Bank Lending and Relationship Capital[*]

Yasser Boualam [†]

January 29, 2015

[†]The Wharton School - University of Pennsylvania. Contact: boualam@wharton.upenn.edu.

Bank Lending and Relationship Capital

Abstract

I develop a general equilibrium theory of bank lending relationships in an economy subject to search frictions and limited enforceability. The model features a dynamic contracting problem embedded within a directed search equilibrium with aggregate and bank-specific uncertainty. The interaction between search and agency frictions generates a slow accumulation of lending relationship capital and distorts the optimal allocation of credit along both intensive and extensive margins. A crisis characterized by a sizable destruction of lending relationships therefore leads to a significant contraction in credit and a slow recovery, consistent with the Great Recession. I calibrate the model to study aggregate and cross-sectional implications and analyze policies aimed at reviving bank lending.

JEL Classification: D86; E02; E22; G01; G21; G28; L26.

Keywords: Credit Markets; Relationship Banking; Dynamic Contracting; Limited Commitment; Directed Search; Financial Frictions.

1   Introduction

The recent financial crisis caused a severe disruption in bank credit markets. Limited access to bank financing impacted households and small and medium-sized enterprises (SMEs), with sharp economic consequences.[1] Banks' persistent reluctance to lend has been at the heart of policy debate and academic research, with recent macro-finance literature highlighting the critical role of deteriorated bank balance sheets and the scarcity of financial intermediary capital in amplifying the crisis and restricting bank lending.[2] Yet, despite the stabilization of the banking system, an improving economic outlook and many policy interventions, the flow of business lending has remained markedly low.

This paper argues that two key factors contributing to the sluggish credit recovery are the severance of *bank lending relationships* during the crisis and the consequent process of *credit reallocation*. In particular, I show how an environment characterized by search frictions and long-term financing contracts subject to limited enforcement can generate slow recoveries, consistent with the Great Recession.

My analysis is based on two premises inherent to bank lending markets for SMEs. First, relationship lending matters because banks are a critical source of external finance and the repeated interaction between borrowers and lenders relaxes contracting distortions and gradually enhances credit flow.[3] In this paper, these long-term financing contracts are subject to limited enforceability, which reflects the borrower's inability to commit to a given arrangement. This is notably relevant for small and private firms with limited or opaque collateral. Second, the process of credit reallocation is important because establishing these lending relationships can be costly and time-consuming in decentralized and imperfectly competitive environments. This is the case for credit markets where both borrowers and lenders often devote significant time and resources to locate the right matches.

A salient feature of my approach to bank lending is that it highlights the importance of the market structure and the contracting environment, and does not rely on fluctuating bank balance sheets or firm collateral values. I use the term "relationship capital" to describe a form of intangible capital reflecting the banking sector's aggregate capacity to funnel credit into existing lending relationships. Accumulating relationship capital is tied to the joint effects of frictions hampering *both* the formation (extensive margin) and build-up (intensive margin) of bank-firm pairs. As a consequence, the reallocation of credit in the aftermath of a crisis can be very slow. The model uncovers a propagation mechanism relying on two distinct channels. The first channel ("credit relationship channel") affects the dynamics of credit availability and pricing of existing lending relationships. The second channel ("credit origination channel") operates through search and matching and impacts the bank's decision to offer new credit opportunities as well as the contractual terms at origination.

---

[1] Ivashina and Scharfstein (2010), Chodorow-Reich (2014), Greenstone and Mas (2012).

[2] E.g. Gertler and Kiyotaki (2010), Gertler and Karadi (2011), He and Krishnamurthy (2011), Brunnermeier and Sannikov (2014), Adrian and Boyarchenko (2012), Rampini and Viswanathan (2014), and Bigio (2014).

[3] Petersen and Rajan (1995), Boot and Thakor (2000).

In this paper, I first develop and fully characterize a dynamic contracting problem embedded within a directed search equilibrium with aggregate and bank-specific uncertainty. The model has two interconnected building blocks. The first relates to the dynamic contracting problem with limited enforceability, as in Albuquerque and Hopenhayn (2004). The borrowing capacity of the firm endogenously emerges as part of the optimal contract solution. When firm value is initially low, the agency problem impedes the amount of credit available because the entrepreneur has the option of not repaying the debt and searching for a new financier after a temporary exclusion from credit markets. The optimal contract specifies credit terms that gradually improve over time. Intuitively, by backloading firm claims to future cash flows, the bank can minimize the contract distortions due to the participation constraint of the entrepreneur and therefore extend more credit throughout the lending relationship.

The second building block of the model describes the problem of credit origination preceding the contracting stage. I consider a frictional meeting process modelled through directed search, as in Moen (1997), where heterogeneous banks compete for borrowers by posting long-term credit offers. Banks differ with respect to their funding costs and optimize over the offered contractual terms by taking into account the trade-off between loan profitability and the probability of attracting unfunded borrowers. The nature of this trade-off is endogenously determined through bank entry and the ratio of the number of credit opportunities to the number of applications. The introduction of search frictions delays the formation of lending relationships. More importantly, it endogenizes contractual terms at origination and firm outside option and characterizes the degree of competition in credit markets.

The interaction between agency and search frictions induces credit market conditions to directly affect firm default incentives. It therefore shapes the dynamics of optimal contracts and the transmission of shocks across borrowers and lenders. The analysis exhibits differences between effects at both micro and macro levels. At the bank level, search frictions limit the access to credit for defaulting firms, ease the agency problem, and hence allow for larger credit availability. At the aggregate level, however, this slows down the creation of new lending relationships and can consequently lead to lower total credit supply.

In the second part of the paper, I evaluate whether this mechanism is a meaningful source of persistence in credit markets. I consider two types of aggregate shocks: a productivity shock and a bank funding cost shock. I show that a negative (positive) shock to firm productivity (bank funding costs) can cause a significant decline in credit supply along both intensive and extensive margins. The effect on the intensive margin is short-lived and is directly driven by the diminishing returns to production. On the other hand, shocks are propagated along the extensive margin of credit as they negatively impact the stock of lending relationships and the number of producing firms in the economy.

In the cross-section, the model allows us to study how aggregate shocks impact the real sector. The analysis reveals an asymmetric treatment between funded and unfunded firms. The banking sector provides insurance against aggregate shocks in the economy and helps smooth out the credit availability

and pricing profiles of ongoing borrowers. The extent of the pass-through depends on the bank's health and the length of the lending relationship. However, unfunded firms are not shielded from shocks. They not only face limited access to lenders, but also experience a sharp decline in credit availability and rising borrowing costs once matched.

Finally, this paper has important policy implications. The model provides a better understanding of the credit reallocation process and is therefore particularly relevant when analyzing the effects of policies targeted toward business lending and banking regulations. Significantly, I show that a policy subsidizing the cost of credit origination - while being effective at incentivizing banks to expand their lending supply in the long-run - can in fact be counterproductive in the short-run.

The model integrates relationship banking to the macro-finance literature. To my knowledge, this is the first paper to jointly study the implications of long-term financing contracts and search frictions. The standard paradigm in the literature studying aggregate implications of financial frictions - starting with Bernanke and Gertler (1989) and Kiyotaki and Moore (1997) - relies on simple one-period interactions between anonymous borrowers and lenders.[4] My paper departs from this line of research along two key dimensions by constructing a dynamic equilibrium model that takes into account the process of formation of bank-firm pairs and the long-term nature of financing contracts. Moreover, by considering repeated interactions between borrowers and lenders, the model allows for endogenous borrowing limits which depend on the history of the lending relationship and credit market conditions.

This novel approach to credit markets emphasizes the importance of relationship banking at the aggregate level.[5] In particular, it is related to Allen and Gale (1997) and Berlin and Mester (1999) who highlight the role of banks as providers of intertemporal insurance for long-term borrowers. The paper is also connected to Bolton et al. (2014) who analyze the difference between relationship and transaction lending during normal and crisis times within a 3-period setting.

From a modelling perspective, my paper builds on the literature of long-term financing contracts in which credit constraints emerge endogenously as a feature of the optimal contract design.[6] Specifically, it draws on insights from Albuquerque and Hopenhayn (2004) and departs from existing literature by constructing a dynamic general equilibrium model which endogenizes the firm value at origination and its outside option and allows for aggregate and idiosyncratic shocks. The focus on the aggregate implications of long-term financing contracts is also shared with Cooley et al. (2004), Jermann and Quadrini (2007), and Monge-Naranjo (2008).[7] Cooley et al. (2004) study a general equilibrium model with limited contract enforceability and analyze how aggregate shocks to technological innovation can be amplified in the absence of market exclusion. In a similar vein, Jermann and Quadrini (2007) investigate

---

[4]Brunnermeier et al. (2012) provide an excellent survey of this literature.

[5]Petersen and Rajan (1995), Boot and Thakor (2000), and surveys in Boot (2000) and Degryse et al. (2009).

[6]Gertler (1992), Thomas and Worrall (1994), Albuquerque and Hopenhayn (2004), Quadrini (2004), Clementi and Hopenhayn (2006), De Marzo and Sannikov (2006), De Marzo and Fishman (2007), and Biais et al. (2010).

[7]Other recent papers also related to the limited-commitment contracting model of Albuquerque and Hopenhayn (2004) include Schmid (2012), Rampini and Viswanathan (2013), and Li et al. (2015).

how these contracts shape the economy's response following a stock market boom and productivity gains. Monge-Naranjo (2008) examines the effects of changes in interest rates, but takes the firm's outside option as exogenous. In contrast, my paper considers the joint aggregate implications of limited enforceability and search frictions in a general equilibrium setup, and examines how market conditions endogenously affect the dynamics of aggregate credit supply.

While a large literature has extensively studied the importance of agency problems in credit markets, little is known about the role of search frictions in this context. Previous work (Diamond (1990), Den Haan et al. (2003), Wasmer and Weil (2004), Becsi et al. (2005)) mainly considers static random search environments with simple contracts.[8] My paper shares some insights with Den Haan et al. (2003), who highlight the lasting damage due to the joint effects of the destruction of credit relationships and coordination failure in investment decisions. In contrast, I provide a richer and dynamic setup by embedding long-term financing contracts within a directed search equilibrium. The property of block-recursivity characterizing this equilibrium provides a numerically tractable solution and allows for the introduction of aggregate shocks and the analysis of the economy's transitional dynamics.[9]

The paper also belongs to the growing theoretical literature studying the interaction between search and agency frictions. These studies have mostly focused on labor markets as in Rudanko (2009), Moen and Rosén (2011) and Lamadon (2014) and goods markets as in Guerrieri et al. (2010). In this paper, I develop a credit markets model with limited contract enforceability and search and show how the interaction between these frictions can provide novel insights on the dynamics of credit along both intensive and extensive margins and lending rates.

Finally, this paper complements the emerging theoretical research explaining credit market freezes motivated by the recent financial crisis. This includes Bebchuk and Goldstein (2011), who show how coordination failure among financial institutions can lead to self-fulfilling credit contractions. Diamond and Rajan (2011) argue that the reluctance to extend credit is related to banks' fear of future fire sales. Benmelech and Bergman (2012) show that the interplay between financial frictions, market liquidity, and collateral values can give rise to credit traps and analyze the credit channel transmission of monetary policy. Philippon and Schnabl (2013) examine the problem of efficient recapitalization when banks are subject to debt overhang. In contrast, my paper provides a novel "flow-driven" theory focusing on frictions hindering the accumulation of relationship capital.

The rest of the paper is organized as follows. Section 2 describes the model. Section 3 characterizes the optimal contract, while section 4 analyzes the equilibrium properties of credit markets. Section 5 discusses comparative statics and testable predictions. Section 6 analyzes the quantitative properties of the model. Section 7 provides additional extensions and comments, and section 8 concludes.

---

[8]Inderst and Müller (2004) and Silviera and Wright (2006) also analyze the role of search frictions within models of venture capital. See also Petrosky-Nadeau and Wasmer (2011) and Petrosky-Nadeau (2014) who introduce search frictions in multiple markets simultaneously.

[9]Menzio and Shi (2011).

2 The model

In this section, I present a full equilibrium model of credit markets with search frictions and dynamic contracting. The contracting problem relies on limited commitment as in Albuquerque and Hopenhayn (2004) and is embedded within a directed search equilibrium, featuring aggregate and bank-specific uncertainty, and where heterogeneous financiers compete for borrowers by posting long-term contract offers. Optimal contracts are history-dependent and specify both the size of the loan and the corresponding lending rate throughout the credit relationship. The frictional meeting environment generates in equilibrium a continuum of submarkets in which borrowers and lenders meet, and where the optimal contract offered by each bank trades off its profits with the probability of matching with a borrower. In equilibrium, unfunded entrepreneurs are indifferent across all active submarkets, and the matching probability depends on the number of firms and banks present within each submarket. Proofs are presented either within this section or in Appendix A.

2.1 Environment

The model is in discrete time with infinite horizon. The economy is populated by two types of infinitely-lived agents: Entrepreneurs and Bankers. The mass of entrepreneurs is normalized to one, while the mass of active bankers is subject to endogenous entry and exit.

Agents and preferences
Both agents share the same discount factor $\beta \in (0, 1)$. Bankers are risk-neutral. Entrepreneurs on the other hand are risk averse and maximize their expected lifetime utility $\mathbb{E}_0[\sum_{t=0}^{\infty} \beta^t u(d_t)]$. d_t are the firm's net profits per period, and the flow utility $u : \mathbb{R} \to \mathbb{R}$ satisfies standard regularity conditions: $u' > 0$, $u'' < 0$, $\lim_{d \to 0} u'(d) = \infty$, and $\lim_{d \to \infty} u'(d) = 0$. The assumption of risk averse agents departs from the literature on dynamic debt contracting which typically analyzes the case where both agents are risk-neutral.[10] This is justified for firms where managers derive their consumption from their business venture, without any ability to diversify firm-specific risk (Stulz (1984)).[11] Entrepreneurs are also assumed to be hand-to-mouth and do not have access to a storage technology.

Technology and shocks
Each entrepreneur has access to a production technology subject to stochastic productivity shocks. However, she is initially cashless, and has to seek out external financing in order to start production. When funded with working capital K, a project can generate gross revenue $F(z, K) = zf(K) + (1-\delta)K$,

---

[10]Marcet and Marimon (1992) and Thomas and Worrall (1994) also allow for risk averse agents.
[11]This is also the case for larger corporations actively engaged in risk management policies as in Froot et al. (1993) or Rampini and Viswanathan (2010).

where z is the realization of the aggregate productivity shock. The gross revenue function F takes into account both current cash flows and capital depreciation at rate δ. The function f is differentiable, strictly increasing and strictly concave in capital, and satisfies $f(0) = 0$, $\lim_{k \to 0} f_k(k) = +\infty$, and $\lim_{k \to \infty} f_k(k) = 0$. The realization of the aggregate productivity $z \in \mathbb{Z} = \{z_1, z_2, ..., z_{N_z}\}$ is publicly observed every period, and follows a Markov process with transition probability $\Gamma_z : \mathbb{Z} \times \mathbb{Z} \to [0, 1]$. For simplicity, the production function abstracts from labor input.

The repeated interaction between the same borrowers and lenders can alleviate the agency friction in place. Banks arise in this economy because they are able to originate and commit to long-term relationships at a cost that is lower than that incurred by a repeated sequence of short-term interactions with direct monitoring. When matched, a given bank i acts as an intermediary channelling funds from depositors to entrepreneurs at a funding cost $r_d^i = \bar{r} + s^i$. The aggregate stochastic component $\bar{r} \in \mathbb{R} = \{r_1, r_2, ..., r_{N_r}\}$ with transition probability $\Gamma_r : \mathbb{R} \times \mathbb{R} \to [0, 1]$ corresponds to the aggregate state of the banking sector, while the bank-specific spread $\{s^i\}_i \in \mathbb{S} = \{s_1, s_2, ..., s_{N_s}\}$ follows independent Markov processes with transition probability $\Gamma_s : \mathbb{S} \times \mathbb{S} \to [0, 1]$. Heterogeneity in bank funding costs can be motivated by differences in deposit technologies and competition across deposit markets, bank size and economies of scale, Too-Big-To-Fail subsidies, and the ability to access interbank lending and repo markets.[12] Since the focus of this paper is on bank lending behavior and its effects on the real economy, bank liability structure is modelled in a parsimonious way and abstracts from potential feedback effects between the bank's asset quality and its funding. Each banker offers a long-term credit contract (specified explicitly below) and serves one entrepreneur at a time. Hence, at any given point in time, active bankers can either be part of a lending relationship or seeking a borrower.[13]

Credit markets

Credit markets are decentralized and subject to search and matching frictions. This assumption captures the 'localized' nature of bank lending markets (Agarwal and Hauswald (2010)) within which banks exert a certain degree of market power. Because this meeting process is constrained by costly search, this environment creates situations of bilateral monopoly between borrowers and lenders, and therefore determines endogenously the degree of competition in credit markets.[14] This allows for a richer and more realistic setting, as opposed to the standard cases of perfect or monopolistic bank competition.

Bank credit markets are modeled as a competitive search market, where bankers advertise contract offers, and entrepreneurs direct their search toward certain offers.[15] The origination process is costly and banks incur cost c whenever they enter the credit market. This parameter captures bank operating costs

---

[12]This is also consistent with evidence in Berlin and Mester (1999) and Gilchrist et al. (2013).

[13]A bank can therefore be thought of as a collection of bankers (or loan officers) with a loan origination technology that is constant-return-to-scale.

[14]These costs can be either explicit (transaction costs, time spent to prepare an application), or implicit (opportunity cost foregone if the firm stays unfunded).

[15]Here, banks cannot renege their initially posted offer and bargain with the firm at the meeting stage.

incurred before or at loan origination, and represents the cost associated with screening applications, loan officer wages, real estate, and advertisement of bank services. Credit markets are organized as a continuum of submarkets (or market segments), consisting of a subset of borrowers and lenders, and are indexed by the firm value V derived from the contract. The matching function $m(u,v)$, taking in mass of entrepreneurs u and bankers v, implicitly accounts for congestion and coordination externalities among borrowers and lenders. It also captures the bank screening technology and its lending standards in reduced-form. The function m is assumed to be continuous, concave, and homogeneous of degree one in both variables. We denote $p = m(u,v)/u = m(1,\theta) = p(\theta)$ to be the probability of getting access to credit (i.e. the bank's approval rate), and $q = m(u,v)/v = q(\theta)$ to be the probability of a bank locating a borrower, where $\theta = \frac{v}{u}$ is the usual credit market tightness. The functions p and q are assumed to be twice continuously differentiable, with p strictly increasing and concave, q strictly decreasing and convex, and $p \circ q^{-1}$ strictly concave. Moreover, when the credit market tightness tends to zero, the probability of finding a loan and finding a borrower tend to 0, and 1 respectively: $\lim_{\theta \to 0} p(\theta) = \lim_{\theta \to \infty} q(\theta) = 0$. Conversely, when the credit market tightness tends to infinity, the probability of finding a loan and finding a borrower tend to 1, and 0 respectively: $\lim_{\theta \to \infty} p(\theta) = \lim_{\theta \to 0} q(\theta) = 1$.

Long-term credit contracts

Borrowers and lenders sign a state-contingent long-term credit contract upon matching. The value of the contract at origination is determined by the aggregate state of the economy and the characteristics of the search market. A lending relationship, associated with current bank funding cost r_d, capital K, and aggregate productivity z generates the following per-period surplus:

$$S(z, r_d, K) = F(z, K) - (1 + r_d)K = zf(K) - (\delta + r_d)K$$

While lenders are assumed to fully commit to established long-term contracts, borrowers on the other hand are subject to limited enforceability. In the event of a default, borrowers can walk away with a fraction η of the existing capital stock. The diverted amount of capital is assumed to be consumed in the same period, and cannot be used for future production.[16] This however triggers the severance of their lending relationship and a temporary exclusion from credit markets. Defaulting borrowers only regain access to credit markets to seek out new financing with a "fresh start" probability $\xi \in [0, 1]$. This punishment captures the cost of bankruptcy and reflects legal and institutional systems in place.[17] The extreme cases are obtained for $\xi = 0$, or 1. When $\xi = 1$, firms are immediately allowed to get back to

---

[16]This assumption can be relaxed at the expense of increasing the dimensionality of credit submarkets. Here, it greatly simplifies the exposition because defaulting and unfunded firms have the same initial conditions and objective and hence search for credit in a similar way. Note also that the entrepreneur is not allowed to directly switch lenders, nor have multiple lending relationships. These aspects can be introduced in a richer version of the model.

[17]Bebchuk (2000) relates this cost to the length of time spent on bankruptcy procedures. Efficient court ruling, fast liquidation procedures, and a short period of discharge can therefore provide entrepreneurs with the opportunity to move forward and potentially start up a new business venture faster (see Peng et al. (2010)).

the credit markets for a fresh start (no market exclusion). Conversely, when $\xi = 0$, firm default leads to a permanent exclusion from credit markets.[18] The value of contract repudiation, which represents the firm's outside option, is given by the sum of contract-specific and aggregate components: (i) the utility value of diverting fraction η of existing capital stock, and (ii) the discounted expected firm value after default $\{H(z)\}_z$:

$$V^O(z, K, W) = u(\eta K) + \beta H(z) \tag{1}$$

$$H(z) = \xi \mathbb{E}_z[W(z')] + (1 - \xi)(u(d_0) + \beta \mathbb{E}_z[H(z')]), \tag{2}$$

where $\{W(z)\}_z$ is the firm value (determined in equilibrium) and d_0 is the per-period "garage" production when unfunded.

In order for the contract to remain enforceable along the equilibrium path, the current firm value generated through the lending relationship should always be at least as high as the utility derived from repudiation. The outside option and the threat of off-equilibrium credit market exclusion discipline the entrepreneur's incentives and shape the dynamics of credit made available to firms. Moreover, the firm value after default H provides an endogenous link between firm dynamics and credit market conditions. These effects will be explained in the next section with greater detail. Notice that full enforceability is a special case of the model, obtained when $V^O(z, K) = -\infty$ (i.e. both $\xi = \eta = 0$).

Let $\omega_t^\tau = \{(z_t, r_{d,t}), (z_{t+1}, r_{d,t+1}), .., (z_\tau, r_{d,\tau})\}$ denote the history of shocks associated with an ongoing lending relationship starting at date t in state $\omega_t = (z_t, r_{d,t})$, and still in place up to period τ. A long-term debt contract is a set of policies for working capital K_τ and payout d_τ:

$$\mathcal{C}(\omega) = \{(K_\tau(\omega_t^\tau), d_\tau(\omega_t^\tau), \ \forall \ \omega^\tau, \ \tau = t, .., \infty \ s.t. \ \omega = \omega_t\}.$$

Timing

An unfunded firm first searches for capital in the credit markets. Upon matching, the entrepreneur signs a long-term financing contract offering a certain firm value V. At the beginning of each period, borrowers and lenders observe the realized aggregate productivity z, and bank funding cost r_d. With probability σ, the lending relationship is exogenously terminated. In this case, the entrepreneur loses her firm as capital is liquidated, becomes unfunded and needs to start searching for new sources of credit. The bank on the other hand receives 0 until it matches with a new borrower.[19] Otherwise, with probability $1 - \sigma$, the lending relationship continues and the bank offers working capital K as prescribed by the contract.[20]

---

[18]ξ is constant in this setup. This assumption can be relaxed by for example allowing ξ to depend on loan size.

[19]The bank does not receive any income once the lending relationship is severed. In addition, the search for a new borrower is costly, and would also deliver 0 profits ex-ante due to the bank free-entry condition.

[20]Note that the model abstracts away from the firm's ability or willingness to switch lenders. While such an aspect would indeed enrich the model and the contractual space, this does not seem to be a first-order effect in banking. This is

Figure 1 displays the model timeline which comprises of two main stages each divided in two steps: first is the origination stage (extensive margin of credit) involving (i) bank entry, and (ii) credit contraction and expansion (screening and matching); second is the dynamic contracting stage (intensive margin of credit) involving (iii) capital intermediation, and (iv) and firm production or default.

