

Pre-crisis credit standards: monetary policy or the savings glut?

Adrian Penalver*

5 December 2013

Abstract

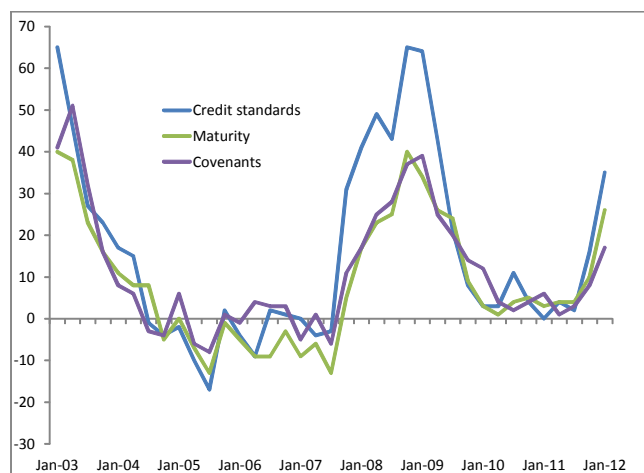
This paper presents a theoretical model of bank credit standards using a framework very different from the established banking literature. It examines how a monopoly bank sets its monitoring intensity in order to manage credit risk when it makes long duration loans to borrowers who have private knowledge of their project's stochastic profitability. In contrast to standard models, it has a recursive structure and a general equilibrium. The bank loan contract considered specifies the interest rate, the monitoring intensity and a profitability covenant. Within this class of contract, the bank chooses the terms which maximise steady-state profits subject to the constraint that it must have as many deposits as loans. The model is then used to consider whether the reduction in credit standards and credit spreads observed before the financial crisis could have been caused by low official interest rates or a positive deposit shock. The model rejects a risk-taking channel of monetary policy and endorses the savings glut hypothesis.

1. Introduction

One of the most pernicious aspects of credit booms is that they are often associated with a reduction in bank lending standards. During credit booms less creditworthy

*Paris School of Economics. This paper has benefitted from comments received at a seminars at the Paris School of Economics, the Australian National University, ESSEC, the Bank of England, the Royal Economic Society annual conference and the European Economic Association annual congress. I am also extremely grateful to John Stachurski for a substantial simplification and generalisation of the proof.

Figure 1: ECB Bank Lending Survey Credit Standards



firms and households obtain bank loans on easier terms than they would at other points in the credit cycle. If boom turns to bust (which doesn't always happen), then these lower credit quality loans suffer larger losses, impair the balance sheets of the banks for some time and retard the growth of new lending and economic recovery. A sustained period of lax credit standards increases the latent credit risk in the banking system and thereby the risk of economic and financial instability¹. The continued weakness of European banks in 2013 can be traced back to the loans made during the prolonged period of low credit standards in the mid 2000s evident in Figure 1.² The previous sentence is hardly controversial yet for these loans to be still causing problems, at least three elements must hold:

1. banks and their borrowers must have agreed at the time to multi-period loan contracts;
2. a substantial proportion of those borrowers have not exercised any option to repay these loans; and

¹Moore (1956) argues that lax credit standards in the 1920s contributed to the severity of the Great Depression citing the fact that loans originating in the second half of that decade defaulted at twice the rate of those originating in the first half.

²A similar picture emerges from US data.

3. banks themselves must have limited options to revoke the loans and demand early repayment.

Figure 1 also includes two components of the ECB credit standards measure and reveals that during their relaxation of credit standards in the mid-2000s, banks extended the maturity of their loans and weakened their collateral requirements. If we are to understand how and why changes in credit standards have such persistent effects on loan portfolios and eventually understand their macroeconomic implications, then we need a modelling framework which explains loan maturity duration and voluntary restriction on loan recall in the presence of rational borrower behaviour. Moreover, such a model ought to be able to explain why banks change their credit standards over time. For example, can it help us differentiate between competing hypotheses on why banks lowered their credit standards in the mid-2000s: was it the prolonged period of low official interest rates after the bursting of the tech bubble (the "risking-taking" channel of monetary policy (Borio and Zhu (2008)) or was it the "savings glut" (as argued by Bernanke (2010) and King (2010)))? The contribution of this paper is to develop a model with all of these features that can help distinguish between these hypotheses.

This model is very different from those found in the standard microeconomics of banking literature and in particular those analysing credit standards.³ The reasons are interrelated. The most profound difference is that the model is recursive rather than set in two periods with a single shock. This recursive structure opens up the possibility that loans can have an expected duration of any length (between one and infinite) and therefore a way to analyse how loan maturity is determined. However, there are immediately several challenges when loans can be subject to a sequence of shocks. We need to consider what information is known to the borrower and lender after each shock and what actions each can take. If both parties are fully informed after each shock and both parties have a veto over continuation - i.e., the borrower can chose to repay and the lender can revoke the loan - then it is effectively a single period loan whatever the nominal maturity of the contract. And if shocks are independent then knowledge of each shock realisation is irrelevant for the loan continuation decision: either a borrower and bank

³Freixas and Rochet (2008) is the comprehensive reference for the microeconomics of banking. Hellwig (1991) in an early survey divided the literature on credit standards into models of *ex ante* screening (following Broecker (1990)) and *ex post* monitoring (such as Diamond (1984), Gale and Hellwig (1985) and (Williamson (1986))). Screening models have tended to dominate recently (for example Ruckes (2004), Dell'Arricia and Marquez (2006)), Gorton and He (2008)), although there is at least one that combines both effects (Acharya and Naqvi (2012)).

will always agree to a loan or they never will depending on the shock distribution and other primitives. So if loan maturity has real economic content, shocks must be serially correlated and at least one party must be either uninformed or unable to act. It is natural to assume that the borrower is freely informed and always has the option to repay the loan (although possibly at a price) and this will be the case in this model. Banks, however, incur costs to learn about their borrowers and offer contracts with loan covenants which limit the circumstances under which they have the right to revoke the loan. Banks, therefore, are the ones with the informational disadvantage and less freedom to act. But these restrictions are voluntary: a bank could monitor more frequently to be better informed or include tighter covenants in its contracts. Deciding how frequently to be informed about each loan, also referred to as the monitoring intensity, will be a key decision variable in the model.

A second departure from standard models is that loan default is endogenous rather than exogenous because it reflects a moral hazard problem rather than inability to pay. This unwillingness to pay is not, as in the sovereign debt literature, because of a lack of enforceability but because borrowers have the option to declare themselves bankrupt. In most models, bankruptcy is equivalent to the worst possible outcome when the borrower simply cannot pay. However, as painful as bankruptcy may be, it is actually intended to be a second-worst option. Borrowers file for bankruptcy *protection* to stop a bad situation getting worse and to limit their losses. So deciding how bad the situation has to be before taking that step involves a comparison of alternatives. In this model, a borrower will be able to exit from production (and the loan) in an orderly fashion or by declaring bankruptcy. Depending on the shock they receive, the borrower will be choosing the best option out of continuing with the loan, orderly exit or default. Therefore, default is a choice, as is the decision to continue in situations which make the decision to default in the future more likely. Credit standards in this model are the means by which the bank tries to combat this moral hazard. However these controls will be costly because monitoring is expensive and the degree of creditor control influences the interest rate borrowers are willing to pay. Credit control will therefore not be complete and there will be default risk in equilibrium.

A third change of perspective which arises when using a recursive rather than static structure is that the bank is comparing steady state portfolio distributions rather than exogenously given probability distributions (for example the different distributions of "good" and "bad" borrower types in Holmstrom and Tirole

(1997)). These steady state portfolios arise endogenously from the equilibrium behaviour of the borrowers in response to the credit standards set by the bank. These endogenous portfolios have a much richer structure of credit risk than standard models.

