.2), respectively.

Property 1. *The zero-profit condition and the excess capital demand are both decreasing in r and w (at least locally around the equilibrium values).*

This means that the partial derivatives of $\Pi(r, w)$ and $X(r, w)$ with respect to r and w are negative:

$$\Pi_r < 0,\ \Pi_w < 0,\ X_r < 0 \text{ and } X_w < 0. \tag{F.1}$$

In a (w, r)-diagram, the slope of the iso-profit curve and of the iso-excess demand curve are given respectively by

$$S_\Pi = -\frac{\Pi_w}{\Pi_r} \text{ and } S_X = -\frac{X_w}{X_r}. \tag{F.2}$$

Therefore, equation (F.1) implies that both loci are downward sloping (i.e., $S_\Pi < 0$ and $S_X < 0$). In addition, a northeast shift of the locus (i.e., an increase in r and w) decreases the corresponding value of the respective iso-curve. Furthermore, the relative position of the two loci is determined by the following property.

Property 2. *The gap between the two equilibrium conditions, $G \equiv X - \Pi$, is decreasing in r and increasing in w.*

[59]An explanation of why the two cases correspond to the upper and the lower bound is given in Footnote 62 and the corresponding part in the main text.

[60]We derive the Tayor's expansion of $\Pi(r, w)$ and $X(r, w)$ at (r', w'), respectively. For (r_0, w_0) in the small neighborhood of (r', w'), we have $\Pi(r_0, w_0) \approx \Pi(r', w') + \frac{\partial \Pi}{\partial r}(r_0 - r') + \frac{\partial \Pi}{\partial w}(w_0 - w')$, and $X(r_0, w_0) \approx X(r', w') + \frac{\partial X}{\partial r}(r_0 - r') + \frac{\partial X}{\partial w}(w_0 - w')$. Letting $\Pi(r_0, w_0)$, $\Pi(r', w')$, and $X(r_0, w_0)$ be zero, we get equation (E.3).

Property 2 implies that the iso-profit curve is steeper than the iso-excess demand curve at all combination of (r, w) locally. To see this, note that the slopes of the iso-profit and iso-excess demand curves are given by equation (F.2). Since Property 2 indicates that the partial derivatives satisfy $G_r < 0$ and $G_w > 0$, we have

$$\Pi_r > X_r \quad \text{and} \quad X_w > \Pi_w. \tag{F.3}$$

Therefore, the slopes of the two loci satisfy $|S_X| < |S_\Pi|$. A direct implication is the single-crossing property of the two loci. If the two curves ever cross they cross only once.[61] This establishes the uniqueness of the stationary equilibrium. Furthermore, the properties of the iso-curves indicate the direction for approaching the equilibrium from any off-equilibrium point.

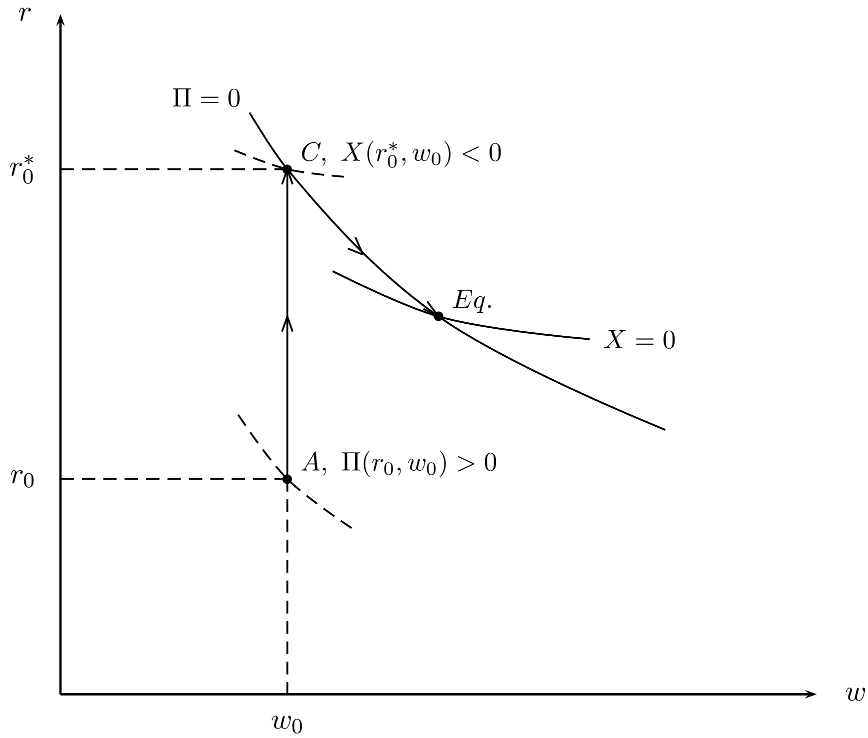


Figure 12: Iso-profit and iso-excess demand curves

Note: The two solid curves represents iso-profit and iso-excess-demand curve, with a value of zero respectively (i.e., $\Pi = 0$ and $X = 0$). Isoquants that lie below the two curves indicate positive values and above them negative ones.

Figure 12 illustrates of iso-profit and iso-excess demand curves and gives an intuition of the algorithm to find the equilibrium, $Eq.$. Suppose that at an initial guess (r_0, w_0) (e.g., point A)

[61] The fact that the two curves cross (i.e., the existence of the equilibrium) is guaranteed in numerical practice. In the region we search for the equilibrium, there always exist combinations of (r,w) on the zero-profit locus, s.t. $X(r, w) > 0$, and combinations, s.t. $X(r, w) < 0$. Since the zero-excess-capital-demand locus must lie between the loci that pass through the above-mentioned two types of combinations, the zero-profit locus, and the zero-excess-capital-demand locus cross.

the value of the iso-profit is $\Pi(r_0, w_0) > 0$. First, we approach the $\Pi = 0$ locus by changing r to r_0^*, s.t.(r_0^*, w_0) is on the locus (Step 4 and 5 in Algorithm E.4). Then at (r_0^*, w_0) the excess capital demand $X(r_0^*, w_0)$ can be positive, negative or 0. In the last case we have found the equilibrium $Eq.$ directly. Now suppose $X(r_0^*, w_0) < 0$ (e.g., point C). Equation (F.1) and (F.3) suggest that the stationary equilibrium lies south-east of C.[62] Therefore, we shift (r, w) - $r \downarrow, w \uparrow$ - along the locus of the iso-profit curve $\Pi = 0$ until the excess demand increases to 0 ($C \rightarrow Eq.$).[63]

G Extension: Capital Requirements on Banks

So far, our discussion focuses on the effect of an exogenous liquidity shock on the allocation of credit among firms, firm dynamics, and the macroeconomy. In this appendix, we show that the model can be modified to analyze the effects of capital requirements on the above-mentioned aspects of an economy as well.