- *Bank entry:* following the realization of aggregate shocks, banks decide to enter credit markets and pay c.

- *Credit contraction/expansion:* following the realization of the idiosyncratic shock, lenders with high funding costs may chose to exit if the contract value is negative. The remaining entrants (with positive contract value) optimally choose the submarket and firm value attached to the loan offers they post. Unfunded entrepreneurs search for credit opportunities within these active submarkets.

- *Capital advanced:* Active banks (either incumbent or newly established lending relationships that did not experience the exogenous destruction shock σ) offer the contractual amount of capital K^*.

- *Firm default/production decision:* In case of default, entrepreneurs divert and consume fraction η of capital K^*, and face a per-period probability of exclusion $1 - \xi$ from credit markets. In case of continuation, they pay back interest rate r^* as prescribed by the contract, and consume d^*.
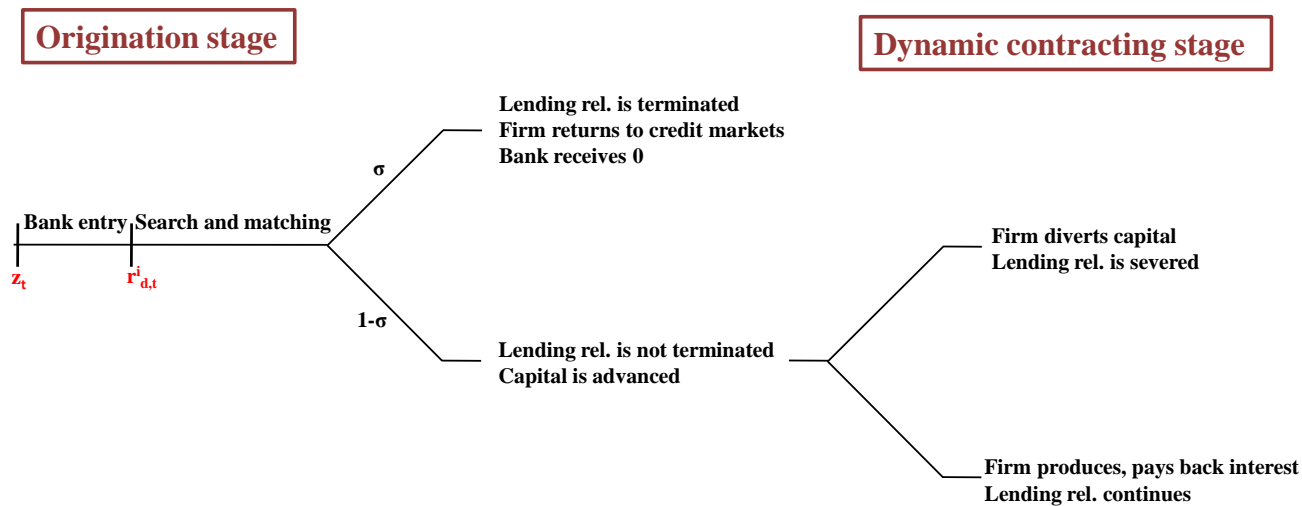

Figure 1. Timeline.

---

particularly true for small firms which typically face high switching costs. See Klemperer (1995) for a theoretical survey on switching costs.

3 Optimal Contracts

In this section, I describe and characterize the optimal lending contract in a partial equilibrium setting, taking as given both firm value V at origination and the vector associated with unfunded firm value $\{W(z)\}_z$ in each aggregate state. These values will be endogenized later once I characterize the credit markets through the joint search equilibrium.

3.1 Intuition

Let us develop the general intuition behind this contracting problem before moving to the analytical characterization. First, the distortion in the credit contract stems from two key ingredients: limited enforcement and banks sunk costs at origination. With full commitment on both sides, the allocation of credit is always optimal. However, a sufficiently low firm value can initially distort credit allocation in the presence of limited enforcement. In this case, receiving the first-best level of capital may induce the borrower to run away if the associated outside option becomes larger than the current value offered by the contract. The bank therefore adjusts its credit allocation downwards, initially constraining the firm, so as to ensure that the contract remains enforceable.

Throughout time, as the firm value increases and the borrower's participation constraint gets relaxed, credit availability also increases up until it reaches the first best allocation level. The dynamics of the contract and the speed at which the firm becomes unconstrained are shaped by the firm's outside option. This depends on the share of capital potentially diverted, the curvature of the utility function and the firm value after default H which reflects credit market conditions.

Why does the firm value start at a low level? The firm value at entry depends negatively on the bank's origination costs. In order for the bank to make credit offers in the first place, the contract must deliver a benefit that is at least as high as the origination cost. The magnitude of this cost is determined in equilibrium and depends on bank origination cost parameters and the credit market tightness conditions. The higher is this cost, the lower is the surplus allocated to the firm and the larger is the contract distortion at origination.

3.2 Contracting problem

The optimal contract maximizes the expected discounted payments to the bank, subject to the promise-keeping, enforcement, and limited liability constraints. I first write this contractual problem in its recursive form using the firm value V as a state variable in the spirit of Spear and Srivastava (1987) and Abreu et al. (1990). To simplify notations, the dependence of the continuation value $\{V'\}_{z',r_{d'}}$ on (z', r_d') is implicitly considered in the following recursive formulation:

$$B(z, r_d, V) = \max_{K, d, \{V'\}} f(z, K) - d - (\delta + r_d)K + \beta \mathbb{E}_{z, r_d} \left[(1 - \sigma)B(z', r'_d, V') \right] \quad (3)$$

subject to

$$V = u(d) + \beta \mathbb{E}_{z, r_d} \left[(1 - \sigma)V' + \sigma W' \right], \qquad \text{(Promise-Keeping)}$$

$$V^O(z, K; W) \leq V, \qquad \text{(Participation)}$$

$$d \geq 0. \qquad \text{(Limited Liability)}$$

The control variables are working capital K, firm profits d, continuation values $\{V'\}_{z', r_{d'}}$. The promise-keeping constraint represents the bank's full commitment to deliver value V. Delivering this value can be decomposed between today's utility from payout $u(d)$, and the discounted promised value: $\beta \mathbb{E}_{z, r_d} \left[(1 - \sigma)V' + \sigma W(z') \right]$, which takes into account the possibility of separation. The second inequality is the participation constraint. In order for the contract to be self-enforcing, this constraint requires that V is always larger than the firm's outside option. Eventually, the contract assumes the firm's payout to be non-negative reflecting the entrepreneur's limited liability. As in Cooley et al. (2004), this assumption is justified by the fact that entrepreneurial consumption cannot be negative, especially when all of the entrepreneur's assets are inside the firm.[21]

3.3 Recursive multiplier formulation

The forward-looking nature of the enforcement constraint and the existence of both aggregate and idiosyncratic shocks make this problem difficult to solve using standard dynamic programming techniques. Here, instead of solving the program in the value space, I adapt the methodology developed in Marcet and Marimon (2011) and its earlier versions, and rewrite the problem in its recursive Lagrangian form, where Λ is the current relative weight associated with the entrepreneur's value and λ is the Lagrange multiplier associated with the enforcement constraint. By introducing the expectational constraint directly into the objective function, this formulation circumvents the time-inconsistency issue present in the original maximization problem since it imposes a law of motion followed by the relative weight Λ associated with firm value V.[22]

Given the model assumptions, Theorems 1 and 2 in Marcet and Marimon (2011) justify the equivalence between the original problem and its recursive multiplier counterpart when separation is exogenous. This formulation will be extremely useful in order to characterize the main properties of the optimal contract and compute the numerical solution to this problem.

---

[21] For example, the outside investment opportunities are assumed to yield lower returns and hence are not held in equilibrium.

[22] Cooley et al. (2004) also uses this technique to solve a similar problem.

Proposition 1. *The maximization problem is equivalent to the saddle-point problem (SPFE):*

$$P(z, r_d, \Lambda) \quad = \inf_\lambda \sup_{K,d} \quad f(z, K) - d - (\delta + r_d)K - \lambda \left[u(\eta K) + \beta H(z) \right]$$
$$+ (\Lambda + \lambda) \left[u(d) + \beta \sigma E_z[W(z')] \right]$$
$$+ \beta(1 - \sigma) E_{z, r_d} \left[P(z', r_d', \Lambda + \lambda) \right]$$

subject to

$$d \geq 0, \quad \lambda \geq 0.$$

The value of the cumulative Lagrangian Λ is strictly increasing - independent of shocks - as long as the firm's participation constraint is binding (i.e. $\lambda > 0$). Its law of motion is given by $\Lambda' = \Lambda + \lambda$, and can be interpreted as an additional condition that the planner imposes in order to follow the optimal path. From the envelope condition, it is straightforward to show that the marginal cost of a one unit increase in firm value is equalized across all continuation states (z, r_d):

$$\frac{\partial P(z, r_d, V(z, r_d))}{\partial V} = -\Lambda \tag{4}$$

3.4 Characterization

Capital policy. The choice of optimal capital K is static. Indeed, K solves the following intermediate problem, which only depends on current firm value V and shock realizations:

$$\pi(z, r_d, V) = \quad \max_K f(z, K) - (\delta + r_d)K \tag{5}$$
$$\text{s.t.} \quad u(\eta K) + \beta H(z) \leq V$$

This problem generates a constrained region where the enforcement condition binds and the level of working capital is suboptimal, and an unconstrained region, where the first-best level of capital K_{FB} is defined for each state (z, r_d) as:

$$K_{FB}(z, r_d) = \arg\max_K f(z, K) - (\delta + r_d)K$$

Let us also define $\bar{V}(z, r_d)$ as the lowest continuation value associated with a non-binding participation constraint:

$$\bar{V}(z, r_d) = u(\eta K_{FB}(z, r_d)) + \beta H(z),$$

and the constrained level of capital $\bar{K}(z, V)$, satisfying $V^O(z, \bar{K}, W) = V$ for $V < \bar{V}(z, r_d)$.[23]

---

[23]Note that for each state (z, r_d), the upper bound $\bar{\Lambda}(z, r_d)$ is reached whenever the firm value is given $\bar{V}(z, r_d)$, and is

Proposition 2. *The firm's optimal capital is characterized by:*

$$K^*(z, r_d, V, W) = \begin{cases} \bar{K}(z, V, W), & if \quad V < \bar{V}(z, r_d), \\ K_{FB}(z, r_d), & if \quad V \geq \bar{V}(z, r_d) \end{cases}$$

Corollary 1. *K is decreasing in ξ, η and W in the constrained region.*

The above results are intuitive since the limited commitment problem becomes more severe as the outside option increases. Given that an increase in either the share of capital diverted η, the probability of fresh start ξ, or the unfunded firm value W implies a larger outside option, the financier lowers the size of capital advanced to prevent the entrepreneur from running away. The link between the firm's equity value, borrowing capacity, and market conditions is apparent from the above expression. The higher is the firm value V, the higher is its borrowing capacity K. Conversely, the higher is the competition in credit markets (high W), the lower is this borrowing capacity. Note also that two firms with same value V, but linked to different bank types $r_{d,1} < r_{d,2}$ receive the same amount of credit as long as $V < \bar{V}(z, r_{d,2})$.

Dividend payout policy. The following proposition shows that the dividend payout d exhibits downward rigidity: in other words, d is never decreasing, and either increases whenever the outside option is binding ($\lambda > 0$), or stays constant once the unconstrained region is reached.

Proposition 3. *Dividend payout is increasing over time whenever the firm's outside option is binding.*

This result is driven by the following envelope condition which exhibits the positive relation between d and V and its relation to the cumulative Lagrangian Λ':

$$\frac{1}{u'(d)} = -\frac{\partial P(z', r_d', V_{z', r_d'})}{\partial V} = \Lambda' \tag{6}$$

The dividend payout is increasing in Λ'. A higher degree of risk aversion is accompanied with a smoother path of dividend payouts.[24] This contrasts with the risk-neutral case, where the entrepreneur is indifferent about the timing of her consumption, and where it is optimal to reinvest all proceeds in the firm in order to grow out of the constrained region faster. In this case, the dividend payout is 0 throughout the constrained region, and jumps as soon as the firm becomes unconstrained. In the

---

given by: $\frac{\partial \Pi(z, r_d, \bar{V}(z, r_d))}{\partial V} = -\bar{\Lambda}(z, r_d)$.

[24]The back-loaded dividend dynamics are analogous to the ones obtained in standard dynamic contracting problems with limited commitment as in Harris and Holmstrom (1982), and where a risk-neutral financier is indeed providing some insurance to risk-averse borrowers. The difference vis-a-vis labor models with limited commitment (for example Rudanko (2009)) comes from the agency problem faced by borrowers, and which generates additional savings incentives, and hence different payout dynamics.

continuation region, the law of motion for dividends is given by

$$\frac{1}{u'(d')} = \frac{1}{u'(d)} + \lambda$$

3.5   Firm value dynamics

Now that we have characterized the relationship between equity value, working capital and dividend payout, we can turn to the analysis of the dynamic aspect of the contract. I first show that, in the absence of shocks, firm value is always increasing in the constrained region. This generalizes the result shown in Albuquerque and Hopenhayn (2004) for the risk-neutral case, and is related to Marcet and Marimon (1992), confirming that the incentives to save in order to outgrow the borrowing constraints are still present despite the consumption smoothing motive.[25] Second, I show that the introduction of shocks implies that firm value can now follow a non-monotonic pattern.

Proposition 4. *Fix* (z, r_d). *For a given firm value* V, V' *is increasing over time whenever* $V < \bar{V}(z, r_d)$.

Keeping aggregate and idiosyncratic shocks constant, firm value is increasing whenever the outside option is binding. In other words, the Lagrange multiplier increases whenever the firms outside option is binding. Hence, $V' > V$ since Π is decreasing in V. This result can be directly generalized to the expectation of promised values when the economy is subject to shocks.

The next proposition states that banks with relatively low-funding costs offer higher continuation values. This result is intuitive since lending relationships established with low-funding cost banks generate more surplus (the marginal product of capital is lower, and the optimal allocation of capital higher) and as a consequence allow for larger firm values.

Proposition 5. *Fix* z. *For a given firm value* V *and funding cost* r_d, $V'(r'_d, .)$ *is decreasing in* r'_d *whenever* $V < \bar{V}(z, r_d)$.

Figure 3 illustrates how working capital, dividends and promised values $\{V'\}$ depend on current firm value V and funding costs (holding z constant), in partial equilibrium.[26]

Effective intra-temporal interest rate on capital.   The contract implicitly specifies an effective intra-temporal interest rate charged to borrowers as a function of the optimal working capital and

---

[25]Marcet and Marimon (1992) also features a dynamic contracting problem with risk averse agents and limited commitment, for a different application.

[26]i.e. taking firm value V at origination and the vector associated with unfunded firm value $\{W(z)\}_z$ in each aggregate state as exogenous.

[28]The aggregate shock z is constant here. Note that promised value V' depends on both current and future funding costs. For simplicity, the promised value profile exhibited in the right panel only corresponds to the case where the current funding costs is at the average level.
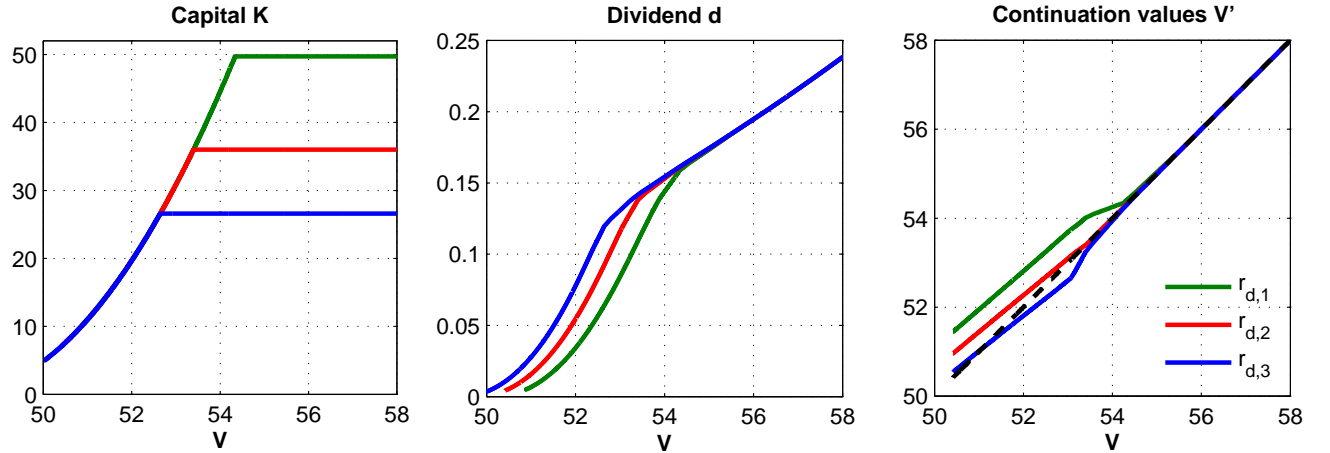
Figure 2. Contract policies as a function of V and funding costs (from low (blue) to high (black)): Working capital (K), firm payout (d), Promised Value (V') [28].

dividend payout policies and given by:[29]

$$r^*(z, r_d, V) = \frac{F(z, K^*(z, r_d, V)) - d^*(z, r_d, V)}{K^*(z, r_d, V)}$$

The dynamics of the firm's capital, dividend payout and value can eventually be summarized as follows. Upon matching at period t_0, the newly-formed lending relationship is started with an initial firm value V_0 (and equivalently an initial cumulative multiplier Λ_0).[30] At that stage, the firm is typically constrained and operates at a suboptimal scale, and Λ_t increases over time whenever the firm's participation constraint is binding. In this setup, entrepreneurs care about the time allocation of dividend payments since they are risk averse. The optimal contract therefore allows for positive dividend payouts even in the constrained region, in order to partially smooth the borrowers consumption profile. The lending relationship matures whenever V becomes sufficiently large to sustain the first-best level of capital across all future possible states. In that case, the firm is unconstrained and operates at its optimal scale.

How do credit market conditions affect contract dynamics? I address this question in the next section where both firm value at origination and after default are endogenized in the full equilibrium model. In this environment, the interaction between agency and search frictions generates two countervailing effects. On one hand, it affects the degree of bank competition and the share of surplus allocated to borrowers and lenders. For example, a high degree of search frictions offers banks more market power, decreases firm value, and hence further distorts the contracting problem. On the other hand, this also decreases the firm's outside option and relaxes the contracting space. As this will become clear,

---

[29]Note however that the implementation of the contract is not unique.

[30]This initial value depends on market conditions, and is determined endogenously in the next section.

the extent of credit misallocation and the speed at which firms become unconstrained depend on the dynamics of the wedge between current firm value at origination and after default.

4   Directed Search Equilibrium

I now introduce the full equilibrium version of the model in order to endogenize firm value at origination and the outside option. The interaction between agency and search affects the firm value in both funded and unfunded stages and generates novel implications. This framework is tractable and amenable to the introduction of heterogeneous agents and multiple shocks due to the property of block-recursivity. Let us first study the search behavior of borrowers and lenders separately.

Firms.   In order to resume production, entrepreneurs search for credit opportunities whenever they are unfunded. Let $\{W\}_z$ denote the firm value when projects are unfunded contingent on the aggregate state z. The Bellman equation for W satisfies:

$$
\begin{aligned}
W(z) &= u(d_0) + \beta\rho(z) + \beta E_z[W(z')] & (7)\\
\rho(z) &= p(\theta(z,V))(V - E_z[W(z')]), \quad \forall\, V & (8)
\end{aligned}
$$

where, d_0 denotes the "garage" production, $\{\rho(z)\}_z$, the added firm value associated with forming a lending relationship, and $\theta(z,V)$ is the credit market tightness in the submarket associated with aggregate shock z, bank type r_d, and firm value V. We can rearrange the two expressions above in order to establish the link between market tightness and contract value as follows:

$$
p(\theta(z,V)) = \frac{W(z) - u(d_0) - \beta E_z[W(z')]}{\beta(V - E_z[W(z')])} \tag{9}
$$

This equation defines a bijective mapping between firm value V and credit market tightness $\theta(z,r_d,V)$ within each submarket. Given that $p(.)$ is strictly increasing in θ, it follows that θ is strictly decreasing in V. This means that submarkets where banks offer low firm value (through a combination of high interest rates and low credit availability), are more liquid and enjoy higher approval rates (or lower lending standards). Conversely, submarkets where banks offer lending contracts with high firm value (e.g. low interest rates and higher credit availability), have lower credit market tightness, lower approval rates, and stricter lending standards. Eventually, a submarket offering a firm value below W cannot attract borrowers and are inactive. In equilibrium, firms are indifferent among the submarkets in which they search, meaning all active submarkets exactly deliver value $\rho(z)$.
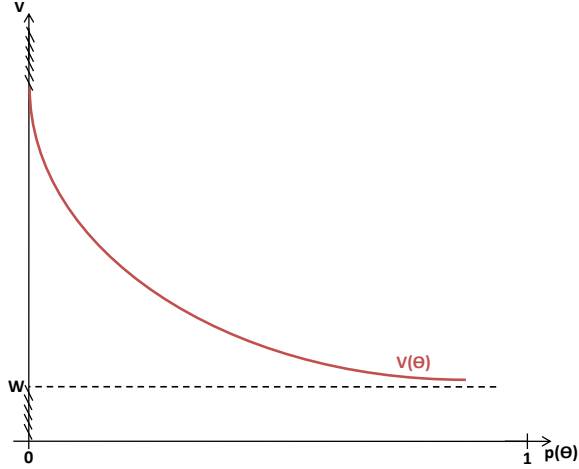
Figure 3. Relationship between approval rates and firm value.

Banks posting problem. Banks decide whether to enter credit markets after the realization of the aggregate shock, but before observing the idiosyncratic component of the funding cost. If the funding cost is not sufficiently low to warrant gains from trade (i.e. $B(z, r_d, W; W) \leq 0$, meaning the lending contract that can jointly deliver a firm value above W and positive bank profits), the bank exits immediately from credit markets and does not offer any lending opportunity. Otherwise, each bank optimally chooses the submarket that maximizes its expected profits taking into account the probability of finding a borrower:

$$\max_{V} q(\theta(z, V; W)) B(z, r_d, V; W) \tag{10}$$

Let us define the compact interval $\mathcal{S}_0 = [\underline{S}, \bar{S}]$, where $\underline{S} = \frac{u(d_0)}{1-\beta}$, and $\bar{S}$ is the maximum value obtained by the entrepreneur when the joint surplus from the match is entirely kept by the firm.

Lemma 1. *The solution $V^*(z, r_d; W)$ to the maximization problem (10) exists and is unique for each $(z, r_d, W) \in \mathbb{Z} \times \mathbb{R} \times \mathcal{S}_0$.*

Free entry. Because a loan offer is only made if it provides positive profits to the bank, we can define the ex-ante expected bank profits as:

$$B^*(z; W) \quad = \quad E_{r_d} \left[\max_{V} q(\theta(z, V; W)) B(z, r_d, V; W)]_+ \right], \tag{11}$$

and the conditions for free-entry and complementary slackness as follows:

$$c = B^*(z, W), \tag{12}$$

$$0 = \theta(V)[B^*(z, W) - c], \quad \forall V. \tag{13}$$

The free-entry condition (12) states that banks keep supplying new loan offers as long as the ex-ante profits from credit origination are at least equal to the cost c. Condition (13) is the standard complementary slackness condition which specifies the set of active ($\theta > 0$) and inactive ($\theta = 0$) submarkets. As this is common in this class of search models, the free-entry condition is critical in allowing for the equilibrium to be block-recursive and hence tractable. It also for the analysis of the equilibrium outside of the steady state. In particular, without this condition, the distribution of banks in the economy becomes critical in order to determine the tightness associated with each active submarket. This in turn dramatically increases the state space of the agents (see for example Krusell and Smith (1998)). In our case, the equilibrium market tightness is independent of the distribution of banks since agents already know that credit relationships are formed until the free-entry condition is satisfied with equality.