A final important difference from the prevailing literature is that the model is general equilibrium rather than partial equilibrium. Standard credit risk models assume that banks can borrow as much as they want at a given interest rate and that the size of the balance sheet is irrelevant. Bank credit standards, therefore, have no effect on aggregate outcomes nor are they constrained by macroeconomic circumstances. By contrast, in this model, there is a monopoly bank that needs to have the same measure of deposits as loans. So the credit standards and loan interest rates offered by the bank will have to deliver a steady state distribution which satisfies this balance sheet constraint. Intuitively, the terms available to borrowers do not just affect borrowers themselves but also the incentives to be savers. This general equilibrium perspective is crucial in distinguishing between the risk taking channel of monetary policy and the savings glut hypothesis.

Evidently the model does not build on the theoretical foundations of the existing banking literature. Instead it adapts the firm dynamics model of Hopenhayn (1992) to incorporate features of banking and in particular credit risk control. (It shares this approach with the recent model of Arellano, Bai and Zhang (2012)). With this framework, the balance sheet of the bank is a counterpart of the distribution of firms in the economy. Firm (and thus loan) dynamics are driven by two effects. There is a natural component because of persistent idiosyncratic shocks to firm profitability. But the terms of the loan contract and the loan interest rate charged also influence the decisions to enter and to exit production, and therefore the shape and size of the steady state distribution of firms and the credit risk facing the bank. The objective of the bank is to find the profit maximising credit standards and loan interest rate subject to the balance sheet constraint.

The paper is organised into seven sections as follows. Section 2 sets up the model by describing the state space, the heterogeneous agents and the banking contract. Section 3 explains the equilibrium behaviour of the agents for a given loan contract. To assist in understanding the subsequent choice of the bank, an illustrative numerical example is presented. Section 4 explains equilibrium bank behaviour and the profit-maximising choice of credit standards. Section 5 describes the general equilibrium properties of the model. Section 6 uses the model to consider whether the global savings glut or the risk-taking channel of monetary

policy are possible culprits for explaining the loosening of credit standards in the lead up to the crisis. The results show that in this model only the savings glut hypothesis is consistent with lower credit standards and falling credit spreads. Section 7 concludes.

2. Model

2.1. Space

The model takes place in an economy comprising a measure 1 of infinitely small, *ex ante* identical and infinitely living risk-neutral agents and a bank. Time is discrete and future payoffs are discounted at rate β . In each period every agent chooses whether or not to run a "project" (or "firm"). Existing or potential projects are agent specific and are indexed by an idiosyncratic "profitability" state, a , which can be thought of as including technical productivity, consumer preferences, degree of market power and managerial talent. a is drawn stochastically every period from the compact set $\{A \in \mathbb{R} : 0 \leq a \leq 1\}$ and completely defines the circumstances of each agent (and thus *ex post* heterogeneity). Profitability levels are private information and can only be verified by the bank by paying a monitoring cost m per loan. All agents are endowed with a unit of capital which they cannot add to or lose. The purpose of this assumption will be explained once more of the structure of the model is in place.

2.2. Agents and production

At any point in time the agents in this economy are partitioned endogenously into two situations.

One group of agents is considering whether to enter production. During each period these "inventors" receive an "idea" for a project with a profitability state drawn from the continuous i.i.d distribution $G(a)$.⁴ Inventors weigh up paying start-up costs S to enter production next period based on this profitability draw or waiting for a better idea in the future.⁵ Inventors are of measure I (which in

⁴This is similar to the way wage offers are received in labour search models such as McCall (1970).

⁵Start up costs are large in many countries. In Djankov *et al* (2002) just the official costs of entry range from 0.5% of per capita GDP in the US to 460% of per capita GDP in the Dominican Republic.

equilibrium will equal a $\frac{1}{2}$) and while they wait, they deposit their capital at the bank on which they receive an exogenously fixed deposit rate, r_d .

The other group of agents is currently in production for which they will have borrowed capital from the bank at an *endogenously* determined loan interest rate, r . These agents are called "entrepreneurs" and their per period payoffs ("profits" and "losses") are a function of their idiosyncratic profitability states $q(a)$ and the loan interest rate r . These profitability states evolve according to a first-order Markov process $F(a', a)$. The following assumptions on the transition process and profit function are made:

- A** (i) $F(a', a)$ is continuous in a and a' ; (ii) Profitability shocks are persistent and so $F(a', a)$ is strictly decreasing in a . (iii) But profitability shocks eventually die out and the monotone mixing condition is satisfied: $F^n(\epsilon, a) > 0 \forall \epsilon$ for some n where $F^n(\epsilon, a)$ is the conditional probability distribution of profitability in n periods time given a . So from any given level of profitability, it is possible to transit to any other profitability level in a finite number of periods. Since there are exit thresholds, this assumption implies that all projects will almost surely close at some future point.
- B** (i) $q(a)$ is continuous and; (ii) strictly increasing in a .

In each period, entrepreneurs decide whether to continue in production next period or to exit. If they do not wish to continue, there are two exit options.

- "Orderly" exit occurs if the entrepreneur absorbs current period payoffs - naturally losses - and pays a liquidation cost L to close the project. These liquidation costs might be pecuniary such as termination pay, liquidating stock at below cost and administrative costs or non-pecuniary such as lost human capital and reputation.⁶
- "Default" occurs if the entrepreneur pays an exogenous cost B to file for bankruptcy protection in which case current period losses are excused (including repayment of loan interest). If the entrepreneur defaults, the bank has to pay the liquidation costs instead. (For notational convenience it is assumed that S , L and B are paid in the following period.)

⁶Ramey and Shapiro (2001) describe the heavy discounts on machinery sold during the closure of aircraft manufacturing plants.

These options for the entrepreneur will partition the set A into three regions - continuation, orderly exit and default - delineated by threshold values a_X and a_δ (exit and default, respectively) and the value of B is calibrated for the sensible ordering $a_X > a_\delta$. If the model had a continuous time shock process, then agents experiencing negative profitability shocks would always enter the voluntary exit region first and default would never arise. In discrete time, there is default because some projects jump over the orderly exit region $[a_\delta, a_X]$ (which could be considered as a proxy for discontinuous shocks). By assumption A(iii), project profitability will almost surely pass below one of these thresholds and the project is terminated on the first occasion. When entrepreneurs exit, they become inventors again so the value of an entrepreneur's outside option is the expected value of being an inventor. Defaulting entrepreneurs are indistinguishable from other inventors.⁷

To keep the model as simple as possible and to focus on the main mechanism of interest, all decisions take place at the extensive margin. Aggregate variables, which are used for market clearing, are thus the sums over measures of agents entering, exiting or continuing. Likewise, expectations of payoffs are simple integrals over profitability states only. It was to turn off the intensive margin that agents were assumed to be unable to change their capital holdings.⁸ Relatedly, it will be assumed that projects require 2 units of capital. Entrepreneurs do not borrow directly from inventors (perhaps for the reasons described by Diamond (1984)) and must get a loan of 1 unit from the bank.⁹ These simplifications ensure that the only decisions that agents are making are how to allocate their units of capital between bank deposits and production, that borrowing and lending take place and that the loan size remains fixed over the lifetime of the project. More technically, it means that the only endogenous distribution involved in the model is the measure of entrepreneurs over the set of profitability states.