Suppose that banks are required to maintain an exogenous capital ratio, ξ. The capital ratio is defined as the ratio between bank equities and total assets in a stationary equilibrium. An increase in the capital ratio represents a tightening of capital requirements on banks. According to banks' balance sheet in Table 3, this implies that $\frac{E}{D+E} = \xi$, or equivalently

$$E = \frac{\xi}{1-\xi} D. \tag{G.1}$$

To accommodate the new setup, we relax the assumption that the banking sector is perfectly competitive. This implies that the zero-profit condition defined in (13) no longer holds and (G.1) has become a new equilibrium condition. The equilibrium factor prices, $\{r, w\}$, and firm entry, λ, are now determined by (G.1), the labor- and capital-market-clearing conditions in (16) and (18), respectively, and households indifferent occupational decision in (14).

Using the same parameter values as in Section 3 and a capital ratio of $\xi = 10\%$, we solve the stationary equilibrium numerically:[64]

[62] This also means that $w_0 < w_{Eq}$. Therefore, w_0 is one lower bound of the equilibrium wage. We will update the lower bound if a new w', s.t. $w_0 < w' < w_{Eq}$ is found. The arguments apply for the upper bound as well.

[63] Similar to the situation described in Footnote 55 it is not possible to find the correct adjustment in (r, w) in one step. Several iterations may be needed and we apply similar technique (i.e., recording upper and lower bounds and updating w in each iteration with binary search) to guarantee that the equilibrium is found in finite iterations. In addition, as we change (r, w) in each iteration, we need to make sure that the change is along the locus of the iso-profit curve $\Pi = 0$. Otherwise, we need to apply the first step again.

[64] A rough calculation with the H.8 release from the Federal Reserve on the assets and liabilities of commercial banks in the United States suggests that the ratio of total assets minus total liabilities and total assets is in the range of $[4\%, 14\%]$ between 1973 and 2020. Karmakar and Mok (2015) investigate various measures of capital ratio including the capital adequacy ratio (CAR), defined as the total capital as a fraction of risk weighted assets.

Table 4: Equilibrium variables in an economy with capital requirements

Equilibrium factor prices and firm entry		Value
Interest rate	r^*	4.25%
Wage	w^*	0.146
Share of entrepreneurs	λ^*	2.72%
Aggregate variables		Value
Aggregate bank loans	B	0.06
Bank equities	E	0.022
Aggregate labor employment	L^D	0.258
Aggregate capital employment	K^D	0.154
Aggregate output	Y	0.063

Note: The lifetime utility of an entrepreneur $V_0^E = -8.42$, equals to that of a worker, $V^W(0, r^*, w^*)$. The value of bank equities is calculated using Equation (23).

Similar to the analysis in Section 3.3, we derive the firm dynamics in the current scenario using simulated paths of productivity shocks on firms and the optimal lending contracts calculated with the equilibrium factor prices in Table 4. The results are summarized in Figure 13.

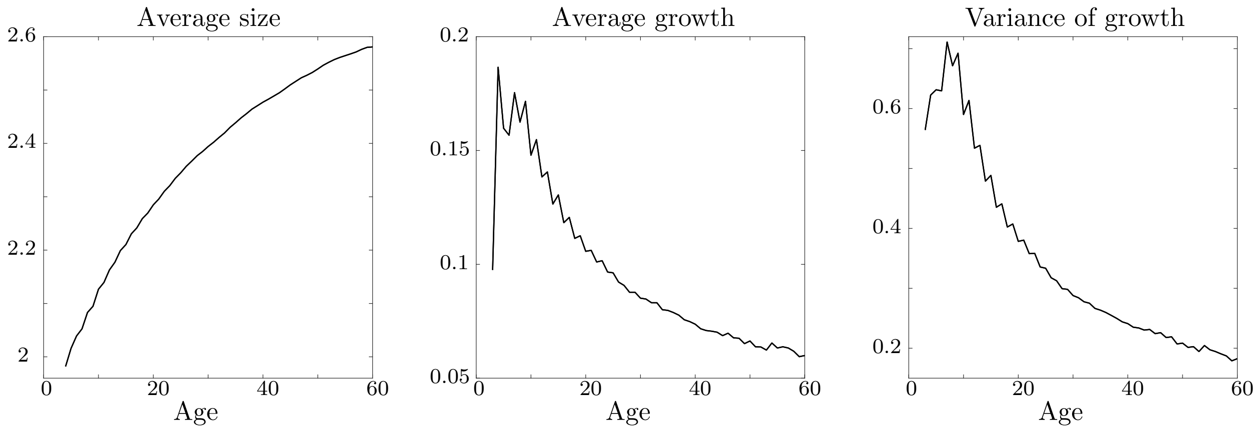


Figure 13: Firm dynamics in an economy with capital requirements

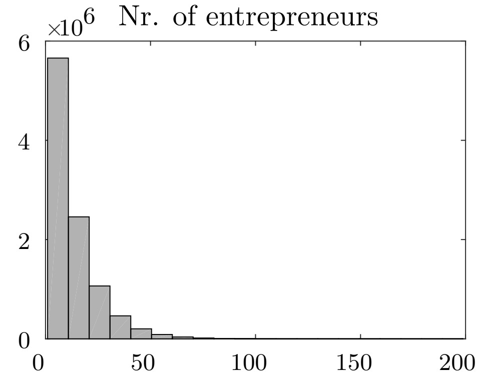
Note: The variables illustrated are the 5-year moving-average values.