4.1 Equilibrium

4.1.1 Characterization of credit markets

The following lemma states that each bank chooses the unique submarket that maximizes its expected profits from the loan offer. There is therefore one active submarket associated with each bank funding level. Markets that offer higher value to their borrowers exhibit higher market tightness and lower approval rates.

Lemma 2.
(i) For c sufficiently small, a solution W satisfying the free entry condition (12) exists and is unique.
(ii) The value of the loan contract to the firm $\rho(z)$ is equalized across all active submarkets, and there is an optimal firm value $V^*(z, r_d)$ for all $(z, r_d) \in \mathbb{Z} \times R_d$ such that:
(a) An active submarket $V = V^*(z, r_d)$ satisfies $\theta(V, z) > 0$ and $p(\theta(V, z))(V - W(z)) = \rho(z)$.
(b) With the expected contract value to a bank in state (z, r_d) $a(V^*) = q(\theta(z, V^*))\Pi(z, r_d, V^*; W))$, we have

$$\theta(z, V) = \begin{cases} 0, & if \quad a \leq 0, \\ q^{-1}\left(\frac{a}{\Pi(z, r_d, V; W)}\right), & if \quad a > 0. \end{cases}$$

In equilibrium, all active submarkets offer the same ex-ante value ρ to borrowers, and borrowers are indifferent across all of them. Unfunded entrepreneurs keep entering a given submarket until the prob-

ability of finding a credit opportunity becomes so low that they would eventually prefer accepting less generous contractual terms from other banks. As a corollary, when the degree of bank competition increases (for example due to a higher matching elasticity or lower origination costs), the unfunded firm value W becomes so high that it may actually deter high-funding banks from entering credit markets. In this case, the distribution of banks supplying credit becomes more concentrated.

Proposition 6. *Credit markets in the cross-section:*

- $V^*(z, r_d)$ *and capital level at origination* $K_0(z, r_d)$ *decrease with* r_d.
- $p(\theta(z, V^*(z, r_d)))$ *increases with* r_d.

Figure 4 summarizes the main equilibrium properties of these markets. The introduction of search frictions provides a natural and realistic setting to describe the coexistence of multiple credit market segments. First, the equilibrium allows for the existence of multiple contracts offering different firm values, loan sizes and lending rates to their borrowers. Second, approval rates and firm values are inversely related. Banks with low funding costs offer high firm value, and attract borrowers faster.
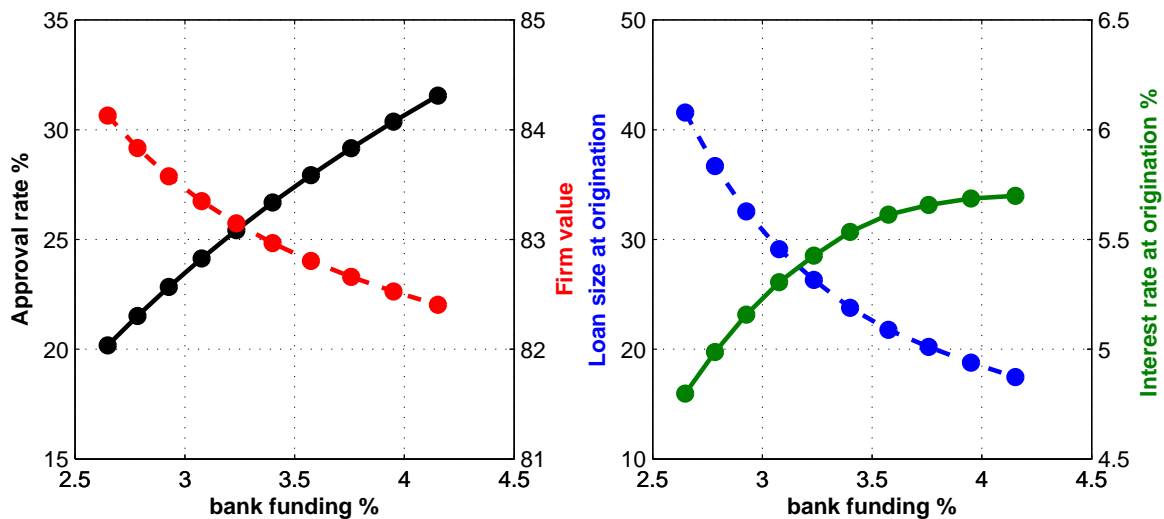

Figure 4. Approval rate and firm value at origination as a function of bank funding cost r_d.

Because of the longer queue length, firms applying to these banks face lower approval rate - or conversely stricter lending standards - relative to high-funding banks.[31] Because firms are ex-ante identical, the matching function can be interpreted as a reduced-form screening process with unobserved heterogeneity. Since firms matched with low-funding banks start with a relatively higher value, the wedge between firm value at origination and after default is larger and so is the initial level of capital.

---

[31]Considering potential economies of scale and the fact that the funding cost here reflects all costs incurred by banks, this result is consistent with industry practices which exhibit stricter lending standards for large banks relative to small and community banks.

I can now turn to the definition of the directed search equilibrium in the spirit of Moen (1997) and establish its existence.

Definition 1. *Directed Search Equilibrium*

A directed search equilibrium of the economy consists of: the value for unfunded $\{W(z)\}_z$ and funded $\{V(z, r_d)\}_{z,r_d}$ firms, market tightness $\{\theta(z, r_d)\}_{z,r_d}$, and lending contract policies $\{(K^*, d^*, V'^*)\}_{z,r_d}$, such that:

a. Credit search strategy of unfunded entrepreneurs maximizes their firm value. That is, the value provided by a lending contract is consistent with (7), and the relationship between this value and market tightness satisfies (9).

b. Bank lending policy is optimal. Given $\{W(z)\}_z$, and for all z and r_d, banks maximizes their profits, by solving problem (3).

c. Bank entry is consistent with (12), and is strictly positive for all z.

d. Measure of unfunded firms in the economy evolves according to:

$$v_{t+1} = v_t\big(1 - \int p(\theta(V^*(r_{d,t})))J_t(r_{d,t})dr_{d,t}\big) + \sigma(1 - v_t),$$

where $J_{t+1}(r_{d,t+1}, V_{t+1})$ is the measure of new bank entrants with firm value at origination V_{t+1}, and funding cost $r_{d,t+1}$.

Proposition 7. *A Directed Search Equilibrium exists for a sufficiently small entry cost c.*

Proposition 7 establishes the existence of a solution consistent with the definition of a directed search equilibrium. Note that this equilibrium solution is well-defined - in the sense that it allows for block-recursivity - only when bank entry is strictly positive across all possible histories of the aggregate shock.

4.2 Efficiency

In this section, I analyze the efficiency of the directed search equilibrium defined above. In order to do so, I explicitly introduce the social planner's problem. The social planner maximizes the discounted sum of utilities derived by banks and firms for incumbent lending relationships, utility derived by rationed entrepreneurs, less total entry costs incurred by loan origination. The problem is subject to the dynamics of the existing lending contracts represented by the function f_c (which only depends on (V_{t-1}, r_{t-1}, r_t)), and the laws of motion for credit rationing v_t, and the distribution of lending relationships g_t.

In order to simplify notations, notice that the social planner's faces the same contracting frictions as each individual bank, hence I can immediately replace the original problem with the corresponding

solution to its Lagrange multiplier formulation (SPFE), and by taking the optimal weights on firm value to be $\Lambda_{t+1} = \frac{1}{u'(d_t)}$.

The social planner therefore maximizes the following objective function:

$$\max_{u_t, g_t, \theta_V, J_t, V_t} \mathbb{E}_z \sum_t \beta^t \left[\sum_{r_{d,t}, V_t} (1-\sigma)g(r_{d,t}, V_t)[S(V_t, r_{d,t}) - d(V_t, r_{d,t}) + \Lambda(V_t, r_{d,t})u(d(V_t, r_{d,t}))] - cJ_t + v_t u(d_0) \right]$$

$$s.t. \ \forall \ (t, z^t)$$

$$\lambda(V_t, r_{d,t}) = \frac{1}{u'(d(V_t, r_{d,t}))}, \qquad\qquad \forall \ (r_t, V_t)$$

$$V_t = f_c(V_{t-1}, r_{t-1}, r_t), \qquad\qquad \forall (r_{t-1}, V_{t-1}, r_t)$$

$$v_t = v_{t-1}(1 - \sum_r \Gamma_r(r)p(\theta(V_r))) + (1 - v_{t-1})(1 - \sigma)$$

$$g_t(r, V) = \sum_{V_{t-1}|V_t = V} (1-\sigma)g_{t-1}(r_{t-1}, V_{t-1})\pi_r(r_{t-1}, r_t) + J_t q(\theta(V_r))\Gamma_r^0(r)1_{V_t = V_r}, \qquad \forall \ (r, V)$$

where Γ_r is the transition probability and Γ_r^0 is the unconditional distribution of the idiosyncratic funding cost.

Proposition 8. *The Directed Search Equilibrium is constrained-inefficient whenever entrepreneurs are risk-averse. When entrepreneurs are risk-neutral, the directed search equilibrium (satisfying positive entry) exists, is unique and delivers the efficient allocation.*

Proposition 8 shows the existence and uniqueness of a solution to the planner's problem and establishes that the corresponding allocation cannot coincide with that of the directed search equilibrium unless entrepreneurs are risk-neutral.

The inefficiency is due to the combination of risk-averse entrepreneurs and search frictions in credit markets and generalizes the results obtained by Acemoglu and Shimer (1999) and Golosov et al. (2013) in the context of labor markets. It states that credit-rationed entrepreneurs inefficiently choose to search in markets that offer low firm values but high approval rates. This inefficiency result is interesting and leaves room for policy intervention.[32]

5 Comparative statics and testable predictions: equilibrium effects and interaction between search and limited enforceability

This section illustrates the equilibrium effects generated by the key parameters governing contract dynamics and credit origination, namely the share of divertible asset η, the probability of fresh start ξ,

---

[32]A further theoretical investigation of this result and its implications for credit markets and bank lending is part of my research agenda.

the matching elasticity γ_m and entry costs c. I investigate how these parameters affect both intensive and extensive margins of credit. The interaction between search and limited contract enforceability operates through two distinct channels: firm value at origination and contract dynamics through the firm's outside option. In particular, the contracting parameters impact the extensive margin of credit in that a tighter constraint (i.e. higher η and ξ) prevents firms from quickly outgrowing their borrowing constraint, and hence limits the generated surplus. This affects bank entry, market tightness, matching probabilities, and the corresponding firm value at origination.

Conversely, the search parameters not only affect the extensive margin, but also shape the dynamics of incumbent lending relationships through their effect on the unfunded firm value W. For example, highly competitive credit markets (i.e. markets with low search frictions: low c or high γ_m) can create a further distortion in the contract (the contracting space becomes more restricted as the lending relationship becomes less 'exclusive'), which tightens the borrowing constraint and slows down the dynamics of credit growth.

5.1 Effects of contracting parameters η and ξ

Parameters η and ξ directly enter in the firm's outside option and have a similar qualitative impact on both extensive and intensive margins. As explained earlier, the difference between these two parameters resides in the fact that η affects the growth rate of firms as a function of their current loan levels, whereas ξ enters into the common component and affects the sensitivity of the credit growth rate to W (and hence the business cycle). Figure 5 shows how η impacts several variables of interest.

Share of divertible assets η. The share of divertible asset η captures the entrepreneur's default incentives. A higher diversion rate amplifies the agency problem and severely distorts the contract. As a consequence, the total surplus extracted from the match and firm value at origination are also lower.

Fresh start probability ξ. The fresh start probability ξ corresponds to the ability of the entrepreneur to return to credit market after a default. The inverse of ξ reflects the average period of exclusion and can be viewed as a measure of the degree of leniency of the bankruptcy code towards entrepreneurs. A higher probability of fresh start allows for a shorter exclusion from credit markets and distorts the contracting problem further.[33] This also captures the quality of legal institutions and the speed at which bankruptcy procedures are dealt with. While this process may take just a few months as in the U.S., it may take several years as in Mexico. The comparative statics are consistent with the studies of La Porta et al. (1997) and La Porta et al. (1998) which show that economies with entrepreneur-friendly bankruptcy laws typically exhibit lower access to credit.

---

[33]The comparative statics charts for parameter ξ is reported in the appendix.
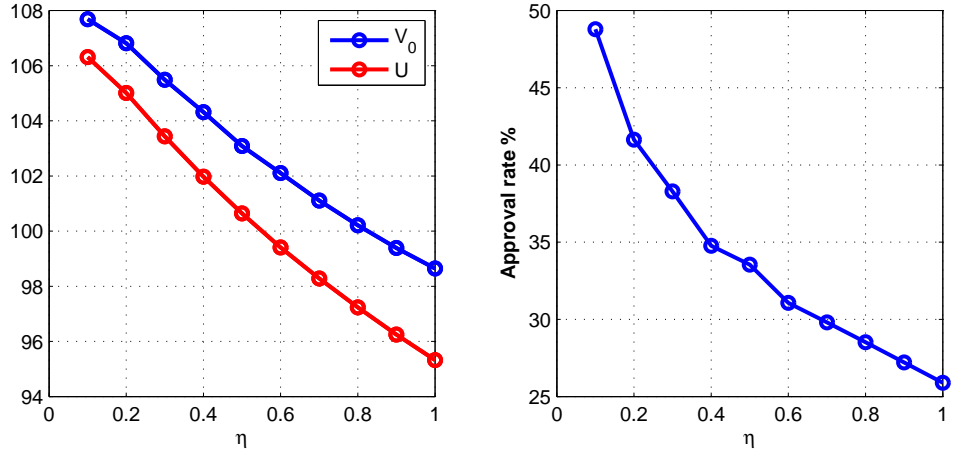
(a) Credit market variables - Firm value (at origination (blue), unfunded (red)) (left panel), and approval rates (right panel), for $\eta \in [0.1, 1]$.
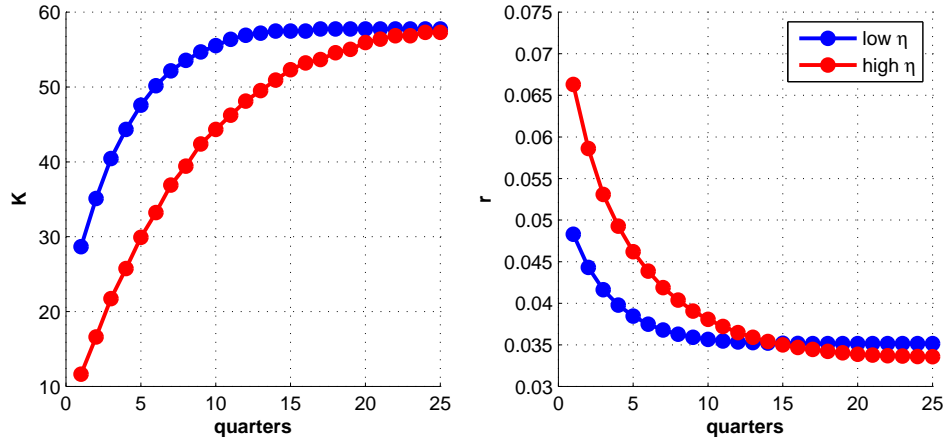
(b) Firm dynamics - Credit availability (left panel) and lending rates (right panel) across the length of the lending relationship, for low (blue) and high (red) levels of η.

Figure 5. Comparative statics - η.

5.2 Effects of search parameters γ_m and c

Higher matching elasticity γ_m and lower origination costs yield the same qualitative features of credit markets, namely, a higher approval rate and firm value. In this context, the degree of competition in credit markets is captured by the equilibrium market tightness θ.[34] In that sense, the model provides a richer environment that is amenable to the analysis of bank competition and therefore departs from more standards assumptions of perfectly competitive or monopolistic banks (as in Dixit-Stiglitz). This

---

[34] Inderst and Müller (2004) provides a similar interpretation.

setup generates an interesting pattern where increased competition implies low credit availability and high lending rates at origination. However, these effects are reversed as the lending relationship matures and the agency problem vanishes.

Matching elasticity γ_m.   The matching elasticity captures the degree of matching frictions in credit markets.[35] The following comparative statics show that markets with a high degree of competition typically generate higher approval rates and better access to lenders. This is consistent with evidence from the Senior Loan Officer Opinion Survey (SLOOS) conducted by the Federal Reserve Board where competition among lenders is often highlighted as one of the major reasons for easing lending standards.[36] The right panel of Figure 6b which highlights how lending rates vary throughout a relationship for both low and high degrees of competition, is also qualitatively consistent with the empirical results in Petersen and Rajan (1994)).[37]

Effect of origination cost c.   The impact of a negative change in c provides similar qualitative results as the ones described above. From the free-entry condition, we can see that a decrease in c is accompanied with lower unfunded firm value W in equilibrium and a higher approval rate.[38] This cost c can be interpreted in several ways. First, c can be viewed as an initial sunk investment needed to start up the firm's project. The above result can therefore provide grounds for potentially explaining why entrepreneurs with initially large level of wealth w_0 ($< c$) may have higher approval rates and larger credit amount at origination. Cost c can also be associated with non-interest expenses incurred by banks. These expenses typically cover loan officers wages, building and administrative expenses, in addition to other costs associated with the screening technology, and can also seriously hamper loan profitability and access to credit.

5.3   Aggregate implications

The model implications at the lending relationship level do not necessarily go through at the aggregate level because of the subtle role of search frictions in the model. In particular, the model exhibits a clear trade-off between intensive and extensive margins of credit. While a high degree of bank competition would indeed generate lower levels of credit supply at the bank-firm level (contracts are initially more distorted because V and W are relatively close), it also generates a higher rate of approval and therefore a higher level of lending relationship creation. Let us analyze these effects at the aggregate level.

[35]See next section for the specification of the matching function used throughout the paper.

[36]"Among domestic respondents that reported having eased either standards or terms on C&I loans over the past three months, the majority of banks cited more-aggressive competition from other banks or nonbank lenders as an important reason for having done so." - from the January 2014 Senior Loan Officer Opinion Survey on Bank Lending Practices: www.federalreserve.gov/boarddocs/SnloanSurvey/201402/.

[37]Note however that Petersen and Rajan (1994) looks at firm age instead of relationship length.

[38]This result is shown in Lemma 3 in the appendix.
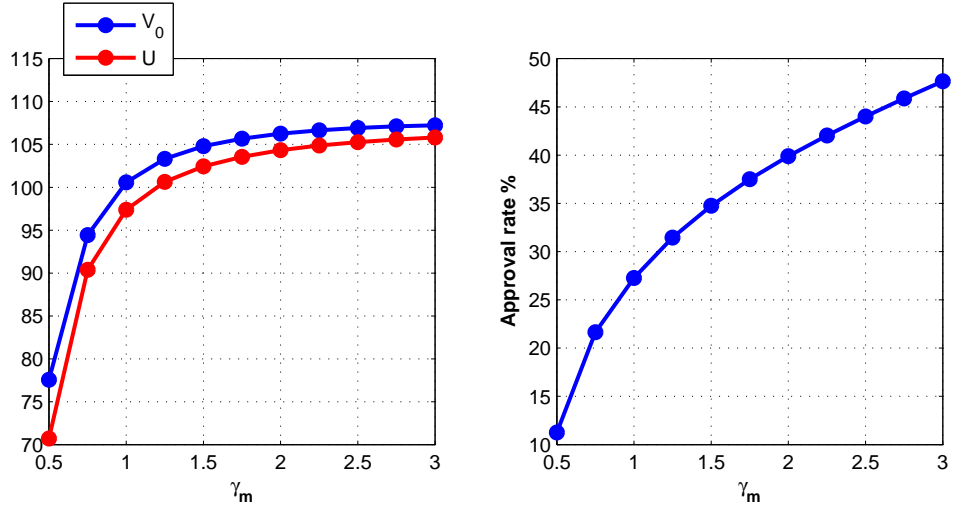
(a) Credit market variables - Firm value (at origination (blue), unfunded (red)) (left panel), and approval rates (right panel), for $\gamma_m \in [0.5, 3]$.
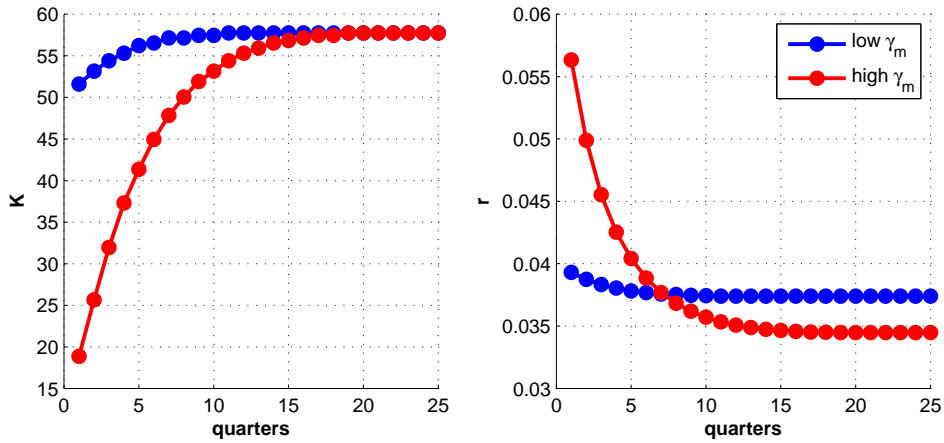
(b) Firm dynamics - Credit availability (left panel) and lending rates (right panel) across the length of the lending relationship, for low (blue) and high (red) levels of γ_m.

Figure 6. Comparative statics - γ_m.

Without loss of generality, let us explore the case of an economy abstracting from shocks and bank heterogeneity, and taking the utility derived from garage production to be 0. From equations (7)-(9) and the firm's participation constraint, we can write:

$$V_0 \;=\; \frac{1 - \beta(1 - p(\theta))}{\beta p(\theta)} W = u(\eta K_0) + \frac{\beta\xi}{1 - \beta(1 - \xi)}W$$

Rearranging terms from the equation above, we obtain the following identity for the total credit supply

due to new bank-firm relationships, which exhibits this tension between intensive and extensive margins:

$$\begin{aligned} \bar{K} &= p(\theta)K_0(\theta) \\ &= \frac{p(\theta)}{\eta}u^{-1}\left(\left[\frac{1}{\beta p(\theta)} + \frac{1}{1-\beta(1-\xi)}\right]W\right) \end{aligned} \tag{14}$$

where K_0 corresponds to the loan size at origination.

Given the model assumptions and previous results, it is straightforward to show that when bank competition increases (either through an increase in matching elasticity, or lower entry costs), the extensive margin effect dominates, leading to an increase in the total bank credit supply.

5.3.1 Numerical illustration: sluggish recoveries

Figure 7 illustrates the evolution of total credit supply $\bar{K}_t$, following the destruction of 20% of lending relationships. In this example, I compare 3 economies subject to either i) search, ii) limited commitment, or iii) both. The half-life of a crisis is 2, 9 and 18 quarters, respectively. The combination of search and limited commitment creates a relatively more sluggish credit recovery because only a fraction of lending relationships are re-established every period, and the corresponding initial firm value is relatively smaller (compared to a setting with perfect competition). It is also worth observing that the introduction of risk-averse entrepreneurs, *mutatis mutandis* leads to an even slower recovery as the growth rate of V decreases with the degree of risk aversion.[39]
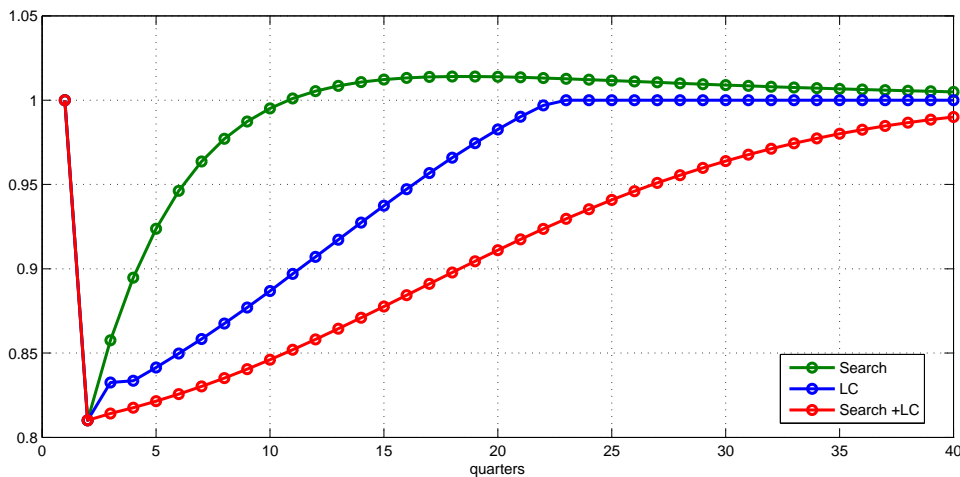

Figure 7. Response (in % change from steady state level) to a one-time match destruction shock (-20%).