Finally, since this is a model based on incentives, it will be (implicitly) assumed that all agents have an exogenous endowment of income every period regardless of their circumstances sufficient to cover any expenses or losses. This simply rules out having to consider situations in which agents want to pay but cannot. Decisions

⁷It is just convenient to recycle defaulters in this way. Nothing of any substance would change by assuming defaulters are excluded forever but new inventors are born at the same steady-state rate.

⁸For a similar model with entrepreneurial wealth accumulation which is used to explain firm size dynamics, see Arellano et al (2012).

⁹This implied leverage ratio could be set to any constant without loss of generality with an appropriate adjustment to the market clearing condition.

are determined by the profitability state alone.

2.3. Banking

Capital in this model is intermediated between inventors and entrepreneurs by a monopoly bank. As a consequence of the assumption of infinitely small agents and the law of large numbers, *ex ante* probability distributions over the bank's loan portfolio are identical to the distribution of *ex post* outcomes. Since this is a model of purely idiosyncratic risk, the distributions faced by the bank are deterministic.

The bank offers the following deposit and loan contracts:

- Deposits earn an interest rate r_d set exogenously by the monetary authority. Deposits can be withdrawn at the end of any period.
- The loan contract specifies an interest rate r , a monitoring intensity φ (where $0 \leq \varphi \leq 1$) and a covenant specifying a minimum profitability level ξ . For notational simplicity, the parameters of the bank contract are summarised by $\psi = \{r, \varphi, \xi\}$. As set out in the introduction, the loan is nominally indefinite but both parties have an option to terminate it each period. The borrower has the option to repay the loan if she decides to exit production. And the bank can demand repayment if it discovers that the covenant condition has been breached. The covenant condition, ξ , will correspond to a threshold value, a_T , at which the bank exercises its right to terminate the loan. Since the bank uses the covenant to protect its interests, it follows that ξ must imply a trigger value of a at least as high as that at which entrepreneurs voluntarily exit or else the covenant would be redundant.

3. Equilibrium behaviour of the agents

This is a recursive model and in each period the move order is the following:

1. Agents enter the period in their previously chosen situation (inventor or entrepreneur) and then draw their idiosyncratic shocks. The inventors get a new idea from $G(a)$ and entrepreneurs get an update of their profitability according to $F(a', a)$.

2. Entrepreneurs decide whether to continue with production next period or to exit either voluntarily or by defaulting. Payoffs are received and loan interest paid by non-defaulting entrepreneurs. Entrepreneurs who exit voluntarily inform the bank that they will repay their loan. Inventors receive deposit interest and decide whether to enter production next period based on their profitability draw.
3. The bank monitors continuing loans at the stochastic rate φ and recalls the loans of all entrepreneurs found below the covenant profitability threshold state a_T .
4. The bank receives deposits from waiting and new inventors and makes additional loans to entering producers.

The following two sections formalise the analysis of the choices of the productive and inventors.

3.1. Entrepreneurs

Depending on the idiosyncratic profitability state, a , an entrepreneur chooses between default, orderly exit and continuation. If she decides to default to escape $q(a) - r$, she pays B next period and switches to being an inventor. The discounted value of defaulting is thus

$$V_B = \beta\{E[V_I(a'; \cdot)] - B\} \quad (1)$$

where the value function of an inventor is denoted $V_I(a; V_E)$ and $E[V_I(a'; \cdot)] = \int_A V_I(a'; \cdot) G(da')$.

If she chooses orderly exit from production, she absorbs current losses, pays liquidation costs L next period and also enters next period as an inventor. The value of orderly exit in state a is

$$V_X(a) = q(a) - r + \beta\{E[V_I(a'; \cdot)] - L\} \quad (2)$$

The remaining option is to continue in production next period. Naturally the conditional value of continuing in production is to receive current payoffs and the discounted expected value of being an entrepreneur in the next period.

$$V_C(a) = q(a) - r + \beta\{E[V_E(a'; \psi, V_I) | a]\} \quad (3)$$

where the value function of an entrepreneur is denoted $V_E(a; \psi, V_I)$.

Given the options available, the value of being an entrepreneur at the moment the shock is revealed is:

$$V_E(a; \psi, V_I) = \max \{V_B, V_X(a), V_C(a)\} \quad (4)$$

There is a natural ordering of the choices facing an entrepreneur. Bankruptcy costs will be assumed to be sufficiently large that entrepreneurs only choose this form of exit when facing a very bad profitability state. So the default threshold will be determined by a comparison of the value of default and the value of orderly exit. It is straightforward to see from equations (1) and (2) that entrepreneurs will default for all values of $a < a_\delta$ where

$$q(a_\delta) = \beta(L - B) + r \quad (5)$$

The threshold for orderly exit, a_X , results from the comparison of $V_X(a)$ and $V_C(a)$. The only challenging aspect of this problem is the conditional expected value of being an entrepreneur next period at the time the decision is made: $E[V_E(a'; \cdot) | a]$. Consider first an entrepreneur with profitability above the loan covenant threshold, $a \geq a_T$. In this case the entrepreneur faces no risk if the bank chooses to monitor, so she can ignore the monitoring probability of the bank and

$$E[V_E(a'; \cdot) | a \geq a_T] = \int_A V_E(a'; \cdot) F(da', a)$$

The calculation is more complex for an entrepreneur with $a_\delta < a < a_T$. In this case, if the entrepreneur decides to continue and escapes monitoring (with probability $1 - \varphi$) then the entrepreneur gets the conditional expected value of being an entrepreneur in the next period. If the entrepreneur tries to continue but is monitored then the loan is recalled by the bank, the project is shut down and the agent involuntarily reverts to being an inventor. Therefore for $a_\delta \leq a < a_T$

$$E[V_E(a'; \cdot) | a_\delta \leq a < a_T] = (1 - \varphi) \int_A V_E(a'; \cdot) F(da', a) + \varphi \left(\int_A V_I(a'; \cdot) G(da') - L \right)$$

The voluntary exit threshold a_X is the value of current period profitability at

which an entrepreneur is indifferent between continuing or exiting voluntarily.¹⁰ Some simple cancelling gives

$$\int_A V_E(a'; \cdot) F(da', a_X) = \int_A V_I(a'; \cdot) G(da') - L \quad (6)$$

3.2. Inventors

We turn now to the decision by inventors whether or not to enter production. Unlike in Hopenhayn (1992), inventors are assumed to draw a profitability level *before* they decide whether or not to enter although they cannot begin production until the *following* period. Each period, an inventor gets one idea with a profitability level drawn from $G(a)$. The agent can either decide to pay the cost of starting up a project, S , and enter production next period or keep his capital on deposit at rate r_d for another period. Profitability next period will be subject to an idiosyncratic shock according to the same function F as existing projects. So the expected value of entering production is equal to the expected value of being an existing entrepreneur at the same level of profitability net of S . An inventor this period receives the interest on his deposit for this period and the discounted expected value of the maximum of the choice between entering or remaining as an inventor the following period. The value function for an inventor with an initial draw of a is therefore:

$$V_I(a; V_E) = r_d + \beta \max \left\{ \int_A V_E(a'; \cdot) F(da', a) - S, \int_A V_I(a'; \cdot) G(da') \right\} \quad (7)$$

The threshold level of profitability at which inventors will choose to enter is where the expected value of being an entrepreneur at that level of profitability net of start up costs matches the expected value of waiting. a_E is thus determined by:

$$\int_A V_E(a'; \cdot) F(da', a_E) - S = \int_A V_I(a'; \cdot) G(da') \quad (8)$$

¹⁰Since the bank only monitors continuing loans, these are always above the default threshold and so any agent forced to repay the loan will prefer to exit in an orderly fashion rather than default.

The right hand side is not contingent on the current state and in equilibrium will be a known constant.