Using data from the call reports database of the Chicago Fed between 1996-2010, they show that the average CAR of U.S. commercial banks is about 15%.

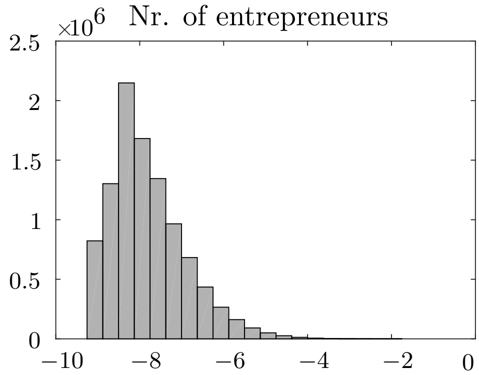
H Figures

H.1 Firm Distributions

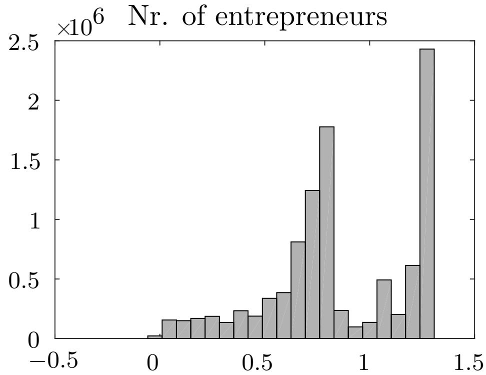
With the simulation of entrepreneurs' life paths described in Appendix E.3 and the optimal dynamic contract solved for the calibrated economy in Figure 3, we derive the distributions of various firm characteristics in equilibrium. Figure 14 illustrates the distribution of all entrepreneurs in the economy with respect to different characteristics: Age, promised values, repayments and bank loans.



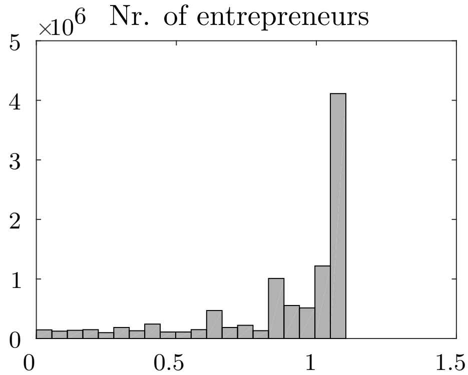
(a) Age



(b) Promised value, V^E, ($V_0^E = -8.36$)



(c) Repayments, m



(d) Bank loan, b

Figure 14: Distribution of age, promised values, repayments and bank loans

Subplot (a) shows the age distribution of entrepreneurs, which is given by the exponential distribution $\{(1-\Delta)\Delta^\tau\}_{\tau=0}^\infty$ in equilibrium.

Subplot (b) shows the distribution of firms' promised value, $\psi^E(V^E; r, w)$, which is used to calculate the aggregate variables in Section 2.4. The mass lies around the starting value $V_0^E = -8.36$, because the newborns account for the largest share in the economy. Firm heterogeneity then arises from the different length and composition of productivity realizations over firms' lifetime. The further away from the starting value V_0^E, the lower is the density of V^E because of the longer and more heterogeneous life paths underlying such values. Moreover, the distribution is skewed to the right. In the calibrated economy, the average firm size increases as firms get older as a result of a higher promised utility on average over time (i.e., $\mathbb{E}(V_S^E) > V$). This is reflected by the larger weight on higher promised values. At the same time, firms have more heterogeneous productivity histories at higher promised values, which corresponds to the more dispersed distribution on the right tail. A less obvious way to consider the underlying drive force is the following: When the high and low state happen with equal probability, $\mathbb{E}(V_S^E) > V$ implies that the increase in firms' promised value after a high productivity shock is larger than the decrease after a low productivity shock.[65] To see the implication of this property for the distribution of promised values, consider an entrepreneur with a promised value of V_0^E. After a high productivity shock, the entrepreneur needs more than one low shocks to come back to V_0^E, whereas after a low productivity shock, one high shock will lead to a promised utility higher than V_0^E. Therefore, on average more firms have a promised utility that is higher than the initial value. Namely, a larger fraction of firms lies to the right of V_0^E.

The distribution of repayments in Subplot(c) follows from the distribution of promised values $\psi^E(V^E; r, w)$ and the optimal dynamic contract. Depending heavily on the current period productivity realizations, the levels of repayments are separated into two groups. This means, the repayments exhibit two distinct sub-distributions, because the gap in repayments of high and low state are relatively large (comparing $m_h(V^E)$ and $m_l(V^E)$ in Figure 3). Moreover, similar to the distribution of the promised value, repayments of new-born entrepreneurs at 1.1 and 0.7, which correspond to $m_h(V_0^E)$ and $m_l(V_0^E)$, respectively, account for the greatest share in the distribution. Furthermore, the average promised value increase as firms get older and there exists a negative relationship between the optimal repayment and the promised value, except for regions close to the lower bound of V^E. Therefore, the repayments decrease over time and the distribution of repayments disperse to the left of the initial mass over time.

The distribution of bank loans in Subplot(d) shows a strong negative skewness resulting from the shape of the optimal bank loans in Figure 3. As is discussed in 3.2, the optimal level of bank loans consists of a rapid-growing part and a relatively flat part: At lower levels of promised values, profit-maximizing banks increase the size of bank loans at a fast pace. For the wide range of promised values following that, however, firms are provided a level of bank loans that is close to the efficient level at $b^* = 1.18$. Since firms' promised value evolves into the second range over time, the mode of the distribution of bank loans accumulates around this value.