---

[39]This effect can however be reversed in general equilibrium as both the probability of approval and firm value at origination would also change.

5.4 Empirical predictions

Before turning to the quantitative analysis of the model, I first present a set of empirical regularities previously documented in U.S. studies on bank lending relationships and credit markets, consistent with the theoretical setting.[40] The model also provides additional testable predictions listed below.

Lending relationships.

- Credit availability and pricing improve with the length of the lending relationship.
- Credit availability increases relatively faster for younger relationships.
- Credit availability and growth rates are higher for firms matched with banks with low funding costs.

Bank lending markets.

- Cross-section. Banks with low funding costs have lower approval rates and higher lending standards. These banks also offer better contractual terms: they provide greater access to credit and charge lower borrowing costs throughout the lending relationship.
- Business cycle. Access to credit becomes more difficult in downturns. The degree of bank competition and lending standards are countercyclical.

Bank competition effects.

- Bank competition generates higher approval rates (or equivalently lax lending standards), increases the creation rate of lending relationships (less credit rationing at the extensive margin), but decreases the amount of credit available at origination (more credit rationing at the intensive margin).
- Controlling for bank type, bank competition increases the dispersion of lending rates across borrowers. Controlling for the length of the lending relationship, bank competition decreases the dispersion of lending rates across banks.

Legal environment effects.

- Economies with more entrepreneur-friendly bankruptcy laws have higher levels of credit rationing (at both intensive and extensive margins).
- Credit availability and pricing improve at slower rates in these economies.

---

[40]See the excellent survey in Degryse et al. (2009).

6 Quantitative analysis

I now move to the quantitative properties of the model and its application to commercial lending. I first specify the functional forms of the model and calibrate its parameters. I then evaluate its steady-state and business-cycle properties. I show that productivity shocks alone are not able to generate long-lived periods of low credit. The economy is instead more sensitive to shocks to bank funding costs, and can generate credit flow patterns consistent with the recent financial crisis.

6.1 Model specification

I set the period utility of entrepreneurs to $u(c) = log(c)$ and specify the firm-level profit function of the form $f(z, K) = zK^\alpha$, where α is the decreasing-return-to-scale coefficient. As it is standard in discrete-time search models, I assume a CES matching function with elasticity γ_m, generating the following meeting probabilities:

$$
\begin{aligned}
p(\theta) &= \theta(1 + \theta^{\gamma_m})^{-\frac{1}{\gamma_m}}, \\
q(\theta) &= (1 + \theta^{\gamma_m})^{-\frac{1}{\gamma_m}}.
\end{aligned}
$$

I also assume that the aggregate productivity follows an AR(1) process:

$$
\log(z_{t+1}) = (1 - \rho_z)\log(\bar{z}) + \rho_z \log(z_t) + \sigma_z \epsilon_{t+1}, \qquad \epsilon \sim \mathcal{N}(0, 1)
$$

Similarly, I assume that the aggregate and idiosyncratic components of the bank funding cost follow AR(1) processes:

$$
\begin{aligned}
\bar{r}_{t+1} &= (1 - \rho_r)\bar{r}_0 + \rho_r \bar{r}_t + \sigma_r \nu_{t+1}^r, \qquad \nu^r \sim \mathcal{N}(0, 1), \\
s_{t+1} &= \rho_s s_t + \sigma_s \nu_{t+1}^s, \qquad\qquad \nu^s \sim \mathcal{N}(0, 1).
\end{aligned}
$$

The process innovations ϵ_t, ν_t^r and ν_t^s are assumed to be uncorrelated, and $\{\nu_t^{s,i}\}_i$ are also independent across banks. All processes are approximated using a finite grid with 5 shock realizations for aggregate productivity ($N_z = 5$) and funding components ($N_{\bar{r}} = 5$) and 9 shock realizations for the bank-specific funding component ($N_s = 9$) following the discretization methodology of Tauchen and Hussey (1991).

6.2 Calibration

The calibration is based on the simulated method of moments (SMM). I set the length of a period in the model to one quarter. I need to assign values to the following set of model parameters

$\{\beta, \alpha, \delta, \sigma, d_0, c, \gamma_m, \xi, \eta\}$, in addition to the parameters governing productivity and bank funding shocks $\{\bar{z}, \rho_z, \sigma_z, \bar{r}_0, \rho_r, \sigma_r, \rho_s, \sigma_s\}$.

In the baseline estimation, I pre-calibrate 5 model parameters. The discount rate β is set to 0.9875 to imply an annual real interest rate of about 5%. The decreasing-returns-to-scale coefficient α is set to 0.75, which is inline with estimates in Hennessy and Whited (2007). This parameter governs the firms optimal scale and the dispersion of firm size distribution.[41] The quarterly depreciation rate δ is set to 0.03, which is in the standard range of values used in the literature.[42] The probability of exogenous separation is set to 3.15% per quarter in order to match its empirical counterpart.[43] Garage production d_0 is set to 1 so that $u(d_0) = 0$.

I also determine the parameters behind the two processes governing bank funding from their empirical counterparts. To this end, I use information contained in the quarterly Consolidated Reports of Condition and Income (Call Reports) and define the realized real funding costs incurred by banks, as the ratio of total interest expenses over total assets over the period 1988 to 2008.[44] The cross-sectional average of the quarterly funding costs of the banking sector is given by $\bar{r} = 0.85\%$. I estimate the standard deviation and autocorrelation of the aggregate funding costs process based on the real effective Fed Funds rates over the same period: $\sigma_r = 0.55\%$, and $\rho_r = 0.90$. Eventually, I extract the bank-specific component of the funding cost by subtracting the cross-sectional average evaluated within each quarter, and computing the autocorrelation and standard deviation for each bank time series available throughout the sample period. The tabulated cross-sectional averages of these moments are $\rho_s = 0.84$ and $\sigma_s = 0.15\%$ (quarterly).

The remaining 7 parameters, namely the productivity process parameters $(\bar{z}, \rho_z, \sigma_z)$, bank origination costs c, share of divertible capital η, probability of fresh start ξ, and the matching elasticity γ_m, are calibrated using a simplex algorithm minimizing the squared distances (in relative terms) between empirical and simulated moments.

I target (i) a yearly net return-on-assets of 1% for bank loans, consistent with Boualam (2014). This moment is matched to the cross-sectional average net return per unit of capital lent, which takes into account the interest rate charged to borrowers minus bank funding and non-interest costs.

The cross-sectional (ii) average of investment rate of 0.145 and (iii) standard deviation 0.139, are taken from Gomes (2001), and calibrated to match the corresponding moments associated with credit growth rate given by $\frac{K'-K}{K}$. These targets are particularly helpful in determining η and ξ. Both of these parameters appear in the borrower's outside option, and govern the firm's growth rate. In particular,

---

[41]α also governs the sensitivity of firm capital and output to productivity and interest rates. A direct calibration of this parameter will be conducted to better match these sensitivities.

[42]Note that this parameter reflects both capital depreciation and non-interest expenses incurred during financial intermediation (which account for about 85bp quarterly according to Call Reports).

[43]See empirical appendix for details. Note also that this is more likely to be a lower bound when considering small and medium-sized businesses given that a sizable fraction of the DealScan database is related to large and public firms.

[44]See Boualam (2014) for details about the construction of the time series.

η governs the speed at which firms reach the unconstrained region (as η tends to 0, the outside option becomes independent of K and firms reach the first-best level almost immediately). On the other hand, because ξ is linked to the firm value when unfunded, it relates the volatility of investment rate to that of the aggregate shocks in the economy.

I also target a (iv) leverage of 0.28, consistent with Bhamra et al. (2010), and map it to its model counterpart at the firm level defined as $\frac{B}{V+B}$, and where the net present value of the contract to the bank B is interpreted as total debt of the firm. This measure captures the degree of distortion in the contract (and reflects the allocation of surplus across borrowers and lenders) and will be helpful in identifying bank origination costs c.[45]

I also identify the matching elasticity parameter by using information on the relationship between approval rates and interest rates.[46] In particular, I target a (v) slope $\beta(p(\theta), r)$ of 0.15, which corresponds to a 15% increase in the approval rate for each additional 1% increase in the annualized lending rate charged at origination.

I finally use the autocorrelation and standard deviation of log-detrended output in order to determine the parameters associated with the aggregate productivity process. In particular, I use real quarterly log-GDP data (seasonally adjusted, and detrended using HP filter with parameter 1600) obtained from the Bureau of Economic Analysis for the period 1947Q1 to 2013Q4, and tabulate the following targets: (vi) autocorrelation of 0.84 and (vii) standard deviation of 0.017.

The bottom panel of Table 1 reports the *preliminary* results of the calibration, while the calibration targets and model fit are reported in Table 2.

6.3 Model properties and validation

The model generates a stationary distribution of bank-firm relationships $g(r_d, V)$ which depends on bank-specific funding cost and firm value. The steady-state level of unfunded firms is 10.5%, while the fraction of unconstrained firms is 62.5%. The remaining 27% are firms that are currently matched but still within the constraint region. The model generates an average approval rate of about 30% per quarter, which corresponds to an expected success rate of 75% annually, and an expected search period of about 3 quarters.[47] Eventually, firms are able to outgrow their borrowing constraint in about 3.5 years.[48]

---

[45] This definition of leverage does not take into account capital accumulation and is therefore biased upwards.

[46] See the following section and the empirical appendix for details.

[47] While I am not aware of an equivalent measure for the U.S., Eurostat provides data for the U.K. showing an annual success rates for firms seeking credit in 2007 and 2010 at 88% and 65% respectively. See epp.eurostat.ec.europa.eu/portal/page/portal/european_business/special_sbs_topics/access_to_finance.

[48] Note that the speed at which firms reach their optimal scale and the sensitivity of credit availability and pricing with respect to the length of a lending relationship appear to be relatively higher than what empirical evidence suggests (for example Hubbard et al. (2002)). A more extensive calibration involving the decreasing-return-to-scale parameter α and the curvature of the utility function will be helpful in adjusting this fact.

Parameter	Value	Description
pre-calibrated		
β	0.9875	Discount factor
α	0.75	Decreasing-returns-to-scale coefficient
δ	0.03	Depreciation rate
σ_0	0.0315	Exogenous separation probability
d_0	1	Garage production
$\bar{r}_0$	0.0085	Average funding cost
ρ_r	0.90	Persistence of common funding shock
σ_r	0.0055	Standard deviation of common funding shock
ρ_s	0.84	Persistence of idiosyncratic funding shock
σ_s	0.0015	Standard deviation of idiosyncratic funding shock
calibrated		
ξ	0.2	Probability of fresh start
η	0.58	Share of divertible assets
γ_m	2.05	Matching elasticity coefficient
c	9.25	Bank origination cost
$\bar{z}$	0.1965	Average aggregate productivity
σ_z	0.011	Standard deviation of aggregate productivity shock
ρ_z	0.92	Persistence of aggregate productivity shock

Table 1. Parameter values (quarterly) - Preliminary calibration.

	Data	Model
Return on Assets	0.01	0.01
Leverage	0.28	0.23
$\Delta K/K$	0.145	0.136
$\sigma(\Delta K/K)$	0.139	0.045
$\beta(p(\theta), r)$	0.15	0.15
$\rho(log(output))$	0.84	0.83
$\sigma(log(output))$	0.017	0.017

Table 2. Targeted moments. This table reports empirical and simulated moments. Data on ROA are tabulated from Call Reports and taken from Boualam (2014). Data on investment rates come from Gomes (2001). Leverage data are consistent with Bhamra et al. (2010). Approval rates are quarterly. log(output) is detrended using HP filter parameter 1600, and autocorrelations and standard deviations are computed based on log-deviation from trend (quarterly).

6.3.1 Approval rates and contractual terms

In order to test the validity of the model, I first explore the relationship between approval and lending rates. Figure 8 displays both empirical and model-generated moments. Detailed data on approval rates and lending standards are neither reported by banks nor publicly available. I attempt to overcome this limitation by using information on approval rates for both small and large banks published by a loan broker (Biz2Credit) on a monthly basis over the period February 2011 - May 2014.[49] I then construct the corresponding lending rates using the effective weighted-average interest rates with minimal risk, for the period Q2 1997 - Q1 2014.[50]
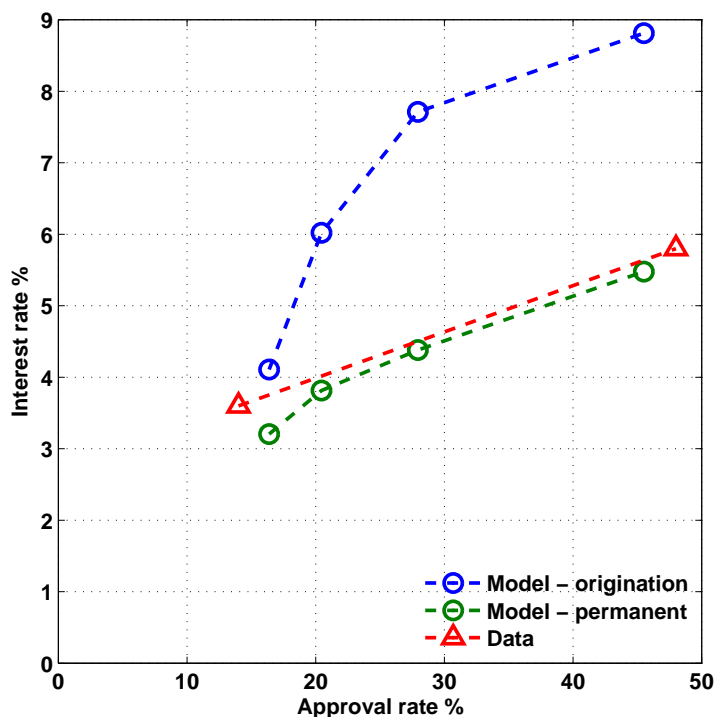

Figure 8. Lending rates vs. approval rates (Model (lending rates at origination (blue), permanent (green)), Data (red)).

For each level of bank funding cost, I construct two types of model-generated moments for lending rates. First, a lending rate at origination, which corresponds to the rate charged during the first period of the contract. Second, I define a permanent lending rate derived from the contract values offered by each bank in the steady state. Given that contracts offered in reality have a certain maturity and are not

---

[49]www.biz2credit.com/small-business-lending-index/may-2014.html.

[50]These time series are obtained from FRED (EEANXSLNQ, EEANXSSNQ). I report here the corresponding interest rates across a larger sample period to be able to compare the model-generated output (based on a steady state level of interest expenses at 3.3%) to the data (the average interest rates associated with the period 2010-2013 being historically low). Details about data construction and corresponding time series are in the appendix.

necessarily adjusted every quarter, this measure seems to be better suited for comparison with the data. For each bank funding cost r_d, it is formally defined as:

$$\hat{r}(r_d) = \frac{z\hat{K}^\alpha - \hat{d}}{\hat{K}},$$

with $\hat{d} = u^{-1}\big((1-\beta)V\big)$, and where $\hat{K}$ solves $(1-\beta)\Pi = \hat{\Pi} = z\hat{K}^\alpha - (\delta + r_d)\hat{K} - \hat{d}$.

The relationship between approval and lending rates is overall satisfied by the model for both the level and the slope. This is particularly the case when considering the measure based on permanent lending rates.

6.3.2 Credit relationship flows

I tabulate measures of creation and destruction rates of lending relationships from LPC's DealScan database over the period 1997-2013.[51] DealScan's coverage is biased towards relatively large firms and may not be representative of the whole economy. Yet, this is to my knowledge, the only comprehensive and publicly available source of information that can be used to construct proxies for these flows. While it is unclear whether positive flows are higher or lower in the whole economy, it is reasonable to assume that the destruction rates are a priori understated since smaller firms are typically more likely to default or have their loans being terminated due to a potentially higher propensity to violate loan covenants.
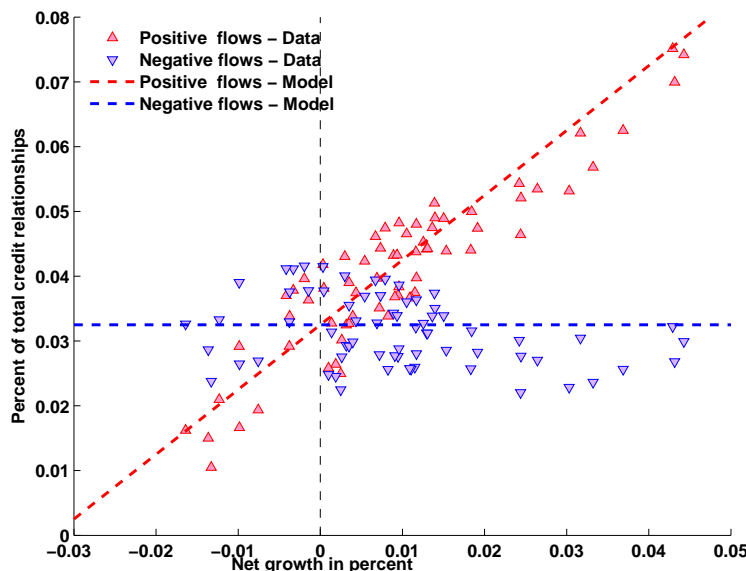

Figure 9. Positive and negative flows of lending relationships as a function of net flows.

---

[51]Although the database starts well before, I only begin my sample in 1997 to make sure that the database is already well populated in order to limit the mechanical bias in positive flows induced by the increase in data coverage. See empirical appendix for details about data construction and robustness checks.

Figure 9 displays the quarterly gross rates of creation and destruction of credit relationships against the net growth rate (i.e. creation minus destruction rate). The model generates by construction a one-to-one map between positive and net flows by construction since the model assumes a constant destruction rate σ. This assumption seems reasonable as a first pass.[52] More importantly, Table 3 in the appendix shows that positive flows are more volatile and more sensitive to the business cycle relative to negative flows. This suggests that the adjustment of the stock of lending relationships weighs more heavily on the process of origination and entry rather than destruction. In particular, the relative rigidity in destruction rates may reflect the fact that loan agreements are typically fixed-term and banks may not always have the flexibility to cut lending until maturity.

6.4   Results: Aggregate shocks

In this section, I study the effects of aggregate shocks related to bank funding cost and firm productivity. I examine the response of the economy along both extensive and intensive margins, and analyze how these shocks affect credit availability and lending rates for both incumbent and new borrowers.

6.4.1   Bank funding shocks

I first look at the implications of a 1 standard deviation increase in the common component of bank funding cost. This shock shifts the distribution of bank funding costs in the economy. As a consequence, funded firms (both constrained and unconstrained) scale down their production as the optimal level of capital decreases. This is the intensive margin of credit at play in the model. Here because there are no adjustment costs or capital accumulation, both firm size and credit availability adjust immediately.

However, firms that are initially unfunded are severely impacted. Establishing new lending relationships becomes less profitable for banks since the joint surplus decreases. Only banks with very low idiosyncratic funding costs will be able to offer credit. Credit market tightness therefore goes down and lending standards soar. Access to credit for new entrepreneurs becomes very limited as the approval rates decreases by about 25% and the number of lending relationships in the economy drops given that the origination rate falls below the destruction rate. This is the extensive margin effect.

In contrast to other macro models which rely on exogenous fluctuations in firm collateral value or the corresponding degree of pledgeability, the borrowing capacity is endogenous and responds to firm value and future cash flow claims, which in turn depend on the aggregate state of the economy. This approach which endogenizes the probability of access to credit can also be viewed as an alternative to models with financial or credit shocks as in Jermann and Quadrini (2012) or Khan and Thomas (2013).

The subsequent response of credit rationing is hump-shaped, consistent with the fact that the rate of

---

[52]Introducing endogenous separation in the contracting problem (see section 7) would however better capture the increase in destruction rates during downturns, and is left for future research.

origination of new relationships stays below its steady state level for about 5 periods. The recovery exhibits a relatively strong persistence in the level of credit rationing and total credit supplied in the economy as shown in the top right panel of Figure 11.

Let us focus now on the response of the contractual terms, namely credit availability and pricing. What is particularly interesting is the stark asymmetry between new and incumbent borrowers. On one hand, incumbent borrowers face a relatively small decline in credit availability (their working capital decreases by about 8-10%), while credit offered to new borrowers drops by over 20%. Similarly, banks shield their incumbent borrowers by marginally adjusting their lending rates during the crisis (corresponding to an increase of about 10bp), while lending rates charged to new borrowers soar by over 120bp.
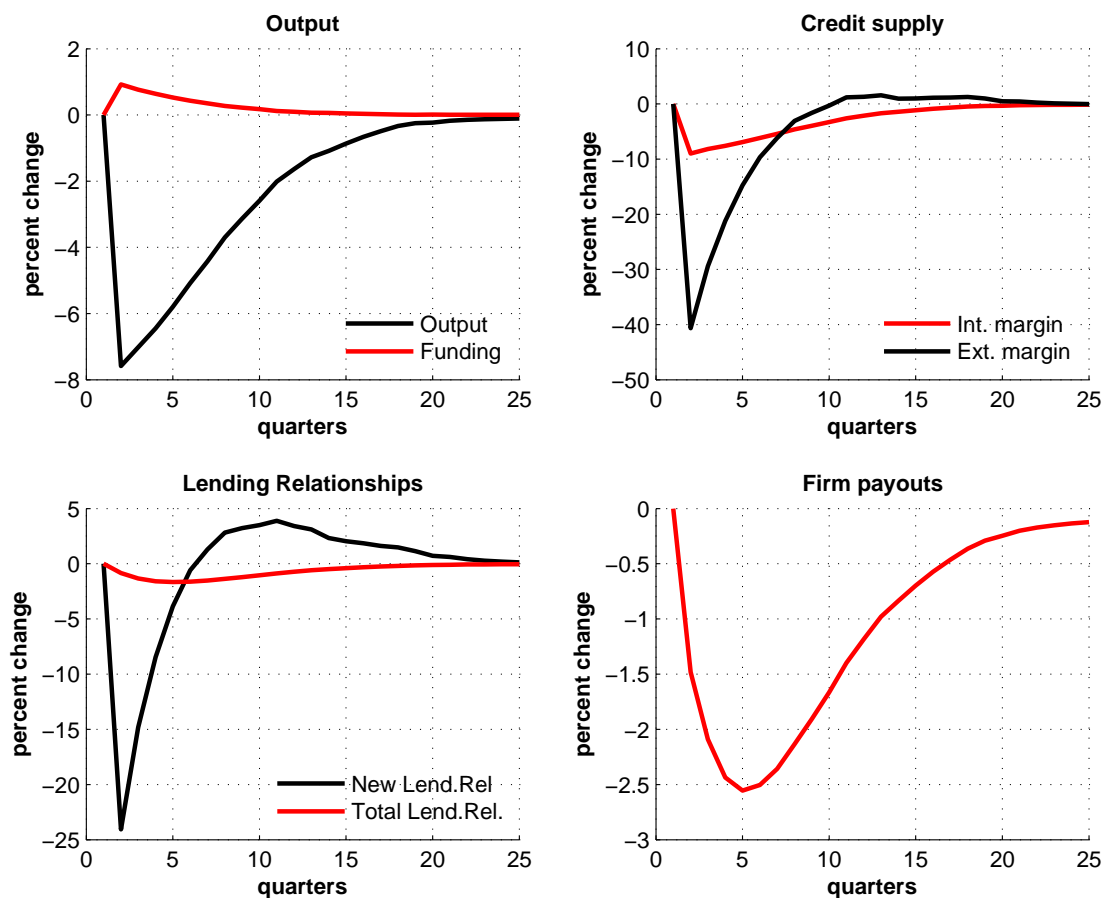

Figure 10. Impulse response - increase in bank funding cost. Aggregate variables.