The equilibrium behaviour of entrepreneurs and inventors is described in the following proposition:

Proposition 1. *Given Assumptions A and B and a banking contract ψ , unique, bounded and mutually consistent functions $V_E(a; \psi, V_I)$ and $V_I(a; \psi, V_E)$ exist. These value functions yield unique and continuous functions in ψ for a_E and a_X . $a_E(\psi)$ and $a_X(\psi)$ are strictly increasing in r .*

Proof. The proofs of all propositions are contained in the Appendix. ■

Proposition 1 states that there are unique values of the entry and exit thresholds, a_E, a_δ and a_X which, along with the profitability threshold implied by the loan covenant, a_T , completely summarise the equilibrium behaviour of agents for a given loan contract ψ and deposit rate r_d . So from here on we can dispense with the value functions.

3.3. Steady state

Define $H([0, a]; \psi)$ as the measure of entrepreneurs at the *end* of each period with profitability levels in the interval $[0, a)$ given loan contract ψ . Define $FH(a'; \psi) = \int_A F(a', a)H(da; \psi)$ which is the distribution of firms evolving into productivity state a' from the distribution $H(A; \psi)$.

With the behavioural assumptions of the model and recalling that I is the measure of inventors, then the transition function for the distribution of entrepreneurs is:

$$H'([0, a']; \psi) = I \int_{a_E}^{a'} G(a) + \int_{a_X}^{a'} FH(da'; \psi) - \varphi \int_{a_X}^{\min(a_T, a')} FH(da'; \psi) \quad (9)$$

The first term on the right hand side measures how many agents enter at profitability levels below a' . The middle term measures continuing entrepreneurs evolving into the profitability interval $[a_X, a')$ - ie above the voluntary exit threshold, a_X . The third term eliminates those entrepreneurs in the interval $[a_X, a_T)$ forced to close down because the loan is recalled by the bank. (Defaulting entrepreneurs

are implicitly removed by the lower truncation of the distribution at a_X .) An invariant steady state distribution occurs if

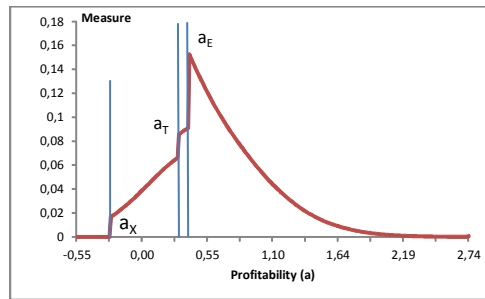
$$H'([0, a]; \psi) = H([0, a]; \psi) = \bar{H}([0, a]; \psi) \quad \forall a \in A$$

Proposition 2. *For each ψ there is a unique invariant distribution, $\bar{H}([0, a]; \psi)$.*

3.4. Illustrative numerical example

Before turning to the decision of the bank, it is useful to illustrate an example of an invariant distribution derived from the model and decompose the transition equation (9). The parameterisation in Figure 2 is entirely illustrative but was chosen to give roughly sensible credit spreads and entry, exit and default rates.¹¹ The monitoring intensity, which will be endogenised later, is here assumed to be strictly positive but less than 1.

Figure 2: Steady-state distribution of firms on the bank's balance sheet



It is easy to see the influence of the three behavioural thresholds on the distribution. Below a_X , there are no entrepreneurs in the distribution at the end of each period because they have either defaulted or exited voluntarily. Between a_X and a_T we have entrepreneurs that are in breach of the loan covenant but

¹¹The example uses a normal distribution for $G(a)$ and an AR(1) process for $F(a, a')$ using a Tauchen matrix approximation (see Tauchen (1986)).

have escaped monitoring. There is a concentration of entrepreneurs just above the entry threshold, a_E .

There is a long right tail to this distribution. These are entrepreneurs who have either entered with a very high initial profitability level or entered and subsequently experienced predominantly positive profitability shocks. Those at the far right are well beyond the initial profitability draws so only exist because of the persistence of shocks and the luck of drawing positive shocks.

The model structure and Figure 2 are broadly consistent with the empirical evidence. Studies using US data, for example Bartelsman and Doms (2000), Baily, Bartelsman and Haltiwanger (2001) and Foster, Haltiwanger and Krizan (2006)), show that there are wide distributions of profitability and productivity within industry classifications and that firm-level shocks are highly persistent. Fariñas and Ruano (2005) use Spanish manufacturing data and show that the productivity distribution of exiting firms is stochastically dominated by the distribution of continuing firms and that the productivity of entering firms is stochastically dominated by continuing firms.

Figure 2 also helps draw out the importance of start-up and liquidation costs in driving the results of the model. If we combine equations (6) and (8) which determine the entry and exit thresholds then we obtain

$$\int_A V_E(at, r, V_I)F(dat, a_X) = \int_A V_E(a', r, V_I)F(dat, a_E) - S - L \quad (10)$$

from which it can be easily seen that if $S = L = 0$, then $a_E = a_X$. Since a covenant threshold is only relevant in the interval $[a_X, a_E]$, if we have $S = L = 0$ and $a_E = a_X$, then a_T is redundant and so is bank monitoring. In this case, even though there is a positive probability that profitability falls below the default threshold, a_δ , there is no incentive for bank monitoring. This occurs because in the frictionless entry and exit case, borrower behaviour is completely aligned with the interests of the bank. Borrowers only continue in situations in which they would also wish to enter. Frictionless entry and exit always selects the most profitable firms given the profitability processes and thus the lowest possible credit risk. So asymmetry of information has no bite when exit and entry is costless.¹² This ability to rely on borrower behaviour breaks down when there are entry and exit costs because private choice by borrowers no longer selects the lowest credit

¹²With a single F process and bankruptcy costs, the bank does not need to engage in *ex ante* screening or *ex post* monitoring.

risk portfolio. By setting a covenant threshold and monitoring stochastically, the bank can alter the distribution of credit risk.

Figure 3: One period transition of bank balance sheet

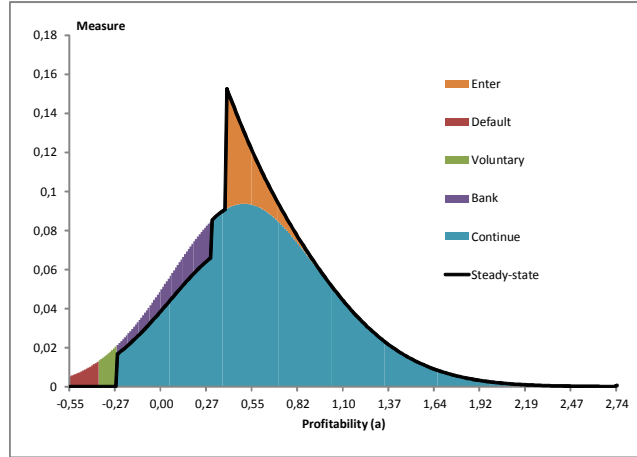


Figure 3 illustrates the one period transition of the distribution in Figure 2 with the invariant distribution overlaid. Looking from right to left, one can see that the upper tail of the distribution is entirely driven by the presence of a small number of existing entrepreneurs experiencing positive shocks. Since on average entrepreneurs with positive profitability experience a reversion towards the mean (of zero), there is a noticeable deterioration in the average quality of existing entrepreneurs - the distribution melts to the left. The distribution is refreshed by the entry of new entrepreneurs clustered above the entry threshold. Moving further to the left, a number of entrepreneurs fall below the threshold a_T but above a_X . These are the entrepreneurs who want to continue but are in breach of the loan covenant. φ proportion of these entrepreneurs are monitored, have their loans revoked and exit and $1 - \varphi$ are able to continue. Moving further to the left, there are entrepreneurs who fall below a_X but above a_δ and exit voluntarily. Finally, there is a portion of the distribution who falls below a_δ and defaults.