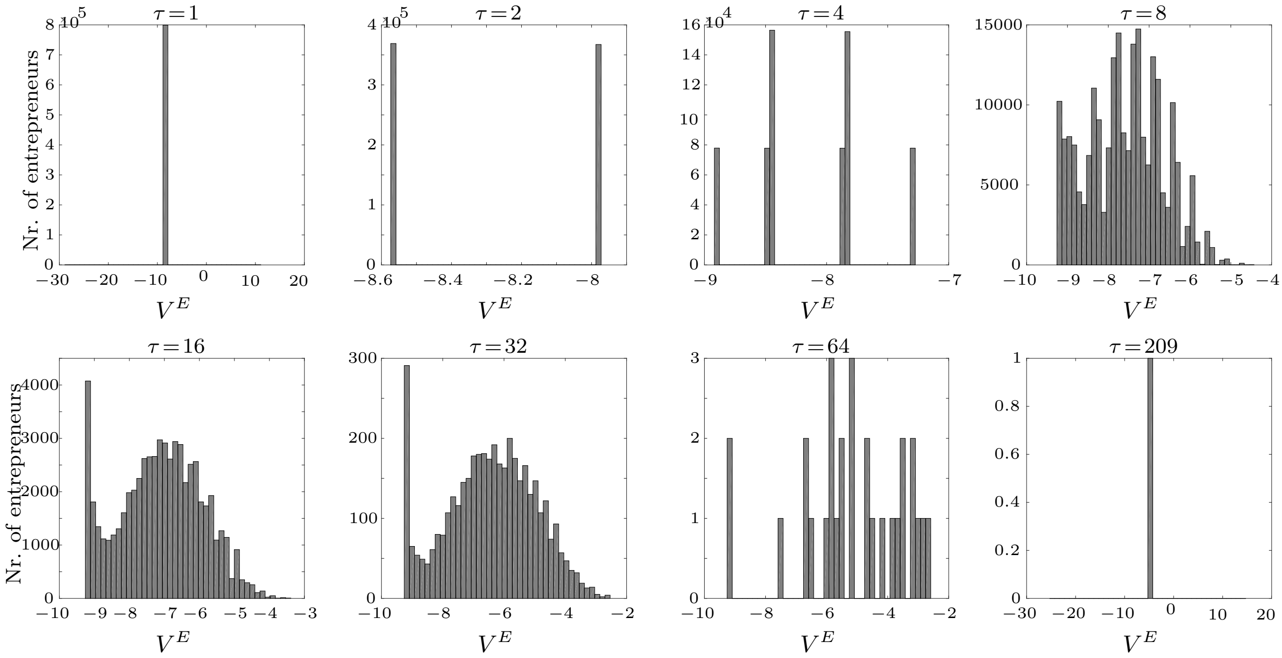


Figure 15: Development of entrepreneurs' promised value distributions

Note: The accumulations of entrepreneurs with a promised utility close to $V_{min} = -9.26$ (as in Subplot $\tau = 16$ and $\tau = 32$) are because $V^E = V_{min}$ is an absorbing state.

H.2 Firm Distribution by Cohorts

In this section, we show the distribution of promised values, the level of bank loans, and repayments across cohorts of firms in Figure 15, 16, and 17, respectively. In particular, the distribution of bank loans across cohorts is the basis of calculating the average firm size (i.e., the mean given the distribution) over lifetime illustrated in Figure 5.

Initially, all newborn entrepreneurs are promised a lifetime utility of $V_0^E = -8.36$. According to the optimal contract solved in Section 3.2, V_0^E determines the level of bank loans, $b(V_0^E)$, provided to newborn entrepreneurs and the required repayments, $m_s(V_0^E)$, following the realization of productivity shocks in the first period, $s \in \{h, l\}$. Therefore, for Cohort 1 (i.e., $\tau = 1$), the distribution of promised values and of bank loans concentrate at V_0^E and $b(V_0^E)$, respectively, where the number of entrepreneurs is the number of newborn entrepreneurs in the economy (see Subplot $\tau = 1$ in Figure 15 and 16).[66] And the repayments distributes with equal number of entrepreneurs at $m_h(V_0^E)$ and $m_l(V_0^E)$, reflecting the equal probability distribution of the two

[65] From $\mathbb{E}(V_S^E) > V^E$ and the fact that high and low state happen with equal probability $\pi_h = \pi_l = 0.5$, we reformulate the equation as $V_h^E - V^E > V^E - V_l^E$.

[66] Since the number of entrepreneurs in the simulation is set as $N = 10,000,000$, the number of newly born entrepreneurs is $(1 - \Delta)N = 8 \times 10^5$, where $\Delta = 92\%$.