The asymmetry in the treatment between these two types of borrowers is partly due to the bank's full commitment and the insurance mechanism induced by long-term contracts. Banks end up subsidizing their long-term borrowers during recessions, and whenever the amount of surplus generated by the match is low, dividend payouts are smoothed out.

This is however not the case for new borrowers. As the shock hits the economy, market tightness plummets and banks offering credit extract a larger share of the match surplus, further impacting these borrowers (this effect can be interpreted as an increase in the surviving banks' bargaining power in downturns). Therefore, the contractual terms offered to unfunded firms are adjusted unfavourably during recessions, and reflect a sharp decline in credit availability in addition to higher lending rates.
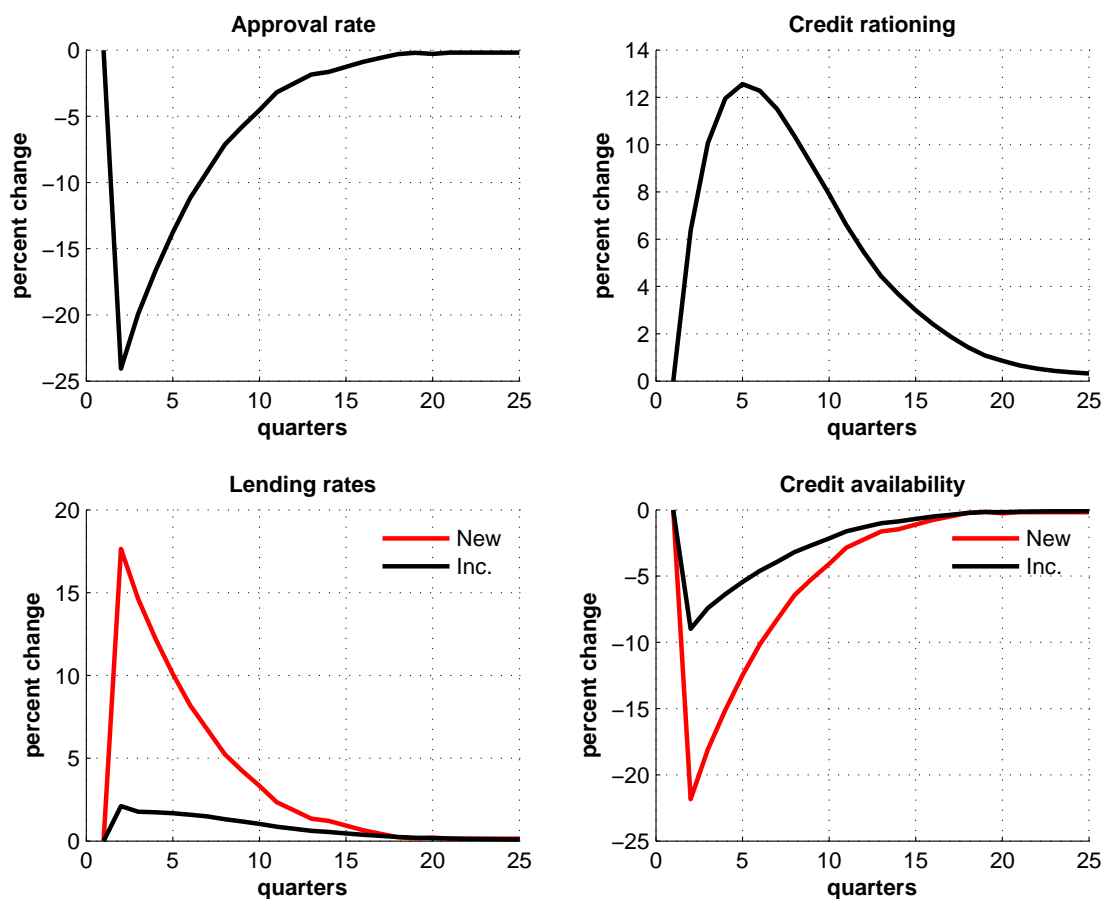

Figure 11. Impulse response - increase in bank funding cost. Credit markets and contractual terms.

6.4.2 Productivity shocks

Let us now consider the response of the economy following a 1 standard deviation decrease in aggregate productivity (The aggregate productivity returns to its steady-state level with persistence ρ_z). Figures 17 - 18 in the appendix display the responses. This shock generates the same qualitative responses as a positive bank funding shock. Overall, responses are however relatively less sensitive.[53] Two main reasons explain this difference. First, productivity and bank funding costs enter in the surplus function

---

[53]The response of aggregate output is again over-stated given that the model does not account for the accumulation of capital stock.

differently. Second and more importantly, aggregate productivity is less volatile relative to bank funding costs.

6.4.3 Unanticipated relationship destruction shock

I finally consider the consequences of an unanticipated destruction shock (i.e. a one-time jump in σ from its initial level of 3.15% to 10% in this example). Figure 12 displays the economy's response following such a "catastrophic" event (due to the default of a couple of major commercial lenders for example).

The goal of this exercise is to test and decompose the persistence due to this lending relationship channel, all else equal. Figure 12 displays a comparison between the persistence generated by the model - which captures both the time-to-form (extensive margin) and the time-to-build (intensive margin) aspects - and models where either one or both frictions are muted.
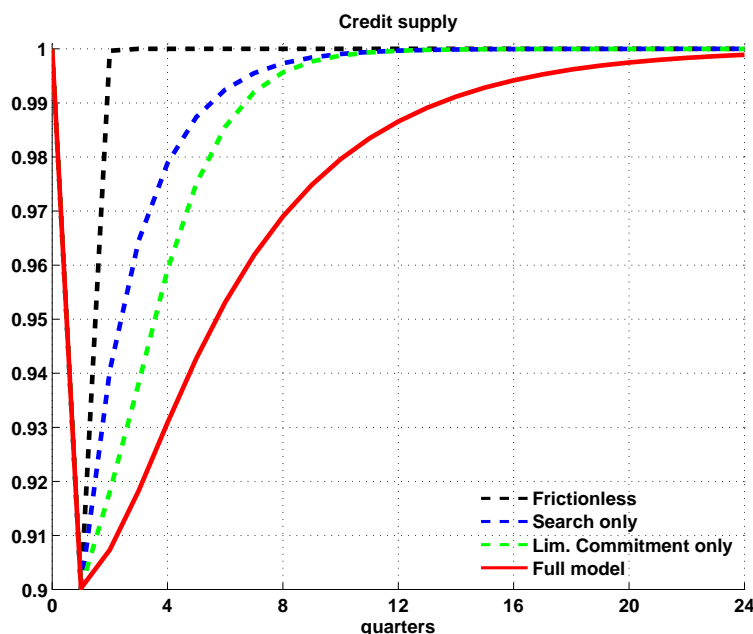

Figure 12. Credit recovery profiles following a one-time unanticipated destruction shock.

The model naturally captures the joint persistence coming from both search and agency frictions. Looking at the 90%-recovery (which corresponds to credit supply going back to 0.99 in Figure 12), it takes about 13 quarters for the full model to reach this level, compared to about 5 and 6 quarters for models featuring only search or limited commitment, respectively. What is also particularly interesting is the fact that the full model exhibits a slow and persistent recovery while the other cases generate an immediate rebound in aggregate credit supply.

6.5 Policy analysis

6.5.1 A simple policy targeting the extensive margin of credit

I analyze the effect of a policy specifically targeted toward credit origination. The goal of the policy I consider is similar to that of the Small Business Lending Fund (SBLF) proposed in the U.S. as part of the 2010 Small Business Jobs Act. More importantly, it is related to the Funding for Lending Scheme instituted in 2012 in the U.K., and to the more recent T-LTRO (Targeted Long-Term Refinancing Operation) program designed by the European Central Bank.[54] While the SBLF only targets relatively small banks, the other programs are generalized across all banks and subsidize their funding whenever certain lending criteria are met.

I run a simple policy experiment where the government subsidizes part of the origination costs of the bank.[55] This closely resembles the policies described above, in the sense that it specifically targets the origination of new credit. The main difference however is the fact that this is a one-time lump-sum subsidy transferred at origination.
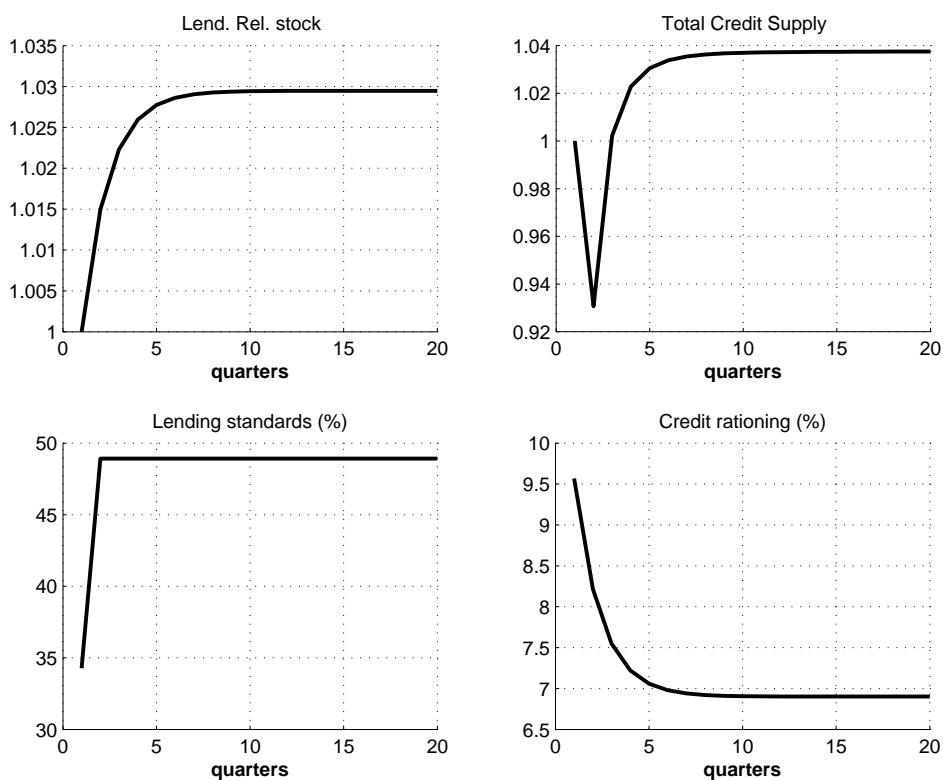

Figure 13. Response to credit origination subsidies - Aggregate variables.

---

[54]See www.bankofengland.co.uk/markets/Pages/FLS/default.aspx for more details.

[55]Note that these costs not only reflect screening and non-interest expenses incurred during the origination, but can be more generally interpreted as an initial sunk investment or long-term debt.

Although the aim here is not to analyze welfare implications, the experiment can already help us gauge the short and long-run effects of this policy on credit allocation and its impact across new and incumbent borrowers. Figures 13 - 14 display the economy's transitional path following the unanticipated introduction of the policy at date 1. While it naturally delivers an increase in the number of lending relationships and credit supplied in the economy in the long run, the policy may actually appear counterproductive in the short-run. By directly affecting credit market conditions, it creates some tension between incumbent and new borrowers. On one hand, the policy is beneficial to unfunded firms and entrepreneurs as it improves their access to credit and credit availability. On the other hand, it may negatively impact currently funded-but-constrained borrowers. As origination costs decrease, bank entry and competition increase as well. This positively impacts credit market tightness and eases lending standards (i.e. approval rates are adjusted upwards).

However, improving market conditions also positively impact firm value after default and banks will temporarily adjust their credit supply downwards to prevent their incumbent borrowers from running away. As these firms gradually grow out of their borrowing constraints, this tension dissipates allowing the aggregate credit supply to eventually increase.
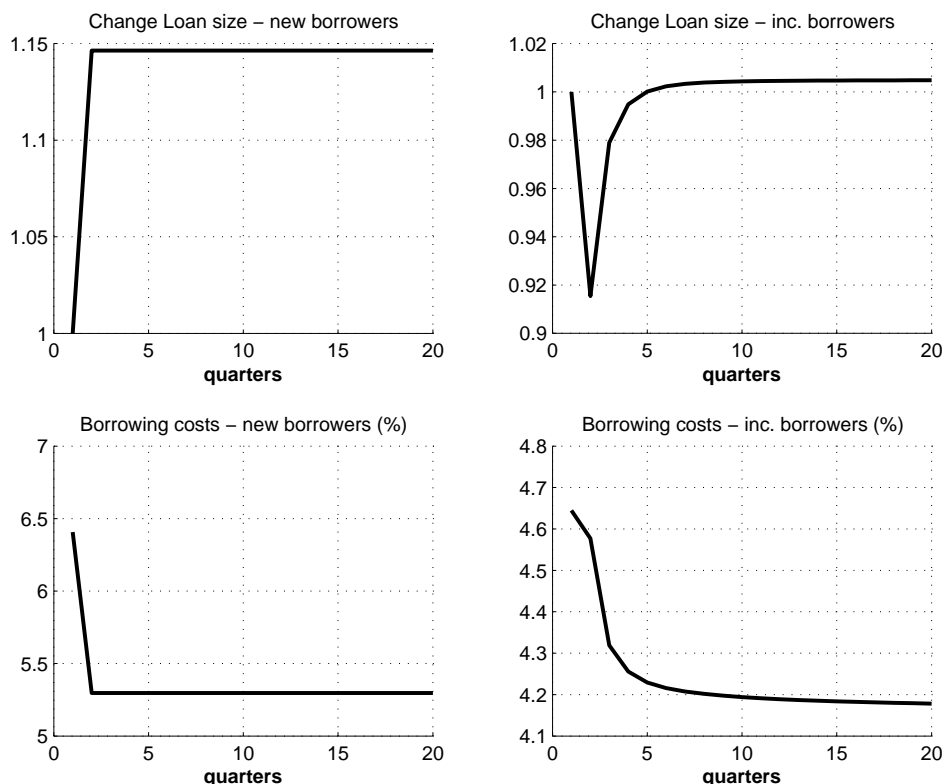

Figure 14. Response to credit origination subsidies - Cross-section.

7 Extensions and Comments

The framework analyzed in this paper can be extended along many directions, which are left for future research. In this section, I describe three possible extensions and explain their general implications.

7.1 Endogenous separation

I consider the possibility of endogenous separation between borrowers and lenders. In this case, the bank also takes into account the firm's discrete decision rule with regards to separation. Following the realization of aggregate and bank-specific shocks, the firm's choice is therefore driven by $\max\{V_{z,r_d}, W(z)\}$ which balances the current value of the contract V_{z,r_d} if the relationship continues and the firms value following separation $W(z)$. Because the contract specifies state-contingent continuation values $\{V_{z',r'_d}\}_{z',r'_d}$ before shocks are realized, this simply translates into an ex-ante probability of termination given by:

$$\sigma_{z,r_d} = \sigma(z', r'_d, V_{z',r'_d}) = \begin{cases} \sigma_0 & \text{if } V_{z',r'_d} \geq W(z'), \\ 1 & \text{otherwise.} \end{cases}$$

with σ_0 being an exogenous destruction rate. In equilibrium, the bank can never promise a value that is below the unfunded firm value W. In such case, the firm will always walk away in order to search for a better lending opportunity. However, in certain states where there are no gains from trade, both agents may be better off with separation, in which case, the bank receives 0 (the bank does not derive any stream of income when it is not matched with a borrower), and the firm becomes unfunded with value W.

Because the bank is always fully committed to deliver promised value V, this means that the promised values offered in the continuation states are higher relative to the case where separation is not allowed. In this case, the saddle-point problem (SPFE) is slightly modified to take into account the additional constraint:

$$\begin{aligned} P(z, r_d, \Lambda) = \inf_{\lambda} \sup_{K,d} \quad & f(z,K) - d - (\delta + r_d)K - \lambda\left[u(\eta K) + \beta H(z)\right] \\ & + (\Lambda + \lambda)\left[u(d) + \beta\sigma_0 E_z[W(z')]\right] \\ & + \beta(1 - \sigma_0)E_{z,r_d}\left[\max((\Lambda + \lambda)W(z'), P(z', r'_d, \Lambda + \lambda))\right] \\ \text{s.t.} \quad & (LL), \qquad \lambda \geq 0. \end{aligned}$$

Introducing endogenous separation allows credit market conditions to affect the dynamics of lending relationships through an additional channel. In such case, shocks to the aggregate economy are further amplified since they impact both entry and exit margins, and generate larger declines in the stock of lending relationships.

42

7.2 Capital accumulation

The model is tractable enough to allow for the introduction of capital accumulation or entrepreneurial savings. In this case, the capital stock becomes a state variable and the distribution of firms in the economy would now depend on the level of capital stock, the financing status (funded or unfunded), in addition to the type of lender and the length of the relationship (when it is funded). From a qualitative standpoint, the dynamics of the contract will not be different from the model presented here. However, the introduction of capital accumulation means that entrepreneurs are in general better off when they become unfunded, relative to the baseline model. The level of accumulated capital would also affect the search behavior of unfunded agents and create an additional layer of heterogeneity in the credit markets.

7.3 Credit markets for Rookie vs. Seasoned firms

The model can also be augmented by allowing for market segmentation among unfunded firms, distinguishing between newly created firms ("rookie firms"), and more "seasoned firms" (i.e. initially established and funded but currently searching for new financiers). The market for newly created firms is subject to higher origination costs due to higher screening costs, while more established firms have for example a publicly observable track record and access to credit in markets requiring lower origination costs. This feature is easily implementable and would allow for the distinction between the contractual terms and credit dynamics associated with both entrant and incumbent firms.

8 Conclusion

This paper develops and characterizes a novel dynamic equilibrium theory of bank relationship capital in an economy subject to search frictions and limited enforceability. The model features a dynamic contracting problem within a directed search equilibrium, with aggregate and bank-specific uncertainty, and where heterogeneous financiers compete for borrowers by posting long-term credit offers. The interaction between these two frictions generates a slow accumulation of lending relationship capital and distorts the optimal allocation of credit along both intensive and extensive margins.

This theoretical research sheds light on the process of formation of credit relationships which has been relatively neglected in the literature. Significantly, it highlights how adverse aggregate shocks can be propagated when they negatively impact the stock of lending relationships and producing firms in the economy. Crises characterized by a sizable destruction of lending relationships can therefore generate slow subsequent recoveries. By providing a framework that captures multiple dimensions of the credit reallocation process, the model is particularly relevant when analyzing the effects of policies targeted toward business lending. The paper shows that policies directly subsidizing the cost of origination of

new credit relationships are effective at boosting the aggregate credit supply in the long-run but can also lead to adverse effects in the short-run.

Further empirical investigations focusing on the dynamics of the extensive margin of credit and the process of origination and matching between banks and firms are a fruitful area for future work. Finally, a thorough analysis of the constrained-inefficiency result obtained in the presence of risk-averse entrepreneurs and search frictions yields important policy implications for aggregate lending and risk-taking behaviors, and is left for future research.

References

Abreu, D., D. Pearce, and E. Stacchetti. 1990. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica: Journal of the Econometric Society* pp. 1041–1063.

Acemoglu, D., and R. Shimer. 1999. Efficient unemployment insurance. *Journal of Political Economy* .

Adrian, T., and N. Boyarchenko. 2012. Intermediary leverage cycles and financial stability. Tech. rep., Staff Report, Federal Reserve Bank of New York.

Agarwal, S., and R. Hauswald. 2010. Distance and private information in lending. *Review of Financial Studies* .

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

Allen, F., and D. Gale. 1997. Financial markets, intermediaries, and intertemporal smoothing. *Journal of Political Economy* 105:523–546.

Bebchuk, L. A. 2000. Using options to divide value in corporate bankruptcy. *European Economic Review* 44:829–843.

Bebchuk, L. A., and I. Goldstein. 2011. Self-fulfilling credit market freezes. *Review of Financial Studies* 24:3519–3555.

Becsi, Z., V. E. Li, and P. Wang. 2005. Heterogeneous borrowers, liquidity, and the search for credit. *Journal of Economic Dynamics and Control* 29:1331–1360.

Benmelech, E., and N. K. Bergman. 2012. Credit traps. *American Economic Review* 102:3004–3032.

Berlin, M., and L. J. Mester. 1999. Deposits and relationship lending. *Review of Financial Studies* 12:579–607.

Bernanke, B., and M. Gertler. 1989. Agency costs, net worth, and business fluctuations. *The American Economic Review* pp. 14–31.

Bhamra, H. S., L.-A. Kuehn, and I. A. Strebulaev. 2010. The aggregate dynamics of capital structure and macroeconomic risk. *Review of Financial Studies* 23:4187–4241.

Biais, B., T. Mariotti, J.-C. Rochet, and S. Villeneuve. 2010. Large risks, limited liability, and dynamic moral hazard. *Econometrica* 78:73–118.

Bigio, S. 2014. Financial risk capacity. *manuscript* .

Bolton, P., X. Freixas, L. Gambacorta, and P. E. Mistrulli. 2014. BIS Working Papers .

Boot, A. W. 2000. Relationship banking: What do we know? *Journal of financial intermediation* 9:7–25.

Boot, A. W., and A. V. Thakor. 2000. Can relationship banking survive competition? *The Journal of Finance* 55:679–713.

Boualam, Y. 2014. Bank Dynamics and Credit Market Competition .

Brunnermeier, M. K., T. M. Eisenbach, and Y. Sannikov. 2012. Macroeconomics with financial frictions: A survey. Tech. rep., National Bureau of Economic Research.

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

Chodorow-Reich, G. 2014. The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 Financial Crisis. *The Quarterly Journal of Economics* 129:1–59.

Clementi, G. L., and H. A. Hopenhayn. 2006. A theory of financing constraints and firm dynamics. *The Quarterly Journal of Economics* pp. 229–265.

Cooley, T., R. Marimon, and V. Quadrini. 2004. Aggregate consequences of limited contract enforceability. *Journal of Political Economy* 112.

De Marzo, P. M., and M. J. Fishman. 2007. Optimal long-term financial contracting. *Review of Financial Studies* 20:2079–2128.

De Marzo, P. M., and Y. Sannikov. 2006. Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model. *The Journal of Finance* 61:2681–2724.

Degryse, H., M. Kim, and S. Ongena. 2009. *Microeconometrics of banking: methods, applications, and results.* Oxford University Press.

Den Haan, W. J., G. Ramey, and J. Watson. 2003. Liquidity flows and fragility of business enterprises. *Journal of Monetary Economics* 50:1215–1241.

Diamond, D. W., and R. G. Rajan. 2011. Fear of fire sales, illiquidity seeking, and credit freezes. *The Quarterly Journal of Economics* 126:557–591.

Diamond, P. 1990. Pairwise credit in search equilibrium. *The Quarterly Journal of Economics* 105:285–319.

Froot, K. A., D. S. Scharfstein, and J. C. Stein. 1993. Risk management: Coordinating corporate investment and financing policies. *the Journal of Finance* 48:1629–1658.

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

Gertler, M., and P. Karadi. 2011. A model of unconventional monetary policy. *Journal of monetary Economics* 58:17–34.

Gertler, M., and N. Kiyotaki. 2010. Financial intermediation and credit policy in business cycle analysis. *Handbook of monetary economics* 3:547–599.