4. Equilibrium bank behaviour

We can now turn to the bank's choices of parameters in the loan contract $\psi = \{r, \varphi, \xi\}$. In this framework, the effects of φ and ξ are almost identical. The bank can protect itself against default risk by raising the monitoring rate (equivalent to a cut in the expected duration or maturity of the loan) or tightening the covenant. But it is simpler to present the equilibrium by fixing the value of one parameter and making the other an endogenous choice variable. In what follows, the covenant value is fixed and the monitoring rate is endogenous but making the alternative choice changes nothing substantive about the results. This leaves pairs of r and φ . Proposition 2 asserted that there is a unique invariant distribution for any loan contract and thus (r, φ) pair. The bank, however, is constrained in its choice of loan contract by the need to finance its loans by deposits.¹³ Using the simplifying assumption made earlier that all agents have a fixed unit of capital but projects require 2 units, it follows that there must be as many borrowers as depositors. With measure 1 of agents, the funding constraint faced by banks in equilibrium is:

$$\bar{H}(A; r, \varphi) = \frac{1}{2} \quad (11)$$

Although choosing an optimal (r, φ) pair is a joint decision, for ease of explanation (and proof) it will be assumed that the bank uses the loan rate to equilibrate its balance sheet and then uses the monitoring rate to maximise profits.

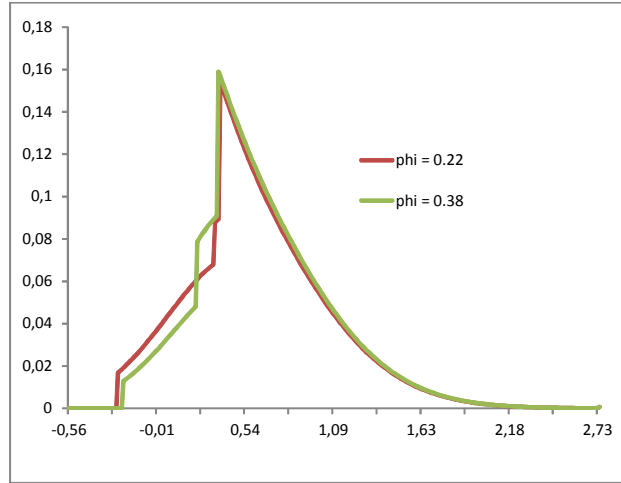
Proposition 3. *There is a unique value \tilde{r} that ensures that the balance sheet of the bank is equal on both sides for given values of φ and ξ .*

This is a very intuitive proposition. If the bank faces an excess demand for loans, then raising the borrowing rate simultaneously reduces the demand for new loans (by increasing a_E), increases the incentive for existing borrowers to repay and exit production voluntarily (an increase in a_X) and effectively tightens the loan covenants (by increasing a_T).¹⁴ These effects work on both sides of the balance sheet by reducing loans and increasing deposits. Uniqueness follows from continuity and monotonicity of the behavioural functions. With r a function of φ , we can denote the subset of invariant balance sheets, \bar{H} , which satisfy the balance sheet constraint as $\tilde{H}(A, r(\varphi), \varphi)$.

¹³The level of equity funding is not relevant in this model.

¹⁴It also increases the default rate by increasing a_δ . Since those going bankrupt are assumed to reappear as depositors in the next period, within the logic of the model, this also reduces the excess demand for loans.

Figure 4: Different Steady-state Distributions



Of course, not all invariant distributions that satisfy the balance sheet constraint are equivalent. To illustrate what is at stake, Figure 4 compares the distributions arising from two different pairs of r and φ which satisfy the balance sheet constraint. When the monitoring rate is lower ($\varphi = 0.22$), it is intuitive that there are more projects in the left tail of the distribution because more borrowers are able to continue in breach of the covenant than otherwise. Since there is a higher potential survival rate, this also increases the interval over which it is rational for the borrower to gamble for resurrection - a_X is lower. But more lenient contract terms are more favourable for the borrower so satisfying the bank balance sheet constraint requires a higher loan rate, r . The higher loan rate explains the other differences in the distribution. With a higher loan rate, the covenant threshold bites at a higher value of a (a_T is further to the right for $\varphi = 0.22$). A higher loan rate is also a disincentive to enter, so the entry threshold a_E is also higher, with the knock-on effect that there are marginally fewer projects with high profitability. Overall, the invariant profitability distribution for a higher monitoring rate stochastically dominates a distribution with a lower monitoring rate and has a lower default rate. In this sense, a lower monitoring rate results in a loan portfolio of lower "credit quality".

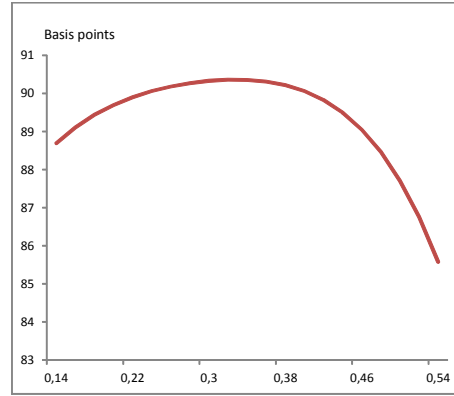
Another way to think about the effect of an increase in φ is to consider how this affects the way entrepreneurs exit. Given the process for $F(a', a)$, entrepreneurs almost surely find themselves in the interval $[a_X, a_T]$ at some point and thus are vulnerable to having their loan recalled by the bank. Since this is above the voluntary threshold, there is a utility loss for entrepreneurs ejected in this way. From the perspective of the individual agent, the recall of the loan results in premature and inefficient liquidation of the project. The parameter φ is thus effectively a distribution of control rights over the decision to exit production - the lower φ is, the higher the control rights allocated to the entrepreneur. Anticipating this, an entrepreneur will be willing to pay more for a loan with a lower monitoring rate because it gives the entrepreneur higher control rights over the exit decision and reduces the risk of premature liquidation. If earnings are volatile, this is potentially a very important consideration for a borrower. An entrepreneur required to reveal her profitability state to the bank every period will not have her loan rolled over and will be forced to pay the liquidation costs as soon as she fails to meet the covenant condition. A less monitored loan is thus a form of insurance against profit volatility and premature liquidation which an entrepreneur is willing to pay through a higher average interest rate.

Having clarified these issues about the influence of r and φ , we can now state the decision problem of the bank more formally. Again to simplify notation, define $F\tilde{H}(a'; \varphi) = \int_A F(a', a)\tilde{H}(da; r(\varphi), \varphi)$ which is the distribution of firms evolving into state a' from the balance-sheet-consistent invariant distribution at the end of the previous period for a given value of φ . ($F\tilde{H}(a'; \varphi)$ is the distribution depicted in Figure 3.) The bank's objective is to:

$$\begin{aligned} \max_{\{\varphi\}} \Pi &= r(\varphi) \int_{a_\delta}^1 F\tilde{H}(da'; \varphi) - r a \frac{1}{2} \\ &\quad - \int_0^{a_\delta} \lambda(a') F\tilde{H}(da'; \varphi) - \varphi m \int_{a_X}^1 F\tilde{H}(da'; \varphi) \end{aligned} \quad (12)$$

where m is the per unit cost of monitoring and $\lambda(a)$ is a parameter measuring loss given default which is decreasing in a . Equation (12) measures steady state bank profits. Since the bank is assumed to be sufficiently large that a law of large numbers applies, the variables in this problem are completely deterministic. The first term in equation (12) measures the interest income received on non-defaulting loans. The second term deducts the payment of interest on all deposits. The third term measures expected loss given default. The final term measures the cost of

Figure 5: Profit per loan



monitoring those entrepreneurs who choose to continue. Since different values of the monitoring intensity imply different invariant distributions, steady state bank profits vary in φ .