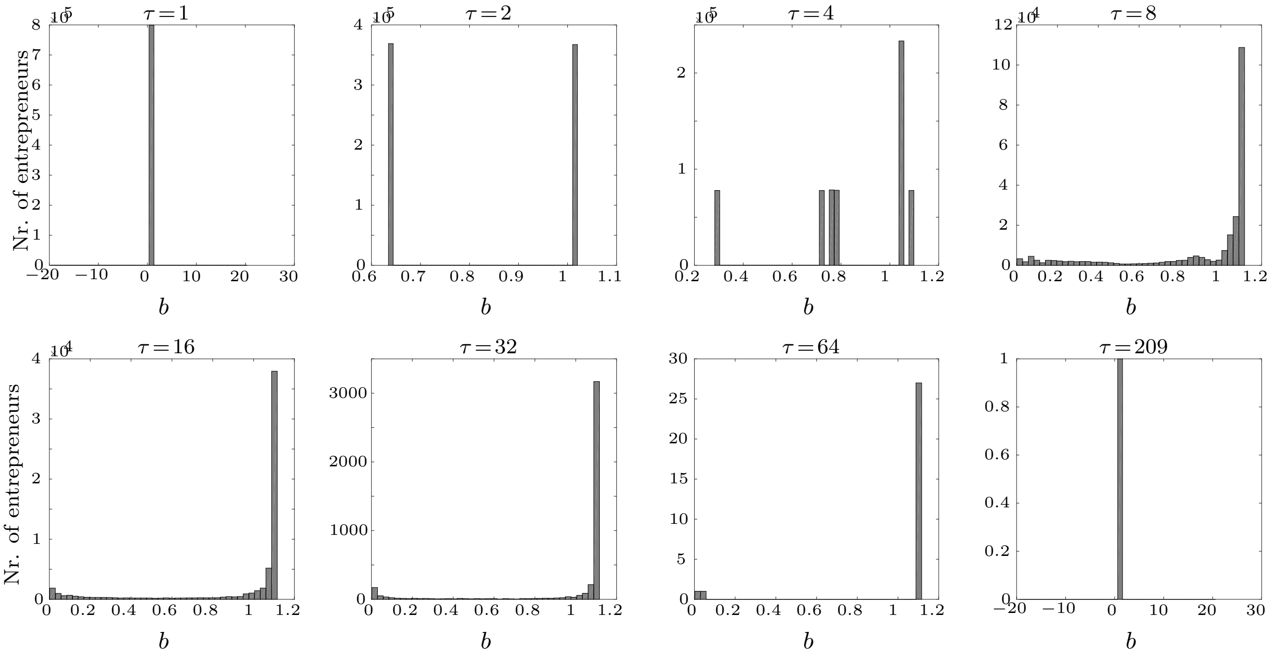


Figure 16: Development of entrepreneurs' bank loans

Note: Similar to the discussion in Appendix H.1, the distribution exhibits an accumulation of firms with a level of bank loans close to the efficient level $b^* = 1.18$.

states (see Subplot $\tau = 1$ in Figure 17).

Depending on the productivity realization at Age 1, the surviving firms are promised a continuation lifetime utility according to the transition function $V_s^E(V^E)$. In particular, a half of the entrepreneurs of Cohort 1 gets a promised utility of $V_h^E(V_0^E)$ and the other half gets $V_l^E(V_0^E)$. Then, banks provide firms the corresponding levels of bank loans. These are shown as the two bars in Subplot $\tau = 2$ Figure 15 and 16, respectively. The required repayments depends on the promised utility and the productivity realizations in the current period. Hence, there are four potential levels of repayments in Subplot $\tau = 2$ Figure 17.

Similar process goes on as τ gets larger. Over time, the paths of productivity realizations get more heterogeneous due to the i.i.d. shocks and the distribution of the variables gets more dispersed. In addition, as age advances, the cohort size becomes smaller because more firms have been exiting the market with the exogenous rate $1 - \Delta$. Eventually, (almost) all firms of a given cohort exit the market and the distribution of old cohorts consists of very few individual observations (e.g., for Cohort $\tau = 64$, only 29 firms are left and for Cohort $\tau = 209$, only one firm is left).

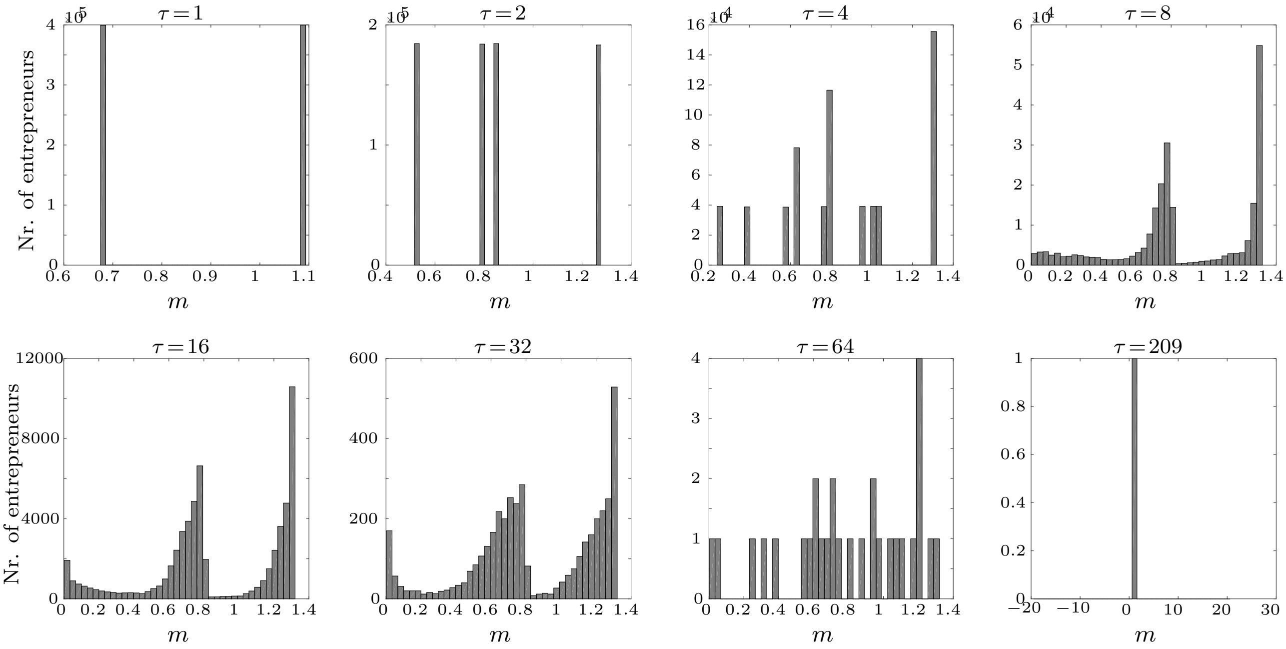


Figure 17: Development of entrepreneurs' repayments to banks

Note: Similar to Figure 14c, the distribution shows distinct sub-distributions (e.g., in Subplot $\tau = 8, 16, 32$). This suggests that the level of repayments relies heavily on the current period productivity realizations and the gap in the required repayments in the two states is large.

H.3 Illustration of Productivity Shock

In this section, we simulate the life path of an entrepreneur for fifty years. Given the life path and the optimal contract in Section 3.2, we calculate the corresponding dynamics in promised value, the level of bank loans, and repayments. Figure 18 to 20 illustrate three circumstances: The life path of entrepreneurs with more high productivity shocks, with similar numbers of high and low productivity shocks, and with more low productivity shocks, respectively.