Gilchrist, S., J. W. Sim, and E. Zakrajšek. 2013. Misallocation and financial market frictions: Some direct evidence from the dispersion in borrowing costs. *Review of Economic Dynamics* 16:159–176.

Golosov, M., P. Maziero, and G. Menzio. 2013. Taxation and redistribution of residual income inequality. *Journal of Political Economy* 121:1160–1204.

Gomes, J. F. 2001. Financing investment. *American Economic Review* pp. 1263–1285.

Greenstone, M., and A. Mas. 2012. Do credit market shocks affect the real economy? Quasi-experimental evidence from the Great Recession and 'normal' economic times .

Guerrieri, V., R. Shimer, and R. Wright. 2010. Adverse selection in competitive search equilibrium. *Econometrica* 78:1823–1862.

Harris, M., and B. Holmstrom. 1982. A theory of wage dynamics. *The Review of Economic Studies* 49:315–333.

He, Z., and A. Krishnamurthy. 2011. A model of capital and crises. *The Review of Economic Studies* .

Hennessy, C. A., and T. M. Whited. 2007. How costly is external financing? Evidence from a structural estimation. *The Journal of Finance* 62:1705–1745.

Hubbard, R. G., K. N. Kuttner, and D. N. Palia. 2002. Are there bank effects in borrowers' costs of funds? Evidence from a matched sample of borrowers and banks. *The Journal of Business* 75:559–581.

Inderst, R., and H. M. Müller. 2004. The effect of capital market characteristics on the value of start-up firms. *Journal of Financial Economics* 72:319–356.

Ivashina, V., and D. Scharfstein. 2010. Bank lending during the financial crisis of 2008. *Journal of Financial Economics* 97:319–338.

Jermann, U., and V. Quadrini. 2012. Macroeconomic effects of financial shocks. *The American Economic Review* 102:238–271.

Jermann, U. J., and V. Quadrini. 2007. Stock market boom and the productivity gains of the 1990s. *Journal of Monetary Economics* 54:413–432.

Khan, A., and J. K. Thomas. 2013. Credit shocks and aggregate fluctuations in an economy with production heterogeneity. *Journal of Political Economy* 121:1055–1107.

Kiyotaki, N., and J. Moore. 1997. Credit cycles. *The Journal of Political Economy* 105:211–248.

Klemperer, P. 1995. Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade. *The Review of Economic Studies* 62:515–539.

Krusell, P., and A. A. Smith, Jr. 1998. Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy* 106:867–896.

La Porta, R., F. Lopez-De-Silanes, A. Shleifer, and R. W. Vishny. 1997. Legal determinants of external finance. *Journal of Finance* pp. 1131–1150.

La Porta, R., F. Lopez-de Silanes, A. Shleifer, and R. W. Vishny. 1998. Law and finance. *Journal of Political Economy* 106:1113–11.

Lamadon, T. 2014. Productivity shocks, dynamic contracts and income uncertainty. Tech. rep.

Li, S., T. M. Whited, and Y. Wu. 2015. Collateral, Taxes, and Leverage. *Available at SSRN 2360391* .

Marcet, A., and R. Marimon. 1992. Communication, commitment, and growth. *Journal of Economic Theory* 58:219–249.

Marcet, A., and R. Marimon. 2011. Recursive contracts .

Menzio, G., and S. Shi. 2010. Block recursive equilibria for stochastic models of search on the job. *Journal of Economic Theory* 145:1453–1494.

Menzio, G., and S. Shi. 2011. Efficient search on the job and the business cycle. *Journal of Political Economy* 119:468–510.

Moen, E. R. 1997. Competitive search equilibrium. *Journal of Political Economy* 105:385–411.

Moen, E. R., and Å. Rosén. 2011. Incentives in competitive search equilibrium. *The Review of Economic Studies* .

Monge-Naranjo, A. 2008. Limited commitment, firm dynamics and aggregate fluctuations. Tech. rep.

Peng, M. W., Y. Yamakawa, and S.-H. Lee. 2010. Bankruptcy laws and entrepreneur-friendliness. *Entrepreneurship Theory and Practice* 34:517–530.

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

Petersen, M. A., and R. G. Rajan. 1995. The effect of credit market competition on lending relationships. *The Quarterly Journal of Economics* 110:407–443.

Petrosky-Nadeau, N. 2014. Credit, vacancies and unemployment fluctuations. *Review of Economic Dynamics* 17:191–205.

Petrosky-Nadeau, N., and E. Wasmer. 2011. Macroeconomic dynamics in a model of goods, labor and credit market frictions .

Philippon, T., and P. Schnabl. 2013. Efficient recapitalization. *The Journal of Finance* 68:1–42.

Quadrini, V. 2004. Investment and liquidation in renegotiation-proof contracts with moral hazard. *Journal of Monetary Economics* 51:713–751.

Rampini, A. A., and S. Viswanathan. 2010. Collateral, risk management, and the distribution of debt capacity. *The Journal of Finance* 65:2293–2322.

Rampini, A. A., and S. Viswanathan. 2013. Collateral and capital structure. *Journal of Financial Economics* 109:466–492.

Rampini, A. A., and S. Viswanathan. 2014. Financial intermediary capital .

Rudanko, L. 2009. Labor market dynamics under long-term wage contracting. *Journal of Monetary Economics* 56:170–183.

Schaal, E. 2012. Uncertainty, productivity and unemployment in the Great Recession. *Federal Reserve Bank of Minneapolis, mimeo* .

Schmid, L. 2012. A Quantitative Dynamic Agency Model of Financing Constraints .

Silviera, R., and R. Wright. 2006. The venture capital cycle. *working paper* .

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

Stokey, N., and R. Lucas. with E. Prescott (1989): Recursive methods in economic dynamics.

Stulz, R. M. 1984. Optimal hedging policies. *Journal of Financial and Quantitative Analysis* 19:127–140.

Tauchen, G., and R. Hussey. 1991. Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica: Journal of the Econometric Society* pp. 371–396.

Thomas, J., and T. Worrall. 1994. Foreign direct investment and the risk of expropriation. *The Review of Economic Studies* 61:81–108.

Wasmer, E., and P. Weil. 2004. The Macroeconomics of labor and credit market imperfections. *American Economic Review* 94:944–963.

A   Appendix - Theory, Derivations and Proofs

A.1   Contracting problem

I rewrite below the general formulation of the contracting problem, and the derivation of the recursive multiplier formulation for both the exogenous and endogenous separation cases. The endogenous separation case introduce an additional difficulty because it potentially features a kink in the bank value function.

$$B(z, r_d, V) \quad = \quad \max_{K, d, \{V_{z', r'_d}\}} zK^\alpha - d - (\delta + r_d)K + \beta(1 - \sigma_{z, r_d})\mathbb{E}_{z, r_d}[B(z', r'_d, V_{z', r'_d})]$$

$$\text{s.t.}$$

$$V \geq u(\eta K) + \beta H(z) \tag{PC}$$

$$V = u(d) + \beta\left[(1 - \sigma_{z, r_d})\mathbb{E}_{z, r_d}[V_{z', r'_d}] + \sigma_{z, r_d}E_z[W(z')]\right] \tag{PK}$$

$$d \geq 0$$

a. Exogenous separation condition.

$$\sigma_{z, r_d} = \sigma_0$$

b. Endogenous separation condition.   In this section, I allow the firm to decide to separate from its current lender and go back to the credit market. This is a binary decision (either stay in the current lending relationships, or look for another bank), with reservation value given by $\mathbb{E}_z[W(z')]$.

$$\sigma_{z, r_d} = \sigma(z, r_d, \{V_{z', r'_d}\}) = \begin{cases} \sigma_0 & \text{if } \mathbb{E}_{z, r_d}[V_{z', r'_d}] \geq \mathbb{E}_z[W(z')], \\ 1 & \text{otherwise.} \end{cases}$$

Notations

$$
\begin{array}{rcl}
V & : & \text{current firm value} \\
V_{z', r'_d} & : & \text{state-contingent continuation value} \\
W(z) & : & \text{firm value when unfunded} \\
H(z) & : & \text{firm value after default} \\
K & : & \text{loan size} \\
d & : & \text{firm payout} \\
z & : & \text{aggregate shock} \\
r_d & : & \text{bank funding shock} \\
\beta & : & \text{discount factor} \\
\alpha & : & \text{DRS parameter} \\
\delta & : & \text{depreciation rate} \\
\sigma_0 & : & \text{probability of exogenous separation} \\
\xi & : & \text{probability of returning to credit markets after default}
\end{array}
$$

Proof. **Proposition 1. Lagrangian formulation.**

The problem above is not 'easily' solved using standard dynamic programming techniques because of the forward-looking nature of the participation constraint. The saddle-point problem methodology developed in Marcet and Marimon (2011) (and its earlier versions) allows for a more tractable approach based on the Lagrange multipliers corresponding these constraints, and provides a recursive formulation to the problem. I adapt this methodology to my problem as follows. Let us first define the Pareto problem $P(z, r_d, \Lambda) = sup_V B(z, r_d, V) + \Lambda V$ as,

$$P(z, r_d, \Lambda) = \sup_{V, K, d, \{V_{z', r_d'}\}} zK^\alpha - d - (\delta + r_d)K + \beta E[(1 - \sigma_{z', r_d'})B(z', r_d', V_{z', r_d'})] + \Lambda V$$

s.t.

$$V \geq u(\eta K) + \beta \xi H(z)$$
$$V = u(d) + \beta \mathbb{E}_{z, r_d}\left[(1 - \sigma_{z', r_d'})V_{z', r_d'} + \sigma_{z', r_d'} W(z')\right]$$

Replacing V in the above equation yields:

$$P(z, r_d, \Lambda) = \sup_{V, K, d, \{V_{z', r_d'}\}} zK^\alpha - d - (\delta + r_d)K + \beta E[(1 - \sigma_{z', r_d'})B(z', r_d', V_{z', r_d'})]$$
$$+ \Lambda \left[u(d) + \beta E_{z, r_d}[\sigma_{z', r_d'} W(z') + (1 - \sigma_{z', r_d'})V_{z', r_d'}]\right]$$

s.t.

$$u(d) + \beta \left[\mathbb{E}_{z, r_d}[(1 - \sigma_{z', r_d'})V_{z', r_d'}] + \sigma_{z', r_d'} W(z')\right] \geq u(zK^\alpha) + \beta H(z) \qquad (\gamma)$$

Eventually, including the participation constraint with weight γ and rearranging terms:

$$P(z, r_d, \Lambda) = \inf_\lambda \sup_{K, d, \{V_{z', r_d'}\}} zK^\alpha - d - (\delta + r_d)K + \beta(1 - \sigma_{z, r_d})E[B(z', r_d', V_{z', r_d'})]$$
$$+ \Lambda \left[u(d) + \beta E_{z, r_d}[\sigma_{z', r_d'} W(z') + (1 - \sigma_{z', r_d'})V_{z', r_d'}]\right]$$
$$- \lambda \left[u(\eta K) + \beta H(z)] - u(d) - \beta E_{z', r_d'}[\sigma_{z', r_d'} W(z') + (1 - \sigma_{z, r_d})V_{z', r_d'}]\right]$$
$$(\text{SPFE - 0})$$

a. Exogenous separation. The problem above becomes:

$$P(z, r_d, \Lambda) = \inf_\lambda \sup_{K, d} \quad zK^\alpha - d - (\delta + r_d)K + (\Lambda + \lambda)\left[u(d) + \beta \sigma_0 E_{z'}[W(z')]\right]$$
$$- \lambda \left[u(\eta K) + \beta \xi E_{z'}[W(z')]\right]$$
$$+ \beta(1 - \sigma_0)\left[E_{z, r_d}[\underbrace{\sup_{V_{z', r_d'}} B(z', r_d', V_{z, r_d}) + (\Lambda + \lambda)V_{z', r_d'}}_{P(z', r_d', \Lambda + \lambda)}]\right]$$

or equivalently,

$$P(z, r_d, \Lambda) = \inf_{\Lambda' \geq \Lambda} \sup_{K, d} \quad zK^\alpha - d - (\delta + r_d)K + \Lambda'\left[u(d) + \beta \sigma_0 E_{z'}[W(z')]\right]$$
$$+ (\Lambda - \Lambda')\left[u(\eta K) + \beta H(z)\right]$$
$$+ \beta(1 - \sigma_0)E_{z, r_d}[P(z', r_d', \Lambda')] \qquad (\text{SPFE})$$

Eventually, we can easily check that all the standard assumptions and regularity conditions needed for the application of theorems 1 and 2 in Marcet and Marimon (2011) are verified in order to justify that a solution to the saddle point problem is indeed equivalent to that of the original maximization problem. □

b. Endogenous separation. The bank now also takes into account the firm decision rule with regards to separation, which is here simply given by:

$$\sigma_{z,r_d} = \sigma(z', r_d', V_{z',r_d'}) = \begin{cases} \sigma_0 & \text{if } V_{z',r_d'} \geq W(z'), \\ 1 & \text{otherwise.} \end{cases}$$

Following the same steps as in the exogenous separation case, we get the following saddle-point problem[56]:

$$P(z, r_d, \Lambda) = \inf_{\lambda} \sup_{K,d} \quad zK^{\alpha} - d - (\delta + r_d)K - \lambda u(\eta K) + (\Lambda + \lambda)u(d) - \lambda\beta H(z)$$

$$+\beta(\Lambda + \lambda)\sigma_0 E_z[W(z')]$$

$$+\beta(1 - \sigma_0)E_{z,r_d}[\max((\Lambda + \lambda)W(z'), P(z', r_d', \Lambda + \lambda))]$$

A.1.1 Link between (SPFE) - (PP)

We can eventually recover the original bank profit and firm lifetime value through the definition of the Pareto problem as follows:

$$V(z, r_d, \Lambda) = \frac{\partial P}{\partial \Lambda}(z, r_d, \Lambda)$$
$$B(z, r_d, V) = P(z, r_d, \Lambda^*(z, r_d, V)) - \Lambda^*(z, r_d, V)V$$

Given the lagrange multiplier Λ', the state-contingent continuation values $V_{z',r_d'}$ can be obtained from the following first-order condition (when there is no separation):

$$\frac{\partial B}{\partial V}(z', r_d', V_{z',r_d'}) = -\Lambda'$$

A.1.2 Properties of the Lagrange multiplier

Conditional on loan rollover, the solution to the optimal contract verifies the following first-order conditions linking the lagrange multiplier Λ to K and d:

$$\frac{\partial F(z, K)}{\partial K} = \delta + r_d - (\Lambda - \Lambda')\eta u'(\eta K) \tag{15}$$

$$\frac{1}{\Lambda'} = u'(d) \tag{16}$$

Equation 15 determines the optimal level of capital as function of Lagrange multipliers (Λ, Λ'). When $\Lambda = \Lambda'$ (or equivalently $\lambda = 0$), the participation constraint is never binding and the firm is unconstrained. In such case, the level of

---

[56]Note that the problem below allows for exogenous separation on top of the endogenous separation choice.

capital is given by K_{FB}.

A.1.3 Properties of B

The following lemmas establish a series of properties of B that are useful to establish the remaining proofs in this section.

Lemma 3. $B(z, r_d, V)$ *is strictly increasing in z and decreasing in r_d.*

This result is straightforward and follows from the fact that the function $\pi(., ., V)$ defined in the intermediate problem 5 is strictly increasing in z and decreasing in r_d.

Lemma 4. $B(., ., V)$ *is strictly decreasing and concave in V in the continuation region, with a slope in $[-\frac{1}{u'(d)}, 0]$.*

This result stems from the observation that an increase in promised value V is always costly to the lender. The lower bound of the slope follows from equation (6). For example, when $u(d) = \frac{d^{1-\gamma}}{1-\gamma}$, and for constant z and r_d, this lower bound is given by $-\bar{d}^\gamma$. As in Lemma 3 in Albuquerque and Hopenhayn (2004), the concavity of $B(., ., V)$ follows directly from the concavity of function $\pi(., ., V)$ in 5 and Theorem 9.8 in Stokey and Lucas.

Lemma 5. $B(V, r_d; z, W)$ *is submodular in V and r_d.*

Proof. Let us fix the aggregate shock z and W without loss of generality and write $B(V, r_d) = (T_B)(V, r_d)$, where T is the operator mapping the set of continuous functions defined over $[W, V] \times [r_0, r_N]$ into itself.

Let us first define the surplus function

$$S(z, r_d, V) = \begin{cases} F(z, K_{cons}(z, V)) - (\delta + r_d)K_{cons}(z, V), & \text{if} \quad V < \underline{V}(z, r_d), \\ F(z, K_{FB}(z, r_d)) - (\delta + r_d)K_{FB}(z, r_d), & \text{if} \quad V \geq \underline{V}(z, r_d) \end{cases}$$

It is straightforward to show that S is continuous and differentiable in both r_d and V. Moreover, it is submodular in V and r_d. Indeed, we have:

$$\frac{\partial^2 S}{\partial V \partial r_d} = \begin{cases} \underbrace{-\frac{\partial K_{cons}(z, V)}{\partial V}}_{\leq 0}, & \text{if} \quad V < \underline{V}(z, r_d), \\ 0, & \text{if} \quad V \geq \underline{V}(z, r_d) \end{cases}$$

Let us now consider a function B to be submodular in (V, r_d), and let us write the cross-derivative of TB with respect to both V and r_d:

$$\begin{aligned} \frac{\partial^2 T(B)}{\partial V \partial r_d}(V, r_d) &= \frac{\partial S}{\partial V \partial r_d}(V, r_d) + \beta(1 - \sigma_0)\mathbb{E}\left[\frac{\partial^2 B}{\partial V \partial r_d}(V, r_d)\right] \\ &\leq 0 \end{aligned}$$

The operator T therefore maps the space of submodular functions into itself and the unique fixed point is also submodular.

$\square$

A.2 Proofs

Proof. **Corollary 1.**

Notice that $K_{cons}(z, V)$ satisfies the following expression in the constrained region:

$$u(\eta K_{cons}(z, V)) = V - \beta H(z) \tag{17}$$

The results follow immediately given that both utility and production functions are strictly increasing in K. Note however that this result is valid in partial equilibrium as both ξ and η can affect (negatively) W in the full model. $\square$

Proof. **Proposition 2.** In section 2. $\square$

Proof. **Proposition 3.** In section 2. $\square$

Proof. **Proposition 4.**

Let us first look at the case where the borrower is risk-neutral as in Albuquerque and Hopenhayn (2004). In this case, both agents are indifferent about the timing of consumption and it is always efficient to postpone dividend payouts in order to allow for the firm value to increase faster until the unconstrained region is reached. In this case, the promise-keeping constraint yields $V' = \frac{V}{\beta}$. Firm value is therefore always increasing in the constrained region at rate $\frac{1}{\beta}$.

Let us now look at the generalization of this result to risk-averse borrowers. In this case, agents are faced with two counteracting motives, namely consumption smoothing and higher savings incentives. The incentive for higher savings however doesn't dominate the agents willingness to grow out of the borrowing constraint. Indeed, assume by contradiction that $V' \leq V$, then $V < \frac{u(d)}{1-\beta}$ from the promise-keeping constraint. But the firm payout is increasing strictly whenever the participation constraint is binding, therefore firm value V must be at least greater than $\frac{u(d)}{1-\beta}$, and V' is strictly greater than V. $\square$

Proof. **Proposition 5.**

To prove this result, I write down the following equality derived from the envelope condition and the first order condition on V':

$$\frac{\partial B(z, r_{d,1}, V_{z',r_{d,1}})}{\partial V} = \frac{\partial B(z, r_{d,0}, V_{z',r_{d,0}})}{\partial V} = -\Lambda',$$

for all states $r_{d,0} \leq r_d, 1$. I then proceed by using the property of submodularity of $B(z, ., .)$ with respect to r_d and V derived in Lemma 5, which gives us immediately that if $r_{d,0} \leq r_{d,1}$, then necessarily $V_{z',r_{d,1}} \leq V_{z',r_{d,0}}$. $\square$

Let us now define $\bar{B}(z, r_d, V; W) = q(\theta(z, r_d, V; W))B(z, r_d, V; W)$ over the compact interval $\mathcal{S} = [W, \bar{S}]$, where $\bar{S}$ is the maximum value obtained by the entrepreneur when the joint surplus from the match is entirely kept by the firm.

Proof. **Lemma 1.** Existence of an interior solution V^*.

$\bar{B}$ is continuous over $\mathcal{S}$ as a product of two continuous functions over $\mathcal{S}$. The problem is therefore well defined and the solution to the maximization problem must also be in $\mathcal{S}$. But $\bar{B}(z, r_d, W, W) = \bar{B}(z, r_d, \bar{S}, W) = 0$, because $q(\theta(z, r_d, W; W)) = B(z, r_d, \bar{S}; W) = 0$ and the supremum of $\bar{B}$ must be strictly positive (for at least some r_d) to warrant bank entry in the first place, hence the solution maximizing $\bar{B}$ must be in $(W, \bar{S})$.

Moreover, we can show that V^* is unique if $\bar{B}(.,.,.;W)$ is indeed strictly concave over $\mathcal{S}$. To that end, let us define the function $f(V) = q \circ p^{-1}(\frac{\rho(z)}{\beta(V-W)}) = q(\theta(V))$. f is strictly increasing and strictly concave in V, thanks to the regularity properties of q and $q \circ p^{-1}$ (in particular the assumption that $q \circ p^{-1}$ is strictly decreasing and concave). Moreover, for z, r_d and W given, we can differentiate $\bar{B}^V(V) = \bar{B}(z, r_d, V, W)$ twice with respect to V, in order to get:

$$\bar{B}^{V''} = \underbrace{f''}_{<0}\underbrace{B}_{\geq 0} + 2\underbrace{f'}_{>0}\underbrace{B'}_{<0} + \underbrace{f}_{\geq 0}\underbrace{B''}_{<0}$$
$$< \quad 0.$$

$\bar{B}$ is therefore strictly concave in V, and has a unique supremum in $(W, \bar{S})$. □

Lemma 6. *B is decreasing in W.*

Proof. Since $\bar{B}(.,.,.,W)$ is continuous and decreasing in W, by the envelope theorem, its maximum over $S = [W, \bar{S}]$ must also be decreasing and continuous over the same interval $[W, \bar{S}]$. Therefore $B(.,W)$ defined in (11) is also decreasing in W over $S = [W, \bar{S}]$ whenever $B(.,W) > 0$. □

Proof. Lemma 2.
(i) We know from above that $\bar{B}$ is continuous and decreasing in W. Moreover, it is strictly decreasing in W whenever it is positive. If no entrant banks offers loans, expected bank profits cannot be positive. Hence, the entry condition implies the existence of at most one solution. A solution exists for sufficiently small c. $B(z, r_d, W, W)$ is strictly positive since $F'(0) = \infty$, and $\pi > 0$. When c is sufficiently small, the intermediate value theorem justifies the existence of a solution given that $B(, W) > 0$ (i.e. there exists a non-empty interval $[\underline{r}, r^*]$ such that $\bar{B}(z, r_d, W, W) > 0$) and $B(.,W)$ is strictly decreasing in W and $\lim_{W \to \bar{s}} B(.,W) = 0$. Moreover, when it exists, the solution is unique given that $B(.,W)$ is strictly monotonic in W.
(ii) Straightforward from equations (8), (12) and (13). □

Proof. **Proposition 6.** Credit markets in the cross-section.

Let us first show the following auxiliary lemmas before establishing this result.