Figure 5 illustrates the general case of an interior solution to the model based on a constant per unit cost of monitoring m and loss given default $\lambda(a) = L - 2q(a)$. Discussion of the importance of the latter assumption is deferred to the following section on the general equilibrium properties of the model. The crucial point is that the bank is less efficient than the entrepreneur in running the project until it can be closed down.¹⁵ In this case, credit quality improves as monitoring increases but at a decreasing rate. This decreasing marginal effectiveness of monitoring occurs because as the credit quality of the portfolio improves, there are fewer and fewer loans at risk of jumping to default and so monitoring is influencing a smaller proportion of borrowers. The equilibrium loan interest rate is falling because more highly monitored loans are less attractive to borrowers. And, of course, monitoring is costly. For an interior solution, the reduction in loss given default dominates initially but is eventually outweighed by the reduction in monitoring effectiveness, the higher monitoring costs and the lower equilibrium credit spread.

¹⁵ $q(a)$ is negative (ie gross losses) whenever default takes places.

5. General equilibrium properties

This section discusses the general equilibrium properties of the model and intuition about how the various features of the model interact. The first step is to describe the social surplus in the model and how it is distributed between the bank and the agents. The second step is to consider the circumstances in which there is either full monitoring or no monitoring. The final step is to examine how the equilibrium relates to social welfare.

Gross surplus in the model is the integral sum of the measure of firms at each level of profitability (or "output") and therefore depends on the equilibrium distribution of firms. For the subsequent discussion it is useful to consider the possibility that there are transfers into this otherwise closed economy from outside which can augment this surplus. From the combined surplus, we need to deduct monitoring costs, entry and exit costs and the excess loss given default arising from the bank liquidating the project.¹⁶ The entry and exit costs depend on the equilibrium turnover rate of firms. Since firms that enter produce more gross surplus than those that exit, a higher turnover rate of firms increases gross surplus. A more profitable distribution also reduces default risk, *ceteris paribus*, and thus excess loss given default. But clearly a higher turnover rate implies higher entry and exit costs. The increase in turnover costs is linear in the turnover rate but the increase in surplus occurs at a diminishing rate because it has a decreasing effect on default risk.

The resulting net surplus is distributed between the inventors, entrepreneurs and the bank.

- The share going to inventors is determined by the deposit rate and the *exit* rate. The deposit rate is set exogenously so this component is independent of the endogenous variables and the equilibrium distribution. The payoffs are common across inventors apart from the exit costs paid by those who have just terminated a project.
- Entrepreneurs get their idiosyncratic profits and losses less the loan interest rate and start-up costs. Defaulting entrepreneurs also pass losses and liquidation costs onto the bank.

¹⁶This assumes that entry and exit costs are deadweight social losses rather than transfers between agents.

- The bank receives the loan spread net of monitoring costs and loss given default. The bank share is not zero in general because it receives a rent due to the assumption that inventors cannot lend directly to entrepreneurs and the latter have higher surplus than the former. This surplus differential occurs because only the marginal entrant is indifferent between the expected discounted sum of payoffs net of start-up costs and expected discounted exit costs. Infra-marginal entrants clearly get more. Inventors want access to this surplus and the bank gets the benefit from the need to ration credit so that the balance sheet condition is satisfied.

The previous section illustrated the general case in which there was an interior solution for the choice of monitoring intensity. But the framework nests the two corner solutions: complete monitoring and no monitoring.¹⁷ Complete monitoring can be the optimal choice through a combination of low monitoring costs, low bankruptcy costs and a high inefficiency cost from the bank closing the firm. Take each of these in turn. If there was no inefficiency in firm default so that the bank was just as efficient at running the project as it is liquidated, then the gain to the borrower from the option to default would exactly match the cost to the bank. Bankruptcy with no inefficiency is simply an option to transfer losses and would be compensated exactly in the equilibrium interest rate. And with a transfer at a fair price (and no risk aversion), neither party has an interest in reducing default risk but the bank has an interest in reducing the monitoring cost to zero by not monitoring. But as soon as there is an inefficiency, the valuation of the default option differs between the two sides: the cost to the bank is more than the benefit to the borrower. As a result, if the bank starts to monitor, the reduction in loss given default is more than the reduction in the loan interest rate the borrower is willing to pay and thus is a profitable action (at least initially). Now turn to the effect of bankruptcy costs. With inefficient default, the bank is exposed to an uncompensated risk from firms going bankrupt. The extent of this risk depends on the incentives for the borrowers to risk bankruptcy which in turn depends on the bankruptcy costs. If these are low (and particularly if they are not much more than the liquidation costs), then the incentive for the borrower to continue is higher. It follows that the higher this moral hazard risk, the more monitoring the bank should undertake. And, finally, how much monitoring the

¹⁷No monitoring but with bankruptcy costs is effectively the same equilibrium as in Diamond (84) and Gale and Hellwig (85).

bank will do will be directly related to the costs of monitoring. All other things equal, a low cost of monitoring increases the monitoring rate. Taking all these points together: if monitoring is cheap, moral hazard risk is high and loss given default is large, then the bank will monitor intensively, possibly fully. Naturally, low or no monitoring will arise in the opposite cases.

An interesting additional example where no (or low) monitoring might arise is if the bank expects to get bailed out in the event of default losses. This is the possibility of external transfers mentioned above. The option to default is valuable to the borrower and, as previously discussed, they are willing to pay a higher loan interest rate for a low rate of monitoring. If the bank can effectively shift some of these default losses onto the taxpayer, then the credit spread increases by more than the bank pays in loss given default. By setting the monitoring rate to zero, the bank simultaneously maximises the implicit subsidy on default risk provided by taxpayers and minimises monitoring costs.

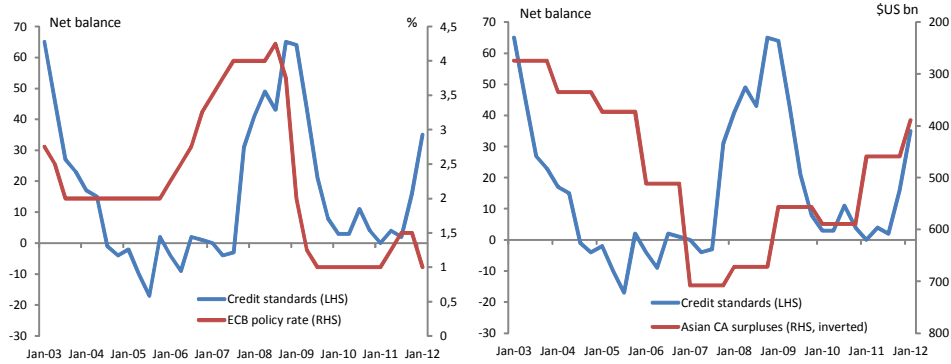
A final, interesting, issue is a comparison between the social welfare maximising level of monitoring and the one chosen by the bank. A natural social welfare benchmark is the maximum net surplus. In the discussion on the origins of the social surplus it was noted that higher monitoring increases the turnover rate of firms and therefore gross surplus is increasing in the monitoring rate. Credit risk is also strictly decreasing in the monitoring rate. Against this, a higher monitoring rate involves higher turnover costs plus the monitoring costs themselves. Just as in Figure 5 there will be a hump-shaped function for net social surplus and a welfare maximising level of monitoring. As might be expected, for interior solutions of the model the socially optimal rate of monitoring is higher than the bank's profit-maximising choice. The bank pays the monitoring costs but part of the increase in social surplus goes to the entrepreneurs. The bank obviously does not take this monitoring externality into account when it sets its monitoring intensity.