When an entrepreneur experiences more high productivity shocks, the promised value increases over time. Hence, firm size increases and the firm grows out of the financial constraint in the end (see Figure 18). The high fluctuation in the repayments comes from the large gap in repayments after a high and a low productivity shock and the alternate high and low productivity shocks featured by the life path of the entrepreneur. On average, the repayments decrease as the promised value increases. The scenario is the opposite when a firm experiences more low productivity shocks (see Figure 20). The only exception is the behavior of the repayments. The strong decrease in repayments implies that the promised value is within the increasing segment of the bank repayments, where the limited liability constraint (LL) and the credibility constraint (CC) are binding. With similar numbers of high and low productivity shocks, the

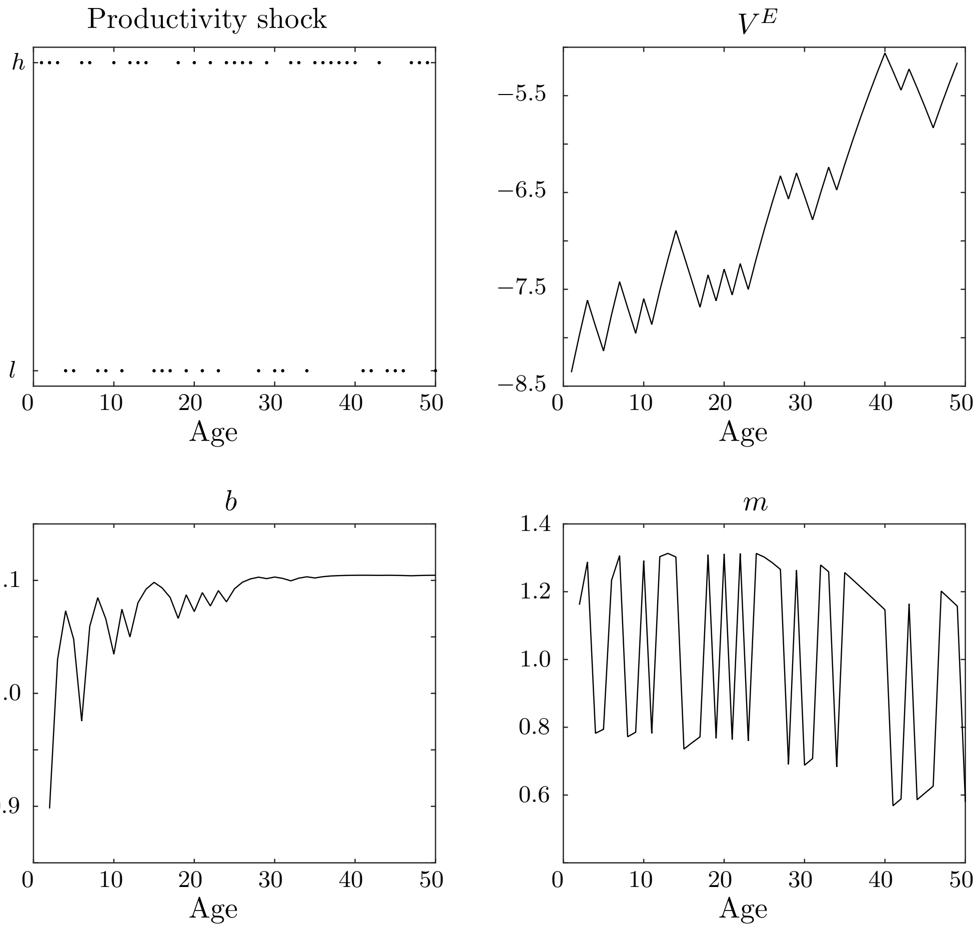


Figure 18: Life path I

Note: This figure shows the life path of an entrepreneur with more high productivity shocks. The promised value and the level of bank loans increase over time.

promised value and the level of bank loans fluctuate around a relatively stable value. In Figure 19, the high productivity shocks at the early age translate into a rapid increase in the level of bank loans. However, the successive low productivity shocks in the 20s and 30s drives the promised value and the corresponding bank loans back to a low level.