Lemma 7. *$\bar{B}(V, r_d; z, W)$ is submodular in V and r_d.*

Proof. The submodularity of $\bar{B}$ with respect to V and r_d is a direct consequence of the submodularity of B and the convexity of q. Indeed, we have:

$$\frac{\partial^2 \bar{B}}{\partial V \partial r_d} = \underbrace{\frac{\partial q(V)}{\partial V}}_{\geq 0}\underbrace{\frac{\partial V}{\partial B \partial r_d}}_{\leq 0} + \underbrace{q(V)}_{\geq 0}\underbrace{\frac{\partial^2 T(B)}{\partial V \partial rd}}_{\leq 0}$$
$$\leq \quad 0$$

□

(i) Let us first show that V^* is decreasing in r_d. Let us fix z, and W without loss of generality, and define $V_0 = \arg\max_V \bar{B}(r_{d,0}, V)$. From the submodularity property of $\bar{B}$ shown in lemma 7, we have

$$0 = \frac{\partial q}{\partial V}B(V_0, r_{d,0}) + q(V_0)\frac{\partial B}{\partial V}(V_0, r_{d,0})$$
$$\geq \frac{\partial q}{\partial V}B(V_0, r_{d,1}) + q(V_0)\frac{\partial B}{\partial V}(V_0, r_{d,1})$$

57

Eventually, since $\bar{B}$ is strictly concave in V, then if $V_1 = \arg\max_V \bar{B}(r_{d,1}, V)$ exists such that $\frac{\partial q}{\partial V} B(V_1, r_d 1) + q(V_1)\frac{\partial B}{\partial V}(V_1, r_{d,1}) = 0$, it must be that $V_1 < V_0$.

Eventually, the properties of capital level at origination $K_0(r_d)$ and approval rate $p(\theta(V^*(r_d)))$ follow immediately from the property of of working capital policy (see proposition 2) and that of the matching probability p.

$\square$

Proof. **Proposition 7.** *Existence of a Block-Recursive Equilibrium.*
I prove the existence of a block-recursive equilibrium using Schauder's Fixed Point Theorem as stated in Stokey and Lucas - Theorem 17.4 and following the general exposition in Menzio and Shi (2010) and Schaal (2012).

Let us first define the set of functions $\mathcal{P} : \mathbb{Z} \times \mathbb{R} \times \mathbb{V} \to \mathbb{R}$ such that $\forall\, B \in \mathcal{P}$, B is: (i) bounded (ii) decreasing and concave in V, (iii) continuous and bi-Lipschitz in V. In order to apply Schauder's theorem, I proceed by showing the following properties: (a) Equilibrium objects W, θ, p and q are well defined and continuous; the operator T defined by the dynamic program in (3) (b-0) maps $\mathcal{P}$ into itself, (b-1) is continuous over $\mathcal{P}$, and (c) the family of functions $T(B)$ is equicontinuous.

(a) Existence, uniqueness and boundedness of $W_B(z)$, given $B \in \mathcal{P}$. First, for a given $B \in \mathcal{P}$, lemma 2 gives us the existence and uniqueness of W_B (assuming c is sufficiently small). The boundedness is immediate since W must lie in the compact set $\mathcal{S}_0 = [\underline{S}, \bar{S}]$.

Let us define $\mathcal{A} = (W, \frac{\bar{S}}{1-\beta}]$. The complementary slackness condition (13) tells us that either $\theta(V, z) = 0$ or $\exists\, a > 0$, such that $q(\theta(V, z))B(z, r_d, V; W) = a$. For $V \notin \mathcal{A}$, such a doesn't exist, and $\theta = 0$ in this region, otherwise for $V \in \mathcal{A}$, the above expression has a unique solution given by: $\theta(V, z) = q^{-1}\big(\frac{a}{B(z, r_d, V; W)}\big)$:

$$\theta(V, z) = \left\{ \begin{array}{ll} 0, & \text{if } V \in \mathcal{A}, \\ q^{-1}\big(\frac{a}{B(z, r_d, V; W)}\big), & \text{if } V \notin \mathcal{A} \end{array} \right.$$

Eventually, the existence and uniqueness of p and q follows immediately from the above results and equations (7)-(9).

(b-0). *The operator T is well-defined and maps $\mathcal{P}$ into itself.*
Let us consider $B \in \mathcal{P}$, and define $T_B = T(B)$.
1. T_B is continuous and concave in V. This is true since T_B is a linear combination of the auxiliary function π and B, which are both continuous and concave in V.
2. From the property above, T_B is differentiable (almost) everywhere and we can use the envelope theorem to show that the first-order derivative $\frac{\partial T_B(z, r_d, V)}{V} = -\frac{1}{u'(d(z, r_d, V))}$. We have already established that for a given pair $(z, r_d) \in \mathbb{Z} \times \mathbb{R}$, the dividend payout policy is bounded. The derivative of T_B is therefore also bounded (on both sides) and strictly negative. This is therefore also the case for T_B given that V is bounded. Eventually, T_B is decreasing in V and the bi-Lipschitz continuity property follows directly given that T_B is differentiable with first-derivative bounded on both sides. This concludes the proof of: $B \in \mathcal{P} \Rightarrow T_B \in \mathcal{P}$.

(b-1). *The operator T is continuous over $\mathcal{P}$.*
Let us introduce the infinite norm $\|.\|$ such that $\|B\| = \sup_{z, r_d, V \in \mathbb{Z} \times \mathbb{R}_d \times \mathbb{V}} B(z, r_d, V)$. Consider two functions $B_1, B_2 \in \mathcal{P}^2$. Fix r_d, z, V and their respective images $\hat{B}_1 = T\hat{B}_1$, and $\hat{B}_2 = T\hat{B}_2$. In order to establish continuity over $\mathcal{P}$, I need to show that $\forall\, l > 0$, s.t. $\|B_1 - B_2\| < l$, $\exists\, \epsilon > 0$ s.t. $\|\hat{B}_1 - \hat{B}_2\| < \epsilon$.

Let $\Phi_1 = (d_1, K_1, \{V_1'\})$ and $\Phi_2 = (d_2, K_2, \{V_2'\})$ be the optimal policies maximizing the bank's contracting problem associated with B_1 and B_2. Let us also consider the suboptimal policy $\tilde{\Phi}_2 = (\tilde{d}_2, \tilde{K}_2, \{V_1'\})$, where the vector of continuation values $\{V_1'\}$ is exactly the same as for policy Φ_1 and where $\tilde{d}_2$ and $\tilde{K}_2$ satisfy the corresponding promise-keeping and

participation constraints for B_2.

$$
\begin{aligned}
\|T_{B_1}(z, r_d, V) - T_{B_2}(z, r_d, V)\| &= \|F_{B_1}(z, r_d, V, \Phi_1) - F_{B_1}(z, r_d, V, \Phi_2)\| \\
&\leq \|F_{B_1}(z, r_d, V, \Phi_1) - F_{B_1}(z, r_d, V, \tilde{\Phi}_2)\| \\
&\leq \|\pi_1(\Phi_1) - \pi_2(\tilde{\Phi}_2) + \beta E[B_1 - B_2]\| \\
&\leq \|\pi_1(\Phi_1) - \pi_2(\tilde{\Phi}_2)\| + \beta \|B_1 - B_2\|
\end{aligned}
$$

We now need to show that there exists a finite upper bound α_T, such that the first component of the right-hand-side $\|\pi_1(\Phi_1) - \pi_2(\tilde{\Phi}_2)\|$ is bounded above by $\alpha_T \|B_1 - B_2\|$.

Technical assumption: F is bi-Lipshitz continuous in K, such that there exists upper and lower bounds $(\underline{\alpha}_F, \bar{\alpha}_F)$ such that:

$$
\underline{\alpha}_F |K_2 - K_1| < |F(., K_2) - F(., K_1)| < \bar{\alpha}_F |K_2 - K_1| \quad \forall (K_1, K_2)
$$

Notice that:

$$
\|\pi_1(\Phi_1) - \pi_2(\tilde{\Phi}_2)\| \leq \|d_1 - \tilde{d}_2\| + \|z K_1^\alpha - (\delta + r_d)K_1 - z\tilde{K}_2^\alpha + (\delta + r_d)\tilde{K}_2\| \tag{18}
$$

Let us first show the following auxiliary result which will be useful for establishing the bounds of the right-hand-side of the expression above.

Lemma 8. *For B_1, $B_2 \in \mathcal{P}^2$, we have*

1. $\|\theta_1 - \theta_2\| < \alpha_\theta \|B_1 - B_2\|$
2. $\|p_1 - p_2\| < \alpha_p \|B_1 - B_2\|$
3. $\|W_1 - W_2\| < \alpha_W \|B_1 - B_2\|$

Proof. Consider (z, r_d) given.

1. Using the definition of $\bar{B}$, we have $c \geq \bar{B}(z, r_d, V, \theta)$.

   Assume market V is open (for both B_1 and B_2). We have

$$
\begin{aligned}
0 &= B_1(V_1)q(\theta_1) - B_2(V_2)q(\theta_2) \\
&= [B_1(V_1) - B_2(V_2)]q(\theta_1) + B_2(V_2)[q(\theta_1) - q(\theta_2)] \\
&\leq [B_1(V_1) - B_2(V_2)] + \|B_2\|[q(\theta_1) - q(\theta_2)]
\end{aligned}
$$

but q is convex, hence

$$
q(\theta_1) - q(\theta_2) \leq q'(\max(\theta_1, \theta_2))(\theta 1 - \theta_2)
$$

and

$$
-\|B_2\|[q(\theta_1) - q(\theta_2)] \leq [B_1(V_1) - B_2(V_2)]
$$

and

$$
-\|B_2\|q'(\max(\theta_1, \theta_2))(\theta_1 - \theta_2) \leq [B_1(V_1) - B_2(V_2)]
$$

We can eventually write the following inequality, using $\|B_2\| \leq c$, and defining $\alpha_\theta = \frac{1}{cq'(\max(\theta_1,\theta_2))}$:

$$|\theta_1 - \theta_2| \leq \frac{1}{cq'(\max(\theta_1,\theta_2))}\|B_1 - B_2\|$$

2. Fix V. We have by definition

$$p_1(z, r_d, V) - p_2(z, r_d, V) = p(\theta_1(z, r_d, V)) - p(\theta_2(z, r_d, V))$$

We can therefore write the following inequality thanks to the concavity of p:

$$|p_1(z, r_d, V) - p_2(z, r_d, V)| \leq p'(0)|\theta_1(z, r_d, V)) - \theta_2(z, r_d, V)| \leq p'(0)\alpha_\theta\|B_1(V_1) - B_2(V_2)\|$$

3. We have:

$$
\begin{aligned}
W_1(z, r_d) - W_2(z, r_d) &= \beta E[W_1(z', r') - W_2(z', r') + p_1(\theta_1)(V_1 - W_1) - p_2(\theta_2)(V_2 - W_2)] \\
&\leq \beta E[W_1(z', r') - W_2(z', r') + p_1(\theta_1)(V_1 - W_1) - p_2(\theta_1)(V_1 - W_2)] \\
&\leq \beta E[W_1(z', r') - W_2(z', r') + V_1(p_1(\theta_1) - p_2(\theta_1)) - p_1(\theta_1)W_1 + p_2(\theta_1)W_2] \\
&\leq \beta E[W_1(z', r') - W_2(z', r') + V_1(p_1(\theta_1) - p_2(\theta_1)) - p_1(\theta_1)(W_1 - W_2) + (p_2(\theta_1) - p_1(\theta_1))W_2] \\
&\leq \beta E[(1 - p_1(\theta_1))|W_1(z', r') - W_2(z', r')| + (V_1 - W_2)|(p_2(\theta_1) - p_1(\theta_1))|]
\end{aligned}
$$

We can eventually proceed by re-arranging terms to obtain the following inequality:

$$
\begin{aligned}
|W_1(z', r') - W_2(z', r')| &\leq \frac{\beta}{1-\beta}(\max(V_1) + \max(W_2))|(p_2(\theta_1) - p_1(\theta_1))| \\
&\leq \frac{\beta}{1-\beta}(\max(V_1) + \max(W_2))p'(0)\alpha_\theta\|B_1 - B_2\|
\end{aligned}
$$

$\square$

We can now go back to inequality (18) to establish our continuity result:

$$\|\pi_1(\Phi_1) - \pi_2(\tilde{\Phi}_2)\| \leq \|d_1 - \tilde{d}_2\| + \|z\tilde{K}_1^\alpha - (\delta + r_d)\tilde{K}_1 - z\tilde{K}_2^\alpha + (\delta + r_d)\tilde{K}_2\|$$

Let us look at each component of the right-hand separately. First, we want to bound $\|d_1 - \tilde{d}_2\|$. To establish this result, let us first notice that

$$\min\{u'\}\|d_1 - \tilde{d}_2\| \leq \|u(d_1) - u(\tilde{d}_2)\|$$

and

$$
\begin{aligned}
\|u(d_1) - u(\tilde{d}_2)\| &= \|V - \beta\mathbb{E}[\sigma_0 W_1(z) + (1 - \sigma_0)V_1'] - [V - \beta\mathbb{E}[\sigma_0 W_2(z) + (1 - \sigma_0)V_1']]\| \\
&\leq \beta\mathbb{E}[\|W_1 - W_2\| \\
&\leq \beta\alpha_W\|B_1 - B_2\|
\end{aligned}
$$

Second, we want to bound $\|zK_1^\alpha - (\delta + r_d)K_1 - z\tilde{K}_2^\alpha + (\delta + r_d)\tilde{K}_2\|$.

Fix (z, r_d, V),

$$u(\eta K_1) - u(\eta \tilde{K}_2) = \beta E_z[H_1 - H_2]$$

But, we have

$$H_1 - H_2 = \xi E_z[W_1 - W_2] + (1 - \xi)\beta E_z[H_1 - H_2]$$

from definition of H given by equation (2), and

$$
\begin{aligned}
|H_1 - H_2| &\leq \frac{\xi}{1 - \beta(1 - \xi)}|W_1 - W_2| \\
&\leq \frac{\xi}{1 - \beta(1 - \xi)}\alpha_W \|P_1 - P_2\|
\end{aligned}
$$

Eventually, using the concavity property of u to get:

$$(K_1 - \tilde{K}_2) \leq \frac{1}{\eta u'(\eta \max(K))}[u(\eta K_1) - u(\eta \tilde{K}_2)]$$

and

$$|K_1 - \tilde{K}_2| \leq \frac{\beta}{\alpha_k}\|B_1 - B_2\|$$

with $\alpha_k = \eta u'(\eta \bar{K})\frac{\xi}{1 - 1 - \beta(1 - \xi)}\alpha_W$, and $\bar{K} = \max K_{FB}$.

Eventually, the result above also goes through for any concave function of K (adjusting the multiplicative term by $\frac{1}{F'(K)}$), with and we can write

$$\|T_{B_1}(z, rd, V) - T_{B_2}(z, rd, V)\| \leq \beta\left(1 + \alpha_W + \frac{\alpha_k}{F'(\bar{K})}\right)\|B_1 - B_2\|$$

which completes the proof.

(c) Equicontinuity of $T(\mathcal{P})$.

Let us show that $\forall\, \epsilon > 0$, $\exists\, \delta > 0$, such that for all $\nu_i = (z_i, r_{d,i}, V_i)$, $i = 1, 2$

$$\|\xi_1 - \xi_2\| < \delta \Rightarrow TB(\nu_1) - TB(\nu_2) < \epsilon, \ \forall\, B \in \mathcal{P}$$

Fix $\epsilon > 0$, and pick $\delta < \min\left(\min_{(z_1,z_2)\in\mathbb{Z}}|z_1 - z_2|, \min_{(r_{d,1},r_{d_2})\in\mathbb{R}_d}|r_{d,1} - r_{d,2}|, \frac{\epsilon}{\alpha_V}\right)$.

For ξ_1, ξ_2 such that $\|\xi_1 - \xi_2\| < \delta$, we have $z_1 = z_2$, and $r_{d,1} = r_{d,2}$.

We can therefore conclude that:

$$\|TB(\xi_1) - TB(\xi_2)\| \leq \alpha_v|V_1 - V_2| \leq \alpha_v\|\xi_1 - \xi_2\| < \epsilon$$

Now that we have shown that assumptions (i), (ii), and (iii) are verified, Schauder's fixed point theorem applies and there exists a fixed point $B^* \in \mathcal{P}$ such that $T(B^*) = B^*$. Eventually, all the remaining equilibrium objects (W^*, ρ^*, θ^*) and policy functions associated with the optimal contract are also well defined; which concludes the existence of a Block-Recursive Equilibrium. $\square$

Proof. **Proposition 8.** Social Planner's problem.

Let us start by simplifying some of the notations of the model before formalizing the social planner's problem. Let us denote $\theta_V = \theta(V)$, the market tightness associated with firm value V, and V_r^0 the optimal firm value offered by the banks with funding costs r. Let us also abstract from the variables dependence on aggregate shocks, and the bank-firm characteristics $(r_{d,t-1}, V_{t-1})$ carried from period $t-1$.

The social planner maximizes the discounted sum of utilities derived by banks and firms for incumbent lending relationships, utility derived by rationed entrepreneurs, minus total entry costs incurred by loan origination. The problem is subject to the dynamics of the existing lending contracts represented by the function f_c (which only depends on (V_{t-1}, r_{t-1}, r_t)), and the laws of motion for credit rationing v_t, and the distribution of lending relationships g_t. In order to simplify notations, notice that the social planner's faces the same contracting frictions as each individual bank, hence we can immediately replace the original problem with the corresponding solution to its Lagrange multiplier formulation (SPFE), and by taking the optimal weights on firm value to be $\Lambda_{t+1} = \frac{1}{u'(d_t)}$. The social planner therefore maximizes the following objective function:

$$
\max_{u_t, g_t, \theta_V, J_t, V_t} \quad \mathbb{E}_z \sum_t \beta^t \left[\sum_{r_{d,t}, V_t} (1-\sigma_0) g(r_{d,t}, V_t)[S(V_t, r_{d,t}) - d(V_t, r_{d,t}) + \lambda(V_t, r_{d,t})u(d(V_t, r_{d,t}))] - cJ_t + v_t u(d_0) \right]
$$

$$
s.t. \ \forall \ (t, z^t)
$$

$$
\lambda(V_t, r_{d,t}) \ = \ \frac{1}{u'(d(V_t, r_{d,t}))}, \qquad\qquad\qquad\qquad \forall \ (r_t, V_t)
$$

$$
V_t \ = \ f_c(V_{t-1}, r_{t-1}, r_t), \qquad\qquad\qquad\qquad \forall \ (r_{t-1}, V_{t-1}, r_t)
$$

$$
v_t \ = \ v_{t-1}(1 - \sum_r \Gamma_r(r)p(\theta(V_r))) + (1 - v_{t-1})\sigma_0
$$

$$
g_t(r, V) \ = \ \sum_{V_{t-1}|V_t=V} (1-\sigma_0) g_{t-1}(r_{t-1}, V_{t-1})\pi_r(r_{t-1}, r_t) + J_t q(\theta(V_r))\Gamma_r^0(r) 1_{V_t=V_r}, \qquad \forall \ (r, V)
$$

where Γ_r is the transition probability and Γ_r^0 is the unconditional distribution of the idiosyncratic funding cost.

The planner's problem is also constrained by the credit market clearing conditions which imply that total number of funded entrepreneurs equals the total number of loans originated within each active submarket. In the context of the model, this is simply given by the following standard condition:

$$
v_{t-1}p(\theta(V_r)) = q(\theta(V_r))\Gamma_r^0(r)J_t, \quad \forall \ \theta > 0
$$

In order to characterize this problem further, let us now denote μ multiplier on law of motion of v, and $\{\zeta_\theta\}_V$ the set of multipliers associated with the market clearing condition for each active submarket and write the corresponding generalized Lagrangian expression:

$$
\max_{u_t, g_t, \theta_V, J_t, V_t} \quad \mathbb{E} \sum_t \beta^t \Bigg(\sum_{r_{d,t}, r_{d,t+1}, V_{t-1}} (1-\sigma_0) g_{t-1}(r_{d,t-1}, V_{t-1})\Gamma_r(r_{d,t-1}, r_{d,t})[S(V_t, r_{d,t}) - d(V_t) + \frac{1}{u'(d(V_t))}u(d(V_t))]
$$

$$
+ \ J_t \sum_r \Gamma_r^0(r)q(\theta(V_r)) \left[[S(V_r, r) - d(V_r, r) + \frac{1}{u'(d(V_r))}u(d(V_r))] - c \right]
$$

$$
- \sum_{V, \theta_V > 0} \zeta_{\theta_V, t} \sum_{r, V(r)=V} [\Gamma_r(r_d)J_t q(\theta(V)) - p(\theta(V))v_{t-1}]
$$

$$
+ \ v_t u(d_0) - \mu_t \Big(v_t - v_{t-1}(1 - \sum_r \Gamma_r^0(r)p(\theta(V_r))) - \sigma_0(1 - v_{t-1}) \Big) \Bigg)
$$

Replacing $\varphi_V = q(\theta_V)$ in the above objective function allows us to have a well-defined and strictly concave problem in all its maximands over a convex set. This in turn is sufficient to establish the existence and uniqueness of an optimal solution to the social planner's problem.

We can eventually decompose the above expression into 3 auxiliary and independent problems for: (i) incumbent banks, (ii) entrant banks, and (iii) unfunded firms; which are all independent from the the distribution of lending relationships at $t-1$, g_{t-1}:

(i) *Incumbent banks*:

$$\mathbb{E}\left[\sum_t \beta^t \pi(r_{d,t-1}, r_{d,t})[S(V_t, r_{d,t}) - d(V_t, r_{d,t}) + \frac{1}{u'(V_t, r_{d,t})} u(V_t, r_{d,t})]\right]$$

(ii) *Entrant banks*:

$$\max_{\theta_V, \{V_r\}_r, J_t} \quad J_t[\sum_r \Gamma_r(r) q(\theta_{V_r}) \left[\sum_{t' \geq t} \beta^{t'-t} \left[S(V_r, r) - d(V_r, r) + \frac{1}{u'(d(V_r))} u(d(V_r))\right] - c\right]$$
$$- \sum_V \sum \beta^t \zeta_{\theta_V,t} \Gamma_r^0(r) q(\theta_{V_r})]$$

(iii) *Unfunded firms*:

$$\max_{v_t} \sum_t \beta^t \left[v_t u(d_0) - \mu_t \left(v_t - v_{t-1}(1 - \sum_r \Gamma^0(r_d) p(\theta_{V_r})) - \sigma_0(1 - v_{t-1})\right) + v_{t-1} \sum_V \sum_r \zeta_{\theta_{V_r}} p(\theta_{V_r})\right]$$

The problem for incumbent banks has been 'maxed out' from the beginning and is independent of the other two problems. This is because once a lending relationship is formed, the problem is such that the optimal contracting decisions do not depend directly on the two other problems.

Turning to the entrant banks problem, we can write for each bank of type r,

$$\max_{\theta_r} \sum_{t' \geq t} \beta^{t'-t} \left[S(V_r, r) - d(V_r, r) + \frac{1}{u'(d(V_r))} u(d(V_r))\right] - \beta^t \zeta_t(V)$$

The above problem is exactly the same as the entrant problems associated with the competitive equilibrium if:

$$\zeta_t(V) = \sum_{t' \geq t} \beta^{t'-t}(1 - \sigma_0)^{t'-t} \frac{u(d(V_r))}{u'(d(V_r))}$$

Eventually, turning to the unfunded firms problem, the first order condition on v_t yields:

$$0 = u(d_0) - \mu_t + \beta \mu_{t+1}[(1 - \sum \pi_0(r_d) p(\theta(V_r)))] + \beta \sum \zeta_{t+1}(V_r) \pi(r_d) p(\theta(V_{r,t+1}))$$

We can now identify $\mu_t = W(z_t)$, and rearrange terms of the above condition to obtain the following identity, which is

satisfied within each active submarket:

$$
\begin{aligned}
W &= u(d_0) + \beta \left[(1 - p(\theta(V_r)))W_{t+1} + p(\theta(V_r))\zeta_{t+1}(V_r) \right] \\
&= u(d_0) + \beta \left[(1 - p(\theta(V_r)))W_{t+1} + p(\theta(V_r)) \sum_{t' \geq t} \beta^{t'-t}(1 - \sigma_0)^{t'-t} \frac{u(d(V_r))}{u'(d(V_r))} \right]
\end{aligned}
$$

The above expression resembles the one obtained in the competitive search equilibrium (7) - (8), with one major difference. Indeed, the equivalence between the two expressions is only true if $\sum_{t' \geq t} \beta^{t'-t}(1 - \sigma_0)^{t'-t} \frac{u'(d(V_r))}{u'(d(V_r))} = V$, which is only the case when entrepreneurs are risk-neutral, and $u' = 1$. Therefore, in general, the above result establishes that the obtained competitive search equilibrium is always constrained-inefficient whenever entrepreneurs are risk-averse.