6. Monetary policy v the savings glut

It is now time to put the model to work on a practical question.

Since the start of the financial crisis there has been a lively debate about whether the low credit standards and narrow credit spreads observed in the years

Figure 6: Risk-taking channel or savings glut?



before it erupted were in part caused by accommodative monetary policy. On one side are those who charge monetary policy makers with having ignored or misunderstood the effects of a prolonged period of low official interest rates on the willingness of banks to take financial risks (Borio and Zhu (2008), Borio and Disyatat (2011), Taylor (2009), Adrian and Shin (2008b), Gambacorta (2009)). The other side, including major monetary policy makers, counter that credit standards deteriorated because of the strong inflows into the western banking system coming from a sharp increase in *ex ante* savings in emerging market countries (Bernanke (2010), King (2010), Portes (2009)). The two hypotheses are illustrated in Figure 6.

There is conflicting evidence on the existence of the risk-taking channel. Jiménez, Ongena, Peydró and Saurina (2008) use data from the Spanish credit registry over the period 1984-2006 and found a statistically significant increase in the credit riskiness of *new* loans (measured by the duration to default of individual loans) when policy rates were low at the time of loan origination. Moreover, if interest rates subsequently rose, then the hazard rate on these riskier loans was materially higher.¹⁸ Ioannidou, Ongena, and Peydró (2008) report similar results using data

¹⁸If the interest rate at origination was 4.13% and the *ex post* interest rate was 4.09%, then the estimated annualised hazard rate is 0.56%. If the loan rate at origination was the lowest in the sample (2.16%) and the *ex post* interest rate the highest (9.62%), then the hazard rate would be around 6 times higher at 3.38%.

from the Bolivian credit register. They also report evidence that borrowers with poor credit histories were more likely to obtain loans when policy interest rates were low.¹⁹ Adrian and Shin (2008a) and Adrian, Moench and Shin (2010) examine the interaction between monetary policy, the size of the balance sheets of leveraged financial institutions, credit risk premia, asset prices and macroeconomic activity. These two papers argue that during the pre-crisis period, non-bank financial intermediaries were the marginal price setter in many risky asset markets. Even commercial banks, which could rely largely on deposits, would borrow on wholesale markets to add to their lending capacity and were thus affected by the market price of risk. A crucial finding of their empirical analysis is that financial sector leverage was highly sensitive to short-term official interest rates.

On the other side of the argument, Lown and Morgan (2008) found no evidence in the US that credit standards were affected by the policy rate. And Dell’Ariccia, Igan, Laeven and Tong (2012) show that the length of the boom and the stance of macroprudential policies rather than monetary policy explain whether a credit boom is followed by a crisis.

The model in this paper differs from recent theoretical models of the risk-taking channel of monetary policy because banks are not the underlying source of friction but exist as an incomplete solution to frictions originating elsewhere. For example, banks are not fundamentally less risk averse than society as in Agur and Demertzis (2012) or value projects in a different way from other investors as in Adrian and Shin (2008b). Indeed, banks are socially useful since they supply maturity transformation and credit risk management. Credit risk arises out of the behaviour of borrowers and exists in equilibrium because it is too costly for banks to eliminate it. The only distortion caused by the bank is the monitoring externality which will result in the bank choosing a socially sub-optimal level of credit standards.

The model developed in this paper can be used to compare the effect of two shocks on equilibrium credit standards and credit spreads.

- The risk-taking channel is introduced through a reduction in the deposit rate r_d . Monetary policy is assumed to directly affect the funding costs of banks and we look for the endogenous equilibrium responses of the loan interest rate and the monitoring intensity.

¹⁹See also Altunbas, Gambacorta and Marques-Ibanez (2010) and Maddaloni and Peydró (2010).

- The savings glut hypothesis is introduced as an exogenous supply of deposits from outside the economy. This alters the balance sheet constraint in equation 11. The deposit rate is kept unchanged and again we look for the endogenous equilibrium responses of the loan interest rate and the monitoring intensity.

The test of these hypotheses is whether these shocks deliver a reduction in monitoring intensity (or equivalently an extension in loan maturity or a weakening in covenants) and a reduction in the spread between the loan and the deposit rate for a given degree of risk. The equilibria being compared in each case are the associated invariant distributions so the implicit assumption is that these shocks are permanent.²⁰ It is also being assumed that we are comparing interior solutions of the model. The results presented below are general and not parameter-specific.

Figure 7 illustrates the effect of the deposit shock. The left panel shows different profit levels for each value of the monitoring intensity. The baseline is the same as depicted in Figure 5. The right panel shows the equilibrium loan spread for each value of the monitoring intensity. The profit-maximising monitoring intensity is clearly to the left of the baseline case and the loan spread is lower for any given monitoring rate.²¹ The intuition for these results is that following a positive deposits shock the bank has to reduce credit terms to induce more loan demand. Clearly one option is to cut only the loan interest rate and leave the monitoring intensity unchanged. But cutting the loan interest rate is the most costly way (in terms of profits) to induce more demand. The profit maximising response is to preserve some of the credit spread by weakening the monitoring intensity. It can do this partly because the lower interest rate reduces default risk for any given distribution of firms.²²

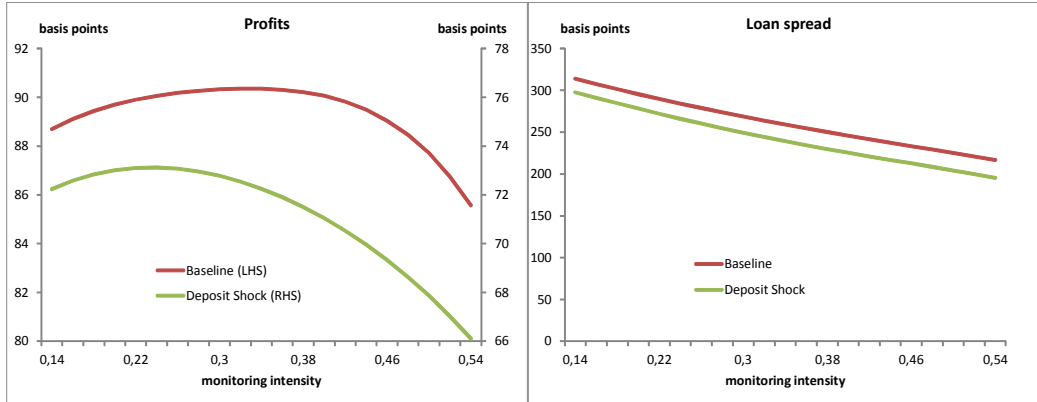
By contrast, the left hand panel of Figure 8 shows that in contradiction of the hypothesis of a risk-taking channel, the profit-maximising monitoring intensity moves to the right in response to a cut in monetary policy. The right hand panel illustrates that credit spreads are higher too. How does such a counterintuitive result arise? The partial effects are in place for a risk-taking channel: the lower deposit rate increases the incentive for inventors to enter production and

²⁰Dynamic versions of the model are currently under development.

²¹The profit maximising loan spread might be higher because the monitoring intensity is lower. The spread conditional on monitoring intensity gives the best measure of risk-adjusted spreads.

²²In equilibrium, however, once the equilibrium reduction in monitoring intensity is taken into account, the default rate is actually higher.