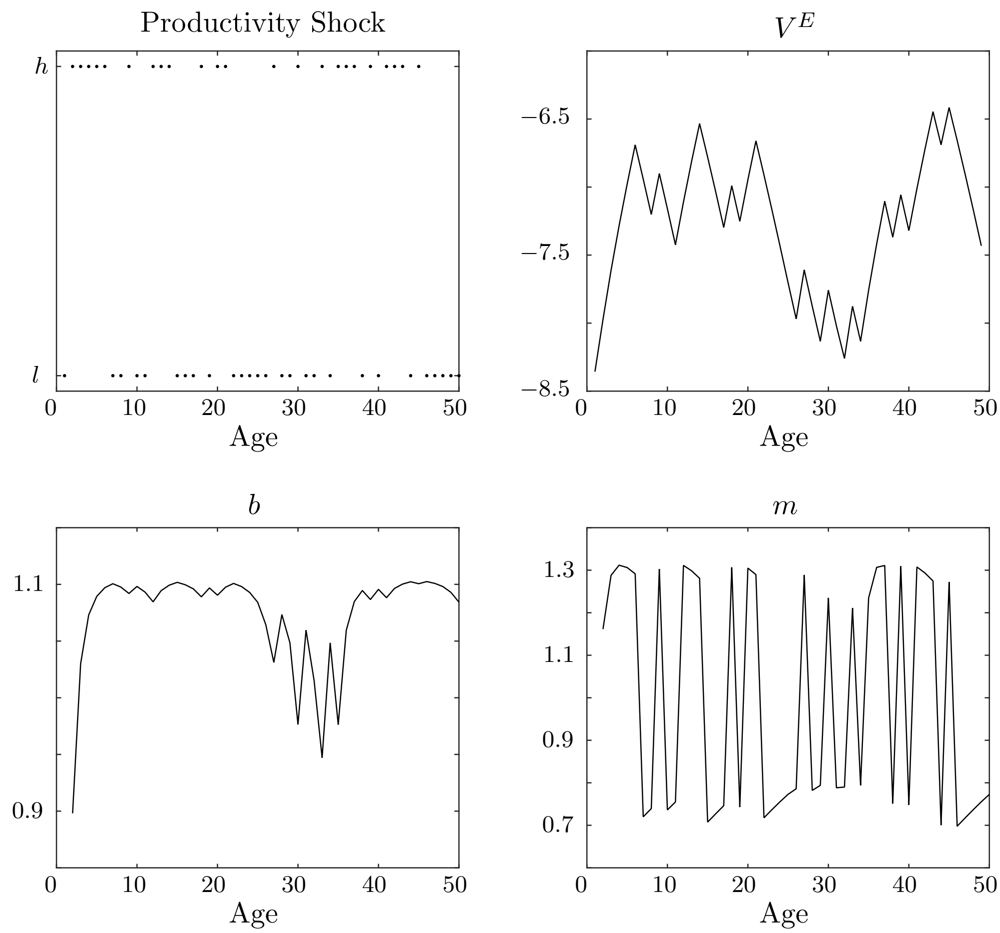


Figure 19: Life path II

Note: This figure shows the life path of entrepreneurs with similar numbers of high and low productivity shocks. The dynamics of the promised value and of the level of bank loans crucially depend on the time order of the productivity shocks.

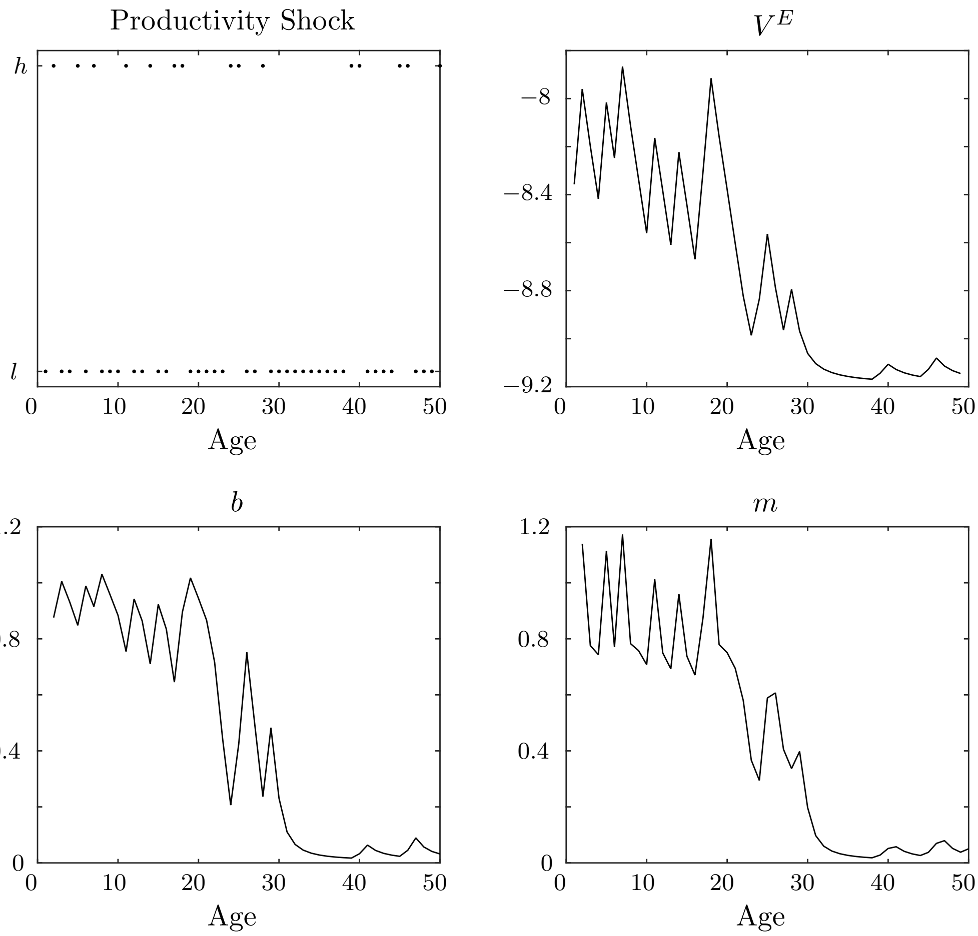


Figure 20: Life path III

Note: This figure shows the life path of entrepreneurs with more low productivity shocks. The promised value and the level of bank loans decrease over time. The strong decrease in repayments implies that the promised value is within the increasing segment of the bank repayments, where the limited liability constraint (LL) and the credibility constraint (CC) are binding.

References

Albuquerque, R. and H. A. Hopenhayn (2004). Optimal lending contracts and firm dynamics. *The Review of Economic Studies 71*(2), 285–315.

Aldasoro, I. and R. Unger (2017). External financing and economic activity in the euro area - why are bank loans special? *BIS Working Papers No. 622*.

Atkeson, A. and R. E. J. Lucas (1992). On efficient distribution with private information.

Review of Economic Studies 59(3), 427–435.

Basu, S. and J. G. Fernald (1997). Returns to scale in us production: Estimates and implications. *Journal of Political Economy 105*(2), 249–283.

Berger, A. N. and G. F. Udell (1995). Relationship lending and lines of credit in small firm finance. *Journal of Business*, 351–381.

Berger, A. N. and G. F. Udell (1998). The economics of small business finance: The roles of private equity and debt markets in the financial growth cycle. *Journal of banking & finance 22*(6-8), 613–673.

Bernanke, B. and M. Gertler (1989). Agency costs, net worth, and business fluctuations. *American Economic Review 79(1)*, 14–31.

Bernanke, B. S., M. Gertler, and S. Gilchrist (1999). The financial accelerator in a quantitative business cycle framework. *Handbook of Macroeconomics*, 1341–1393.

Berrospide, J. M. (2021). Bank liquidity hoarding and the financial crisis: An empirical evaluation. *Quarterly Journal of Finance 11*(04), 2150020.