This concludes the proof. □

B   Appendix - Computational Methodology

This section summarizes the numerical procedure used to solve the dynamic contracting problem and the competitive search equilibrium. This section also explicitly defines the dynamics of the distribution of lending relationships, credit rationing, and aggregate credit in the economy.

B.1   General equilibrium

Solving the model in general equilibrium corresponds to solving a fixed-point problem in unfunded firm value W.

- Initialize $\{W(z)\}$, with W increasing in z.
- Loop over following steps below until convergence:

  1. Given W, solve recursively the contracting problem formulated in the Lagrange multiplier space as described in section to get $P(z, r_d, \Lambda)$, and determine the corresponding contract value to the bank in the promised utility space $B(z, r_d, V)$.

  2. Determine the indifference curve for the firm and compute $\{\rho(z)\}$ based on the following identity:

$$
\begin{aligned}
\rho(z) &= p(\theta(z, V))(V - W(z)) \\
\rho(z) &= \frac{W(z) - \beta E_z[W(z')]}{\beta}
\end{aligned}
$$

  3. For each idiosyncratic state r_d, and taking (IC) as given, solve for the contract value V optimizing bank value:

$$
\bar{B}(z, r_d, V^*(z, r_d)) = \max_V q(\theta(z, V))E[\Pi(z, r_d, V)]
$$

  4. For each aggregate state z, update $W(z)$, either upwards if $k_e < E_{r_d}[\bar{B}(z, r_d, V^*(z, r_d))]$, or downwards otherwise.

B.2   Laws of motion

The model generates a stationary distribution for bank lending relationships. This distribution dynamics of lending relationships and the associated laws of motion for aggregate capital and the credit rationing in the economy (i.e. mass of unfunded entrepreneurs) are given by the following expressions:

Here, I fix z to simplify notations and write these laws of motions in the Lagrange multiplier space:

- **Distribution dynamics of lending relationships**

$$g_{t+1}(r_{d,t+1}, \gamma_{t+1}) = \int_{\gamma_{t+1}=\gamma'(r_{d,t+1},\gamma_t)} (1 - \sigma(r_t^d, \gamma_t)) g_t(r_t^d, \gamma_t) \Pi_r(r_{d,t}, r_{d,t+1}) dr_{d,t} dr_{d,t+1} d\gamma_t + J_{t+1}(r_{d,t+1}, \gamma_{t+1})$$

where $J_{t+1}(r_{d,t+1}, \gamma_{t+1})$ represents the measure of new bank entrants with multiplier γ_{t+1}, and funding cost $r_{d,t+1}$.

- **Credit rationing**

$$u_{t+1} = u_t(1 - \bar{p}_t) + \int \sigma(r_t^d, \gamma_t) g_t(r_t^d, \gamma_t) \Pi_r(r_t^d, r_{t+1}^d) dr_{d,t} dr_{d,t+1} d\gamma_t$$

where $\bar{p}_t = \int p(\theta(r_{d,t})) J_t(r_{d,t}, \gamma_{t+1}) dr_{d,t}$.

- **Aggregate credit supply**

$$L_{t+1} = \int \int \int (1 - \sigma(r_t^d, \gamma_t)) g_t(r_t^d, \gamma_t) K(r_{t+1}^d, \gamma_{t+1}) \Pi_r(r_t^d, r_{t+1}^d) dr_{d,t} dr_{d,t+1} d\gamma_t + \int J_{t+1}(r_{d,t+1}, \gamma_{t+1}) dr_{d,t+1}$$

Note that in the exogenous case, $\sigma(r_t^d, \gamma_t)$ is constant and given by $\sigma = \sigma_0$. The first term of the last expression corresponds to the contribution of the intensive margin to the aggregate credit supply, while the second term takes into account the increase in aggregate supply due to the extensive margin (creation of new lending relationships).

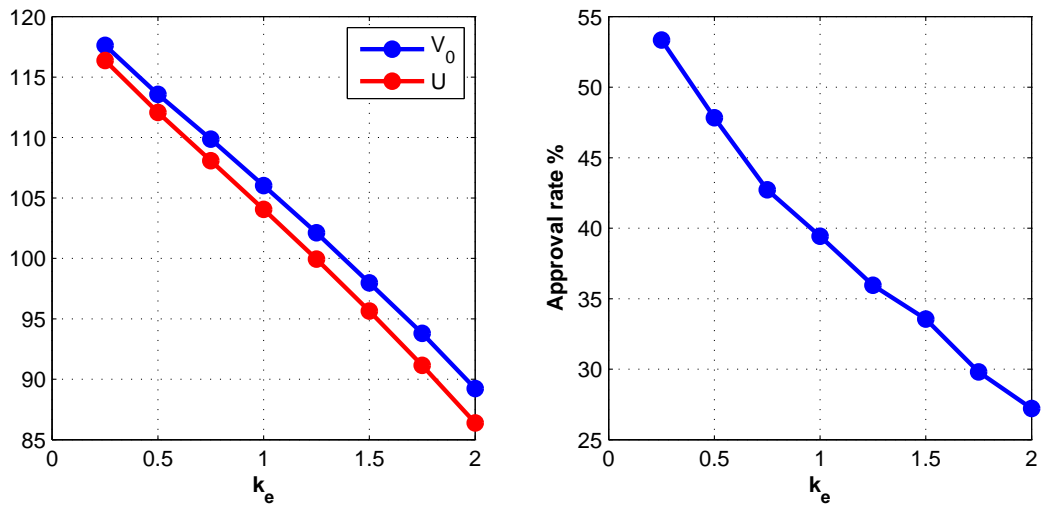# C Appendix - Figures

C.1 Comparative statics
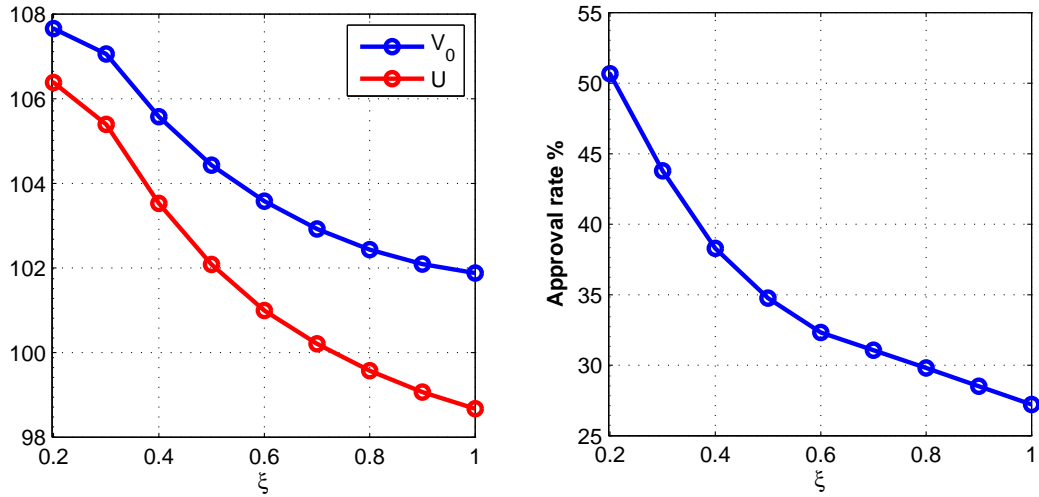
(a) Credit market variables - Firm value (at origination (blue), unfunded (red)) (left panel), and approval rates (right panel), for fresh start probability ξ.
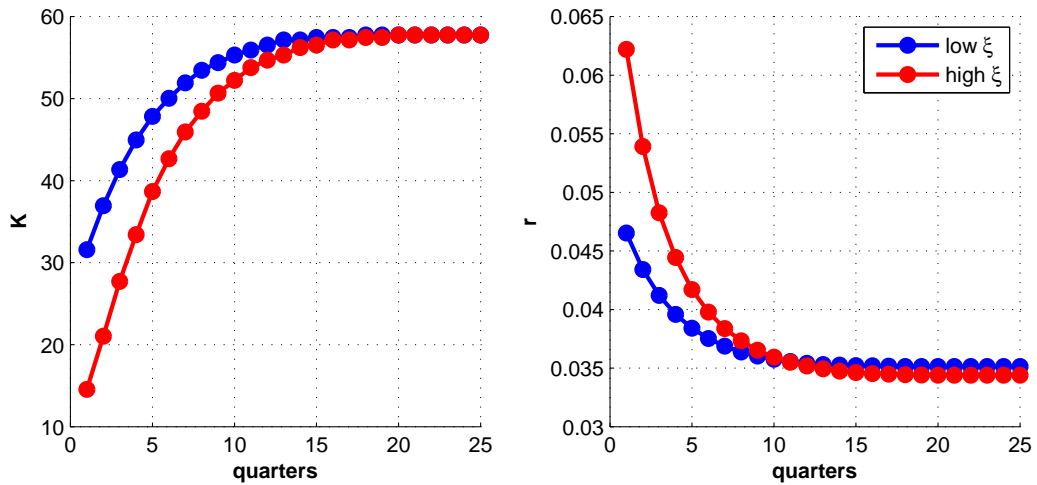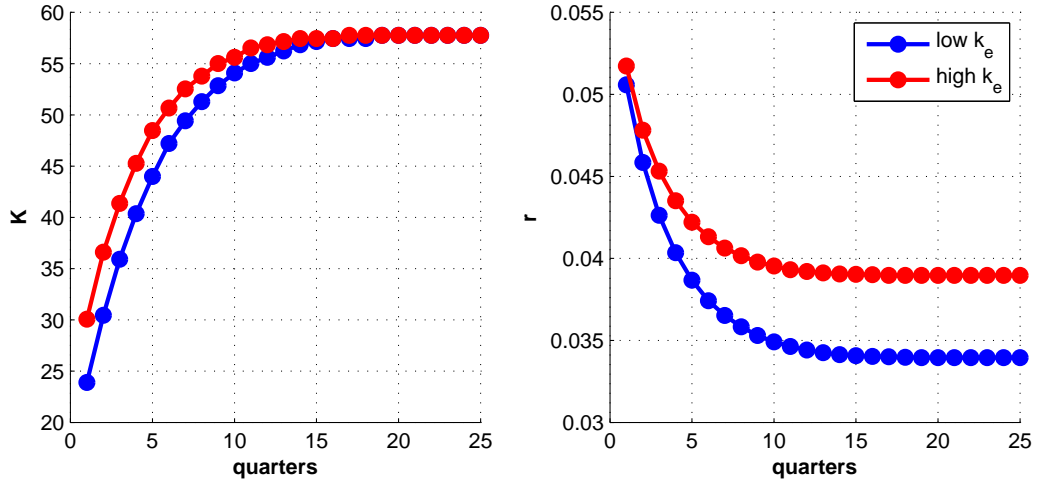

(b) Firm dynamics - Credit availability (left panel) and lending rates (right panel) across the length of the lending relationship, for low (blue) and high (red) levels of fresh start probability ξ.

Figure 15. Comparative statics - ξ.

(a) Credit market variables - Firm value (at origination (blue), unfunded (red)) (left panel), and approval rates (right panel), for entry cost k_e.

(b) Firm dynamics - Credit availability (left panel) and lending rates (right panel) across the length of the lending relationship, for low (blue) and high (red) levels of entry cost k_e.

Figure 16. Comparative statics - c.

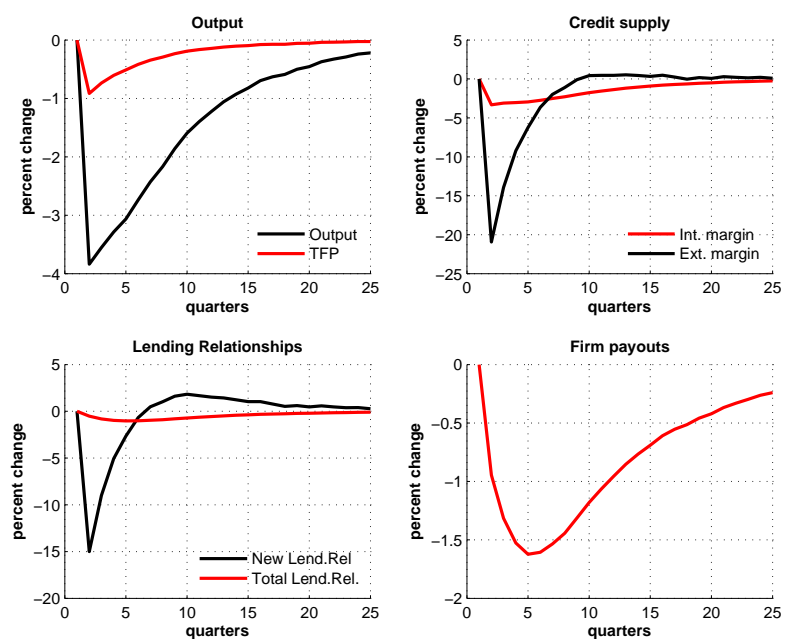## C.2 Impulse response: Productivity shocks

Figure 17. Impulse response - negative productivity shock z. Aggregate variables.
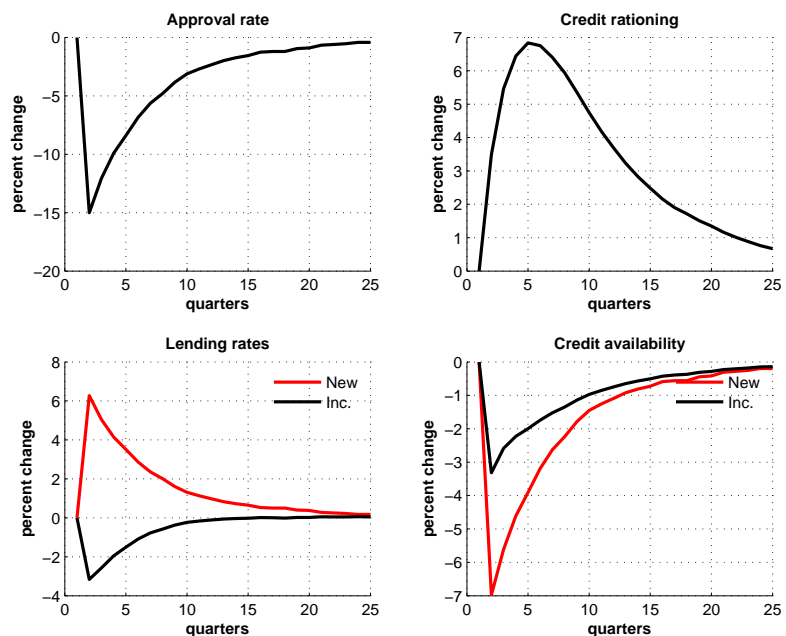

Figure 18. Impulse response - negative productivity shock z. Credit markets and contractual terms.

D Appendix - Empirical section

D.1 Bank lending relationships

I analyze in this section the dynamics of creation and destruction of credit relationships. To that end, I primarily use the Loan Pricing Corporation DealScan database (LPC thereafter) and focus on US loan syndications denominated in US Dollars. I construct time series for positive, negative, and net relationship flows based on a sample covering the period from January 1986 to December 2013.

In order to measure these flows, I need to determine the date of inception and termination of each bank-firm pair forming a lending relationship. First, since LPC mainly comprises syndicated loans with potentially many participants, I only consider the lead arranger and/or main agent for each loan package. When a given loan package involves more than one lead arranger, I consider the ensuing lending relationship for each lender separately.

For a given bank-firm pair, I define $\bar{m}_t$ as the maximum maturity date recorded across all loan agreements made up to a given date t. At date t, a bank-firm pair is considered *inactive* if it has never been matched up to this date, or if it has been matched in the past but no new transaction took place in the 3 years following the maturity date of all previous deals, i.e. $\bar{m}_t < t - 3Y$ [57]. In the former case, the bank-firm pair is considered as inactive starting from date $\bar{m}_t$. Otherwise, it is considered as active between the corresponding dates of inception and termination. The date of inception is defined as the date in which an inactive bank-firm pair is formed, while the date of termination is given by the date in which an existing and active bank-firm pair becomes inactive.

In the benchmark case, I do not control for bank mergers. However, I also consider a more general definition aggregating all lenders matched to a given firm. This allows for the control of possible bank switching and potential bank mergers (and therefore prevents churning effects from driving the results). The results remain qualitatively similar in this case.

I track the date of inception and termination of each lending relationship and construct the aggregate portfolio of lending relationships of the 20 most active U.S. banks. Additional details and robustness checks related to (i) multiple lending relationships, (ii) inactivity periods, and (iii) flow decomposition by firm size are available upon request.

Finally, I construct time series for the stock of lending relationships based on the cumulated net flows over the period 1986 to 1997. I then use the stock at Jan 1997 as my reference point to define creation and destruction rates. This is a reasonable assumption since the average duration of a lending relationship is about 8 years. The time series for creation and destruction rates are considered for the period 1997 to 2013. I chose to start in 1997 to ensure that the database is already well-populated and hence avoid spurious changes in entry rates due to improved firm coverage in LPC. I also consolidate bank names reported in the database to prevent other spurious creations or destructions of lending relationships.

	Net flows	Positive flows	Negative flows
mean (%)	0.94	4.09	3.15
std(%)	1.34	1.28	0.53
corr(x,GDP)	0.44	0.44	-0.18
corr(x,Fed Fund rates)	0.60	0.53	-0.43

Table 3. Lending relationship flow properties (quarterly). GDP refers to the cyclical component of detrended log GDP. Fed fund rates are in levels and correspond to the effective real Fed fund rates. Excess reallocation is defined as the sum of positive and negative flows minus the absolute value of net flows. Sample period: 1997Q1-2013Q4.

---

[57]I also consider alternative measures using 1 and 5 years as a robustness check. These tests yield qualitatively similar flow patterns and are available upon request.
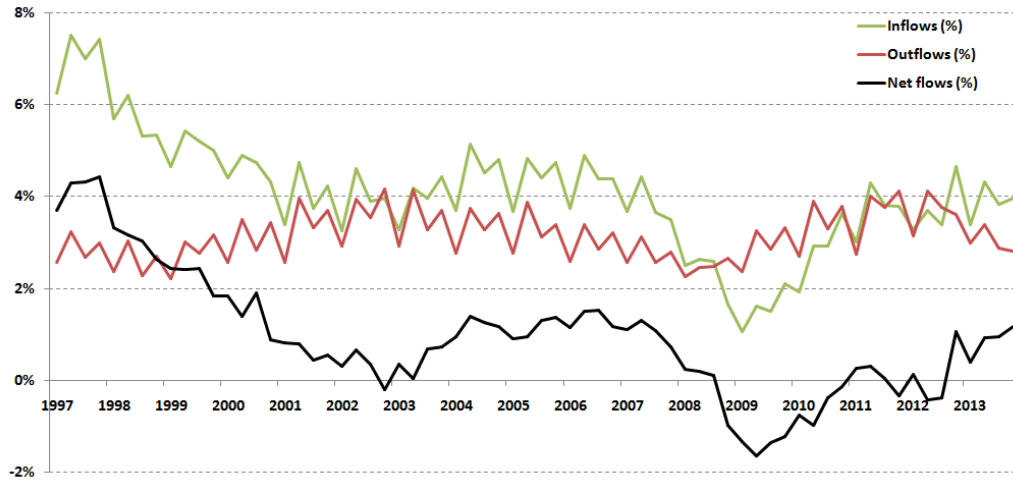
Figure 19. Deconstructing lending relationships flows (1997Q1-2013Q4). Author's calculations tabulated from DealScan database.
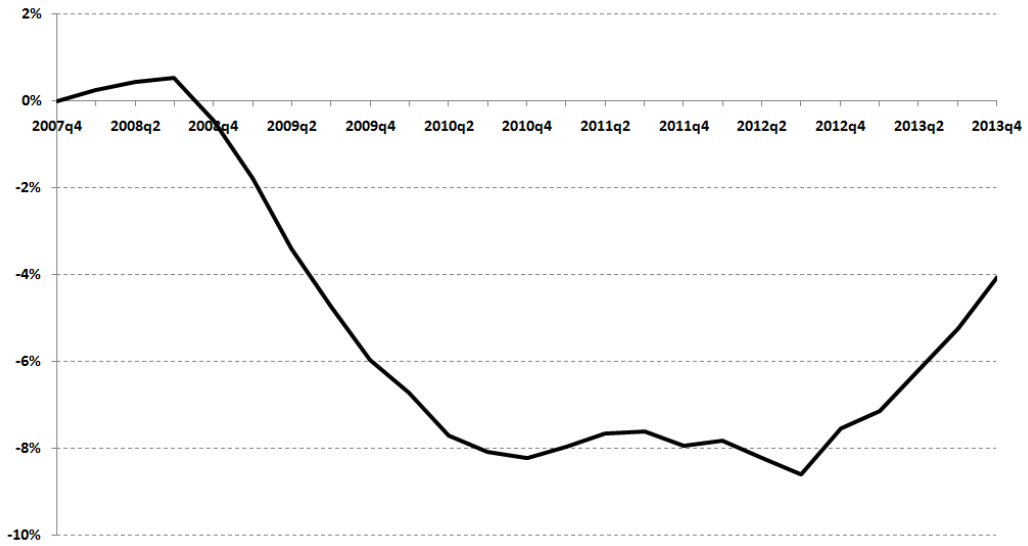

Figure 20. Changes in the stock of lending relationships (log-deviations from trend based on HP-filter with parameter 1600, 2007Q4-2013Q4).

D.2 Approval rates

I use the index constructed by Biz2Credit for Small and Large banks for the period Jan2011-May2014 (detailed information about the composition of applicants and lenders within each category is not available yet).
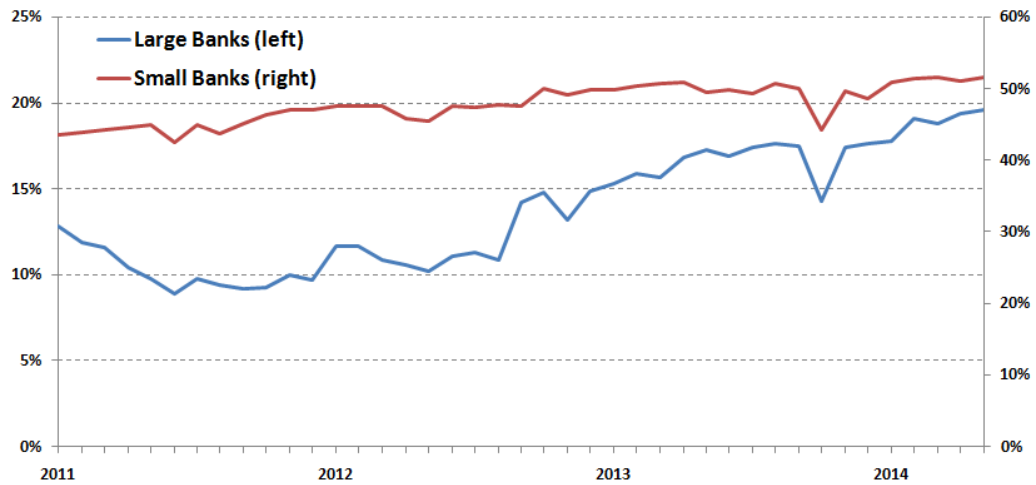

Figure 21. Approval rates for large banks ($10 billion+ in assets) and small banks (January 2011 - May 2014). The approval rate is defined as the share of granted funding requests, within each lender category. These requests are for small business loans (below $3 million). Source: Biz2Credit.com.