Bharath, S. T., S. Dahiya, A. Saunders, and A. Srinivasan (2011). Lending relationships and loan contract terms. *Review of Financial Studies 24*(4), 1141–1203.

Bilbiie, F. O., F. Ghironi, and M. J. Melitz (2012). Endogenous entry, product variety, and business cycles. *Journal of Political Economy 120*(2), 304–345.

Bolton, P., X. Freixas, L. Gambacorta, and P. E. Mistrulli (2016). Relationship and transaction lending in a crisis. *Review of Financial Studies 29*(10), 2643–2676.

Brunnermeier, M. K. and Y. Sannikov (2014). A macroeconomic model with a financial sector. *American Economic Review 104*(2), 379–421.

Campbell, J. R. (1998). Entry, exit, embodied technology, and business cycles. *Review of Economic Dynamics 1*(2), 371–408.

Carlstrom, C. T. and T. S. Fuerst (1997). Agency costs, net worth, and business fluctuations: A computable general equilibrium analysis. *American Economic Review*, 893–910.

Chatterjee, S. and R. Cooper (2014). Entry and exit, product variety, and the business cycle. *Economic Inquiry 52*(4), 1466–1484.

Clementi, G. L. and H. A. Hopenhayn (2006). A theory of financing constraints and firm dynamics. *Quarterly Journal of Economics 121*(1), 229–265.

Clementi, G. L. and B. Palazzo (2016). Entry, exit, firm dynamics, and aggregate fluctuations. *American Economic Journal: Macroeconomics 8*(3), 1–41.

Cohen, J., K. Hachem, and G. Richardson (2021). Relationship lending and the Great Depression. *Review of Economics and Statistics 103*(3), 505–520.

Cooley, T., R. Marimon, and V. Quadrini (2004). Aggregate consequences of limited contract enforceability. *Journal of political Economy 112*(4), 817–847.

De Jonghe, O., H. Dewachter, K. Mulier, S. Ongena, and G. Schepens (2020). Some borrowers are more equal than others: Bank funding shocks and credit reallocation. *Review of Finance 24*(1), 1–43.

Duygan-Bump, B., A. Levkov, and J. Montoriol-Garriga (2015). Financing constraints and unemployment: Evidence from the great recession. *Journal of Monetary Economics 75*, 89–105.

Dyrda, S. (2017). Fluctuations in uncertainty, efficient borrowing constraints and firm dynamics. Unpublished manuscript.

Evans, D. S. (1987). The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries. *Journal of Industrial Economics 35(4)*, 567–581.

Gertler, M. (1992). Financial capacity and output fluctuations in an economy with multi-period financial relationships. *The Review of Economic Studies 59*(3), 455–472.

Gertler, M. and S. Gilchrist (1993). The role of credit market imperfections in the monetary transmission mechanism: arguments and evidence. *The Scandinavian Journal of Economics*, 43–64.

Green, E. J. (1987). Lending and the smoothing of uninsurable income. In E. C. Prescott and N. Wallace (Eds.), *Contractual Arrangements for Intertemporal Trade*. University of Minnesota Press, Minneapolis.

Hachem, K. (2011). Relationship lending and the transmission of monetary policy. *Journal of Monetary Economics 58*(6-8), 590–600.

Hackethal, A. and R. H. Schmidt (2004). Financing patterns: Measurement concepts and empirical results. *Working Paper Series: Finance & Accounting, No. 125, Johann Wolfgang Goethe-Universitaet Frankfurt am Main.*

Hall, B. H. (1987). The relationship between firm size and firm growth in the us manufacturing sector. *Journal of Industrial Economics 35(4)*, 583–606.

Iyer, R., J.-L. Peydró, S. da Rocha-Lopes, and A. Schoar (2014). Interbank liquidity crunch and the firm credit crunch: Evidence from the 2007–2009 crisis. *Review of Financial Studies 27*(1), 347–372.

Karmakar, S. and J. Mok (2015). Bank capital and lending: An analysis of commercial banks in the united states. *Economics Letters 128*, 21–24.

Khwaja, A. I. and A. Mian (2008). Tracing the impact of bank liquidity shocks: Evidence from an emerging market. *American Economic Review 98*(4), 1413–42.

Kiyotaki, N. and J. Moore (1997). Credit cycles. *Journal of Political Economy 105*(2), 211–248.

Ljungqvist, L. and T. J. Sargent (2018). *Recursive Macroeconomic Theory*. MIT press.

Messer, T., M. Siemer, F. Gourio, et al. (2016). A missing generation of firms? aggregate effects of the decline in new business formation. *2016 Meeting Papers. No. 752. Society for Economic Dynamics* (752).

Petersen, M. A. and R. G. Rajan (1994). The benefits of lending relationships: Evidence from small business data. *Journal of Finance 49*(1), 3–37.

Quadrini, V. (2011). Financial frictions in macroeconomic fluctuations. *Economic Quarterly*, 209–254.

Rajan, R. G. (1992). Insiders and outsiders: The choice between informed and arm's-length debt. *Journal of Finance 47*(4), 1367–1400.

Sette, E. and G. Gobbi (2015). Relationship lending during a financial crisis. *Journal of the European Economic Association 13*(3), 453–481.

Sharpe, S. A. (1990). Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *Journal of Finance 45*(4), 1069–1087.

Smith, A. A. J. and C. Wang (2006). Dynamic credit relationships in general equilibrium. *Journal of Monetary Economics 53(4)*, 847–877.

Spear, S. E. and S. Srivastava (1987). On repeated moral hazard with discounting. *Review of Economic Studies 54(4)*, 599–617.

Stokey, N. L., R. E. Lucas, and E. C. Prescott (1989). *Recursive Methods in Economic Dynamics.* Harvard University Press.

Thomas, J. and T. Worrall (1990). Income fluctuation and asymmetric information: An example of a reapeated principal-agent problem. *Journal of Economic Theory 51(2).*

Verani, S. (2018). Aggregate consequences of dynamic credit relationships. *Review of Economic Dynamics 29*, 44–67.