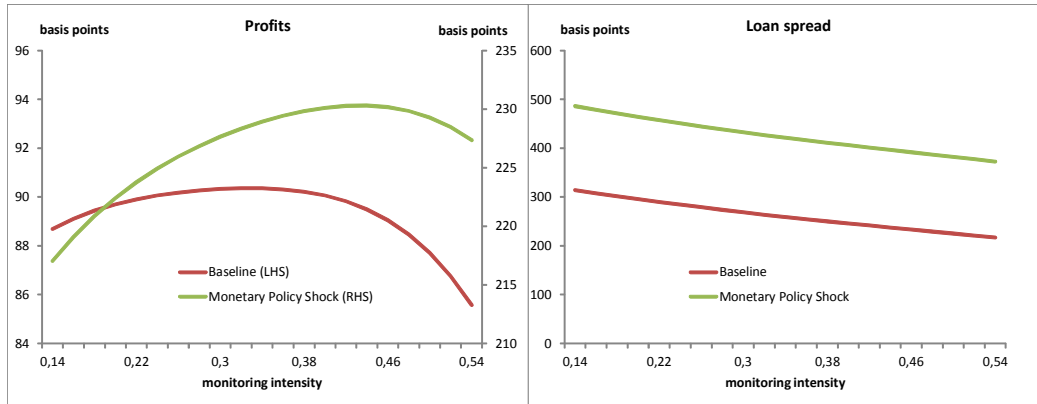
Figure 7: Deposit shock



encourages existing entrepreneurs to continue rather than exit and the default rate would fall *ceteris paribus*. The reduction in the deposit rate increases the lending spread and thus the incentive to lend. The problem is the balance sheet constraint. Without an external source of funds, the only way for the bank to reduce the resulting excess demand for loans is to make borrowing less attractive and so loan rates increase. But since increasing the loan rate is profitable and monitoring is costly, why doesn't the bank rely solely on interest rates to equilibrate its balance sheet? Why does monitoring also increase? The reason is that the credit quality of the loan portfolio is improved by monitoring more but with less increase in the loan interest rate because it increases the equilibrium turnover rate of firms. So a combination of intermediate increases in loan interest rates and the monitoring intensity offers a better mix of loan spread and default risk than higher loan interest rates alone.

Overall, therefore, the comparison of the two shocks shows that, at least as far as the model presented in this paper is concerned, the global savings glut is a more likely explanation for the reduction in credit standards and lending spreads observed in the run-up to the crisis than the risk-taking channel of monetary policy.

Figure 8: Monetary policy shock



7. Conclusion

The model developed in this paper provides a theory of how banks manage credit risk when they make multi-period loans to entrepreneurs who have private information on firm profits and future prospects. If entry and exit were frictionless, then the individually rational choices of entrepreneurs and inventors would select the distribution of firms with the lowest credit risk and the asymmetry of information would have no bite. This is not the case with entry and exit costs since there will now be a segment of firms who choose to continue in circumstances in which they would not decide to start. These firms are at the highest risk of defaulting in the near future and the bank has an interest in trying to reduce this portion of the distribution through its credit standards. The bank monitors continuing loans stochastically to discover breaches of loan covenants and the frequency of monitoring is inversely related to the expected duration or maturity of the loan. But credit controls are costly, directly due to the cost of monitoring and indirectly through the interest rate the bank can charge on loans. Credit standards are a form of control right over the decision to continue a firm - the tighter the standards, the less control exercised by the borrower. Borrowers, therefore, are willing to pay an interest rate premium for greater control rights. So in deciding how to set its credit standards, a bank needs to take into consideration the cost of monitoring and enforcing its covenants, the effect of credit standards on default risk and the loan interest rate it can charge for different contract terms.

The model shows how these competing considerations can be equilibrated whilst ensuring that the bank has sufficient deposits to fund its lending.

In the model there was only one bank and it had complete flexibility to set its terms and conditions subject to the balance sheet constraint. As the only actor capable of influencing aggregate outcomes, the bank effectively determines the distribution of firms in the economy and thus the allocation of resources. Banking is a socially useful activity since lending with no monitoring leads to the highest credit risk distribution which is not welfare maximising except under extreme conditions. However, since there is an externality from higher monitoring, the bank's choice of monitoring intensity will be below that of a social planner.

In the final section, the paper examined the effects of a cut in monetary policy and a deposit shock on equilibrium credit standards and credit spreads. In a closed economy, a cut in monetary policy led to a tightening of credit standards and an increase in credit spreads in contradiction to the hypothesis of the risk-taking channel of monetary policy. By contrast, an exogenous increase in deposits reduced credit standards and credit spreads, consistent with what was observed prior to the financial crisis. Therefore the model rejects the risk-taking channel in favour of the savings glut hypothesis. Macroprudential regulators should clearly take account of both types of shocks but the model suggests that they should be particularly wary of lending practices when there are large capital inflows into the economy.

8. References

References

- [1] Acharya, V. and H. Naqvi (2012), 'The seeds of a crisis: a theory of bank liquidity and risk-taking over the business cycle', *Journal of Financial Economics*, vol. 106(2), pp. 349-366.
- [2] Adrian, T., E. Moench, H.S. Shin (2010), 'Financial intermediation, asset prices, and macroeconomic dynamics' *Federal Reserve Bank of New York Staff Report* no 422, January 2010.
- [3] Adrian, T. and H.S. Shin (2008a), 'Liquidity and leverage' *Federal Reserve Bank of New York Staff Report* no 328, May 2008.

- [4] Adrian, T. and H.S. Shin (2008b), ‘Procyclical leverage and value-at-risk’ *Federal Reserve Bank of New York Staff Report* no 338, July 2008.
- [5] Adrian, T. and H.S. Shin. (2009), ‘Financial Intermediaries and Monetary Economics’ in the **Handbook of Monetary Economics**, Vol. 3, ed. B. Friedman and M. Woodford, North-Holland.
- [6] Agur, I. and M. Demertzis, (2012), ‘Monetary policy and excessive bank risk taking’ *European Central Bank Working Paper Series* no 1457.
- [7] Altunbas, Y., L. Gambacorta and D. Marques-Ibanez (2010), ‘Does monetary policy affect bank risk taking?’, *Bank for International Settlements Working Papers* no 298.
- [8] Arellano, C., Y. Bai and J. Zhang (2012), *Journal of Monetary Economics*, vol. 59(6) pp. 533-549.
- [9] Baily, M., E. Bartelsman and J. Haltiwanger (2001), ‘Labor productivity: structural change and cyclical dynamics’, *The Review of Economics and Statistics*, vol. 83(3), pp. 420-433.
- [10] Bartelsman, E. and M. Doms (2000), ‘Understanding productivity: lessons from longitudinal microdata’, *Journal of Economic Literature*, vol. 38, pp. 569-594.
- [11] Bernanke, B.S. (2010), ‘Monetary policy and the housing bubble’, speech at the Annual Meeting of the American Economic Association, Atlanta, Georgia.
- [12] Borio, C. and P. Disyatat (2011), ‘Global imbalances and the financial crisis: link or no link?’ *Bank for International Settlements Working Papers* no 346.
- [13] Borio, C. and H. Zhu (2008), ‘Capital regulation, risk-taking and monetary policy: a missing link in the transmission mechanism?’ *Bank for International Settlements Working Papers* no 268.
- [14] Broecker, T. (1990), ‘Credit worthiness tests and interbank competition’, *Econometrica*, vol 58, pp. 429-452.
- [15] Dell’Ariccia, G., D. Igan, L. Laeven and H. Tong with B. Bakker and J. Vandenburg (2012), ‘Policies for macrofinancial stability: how to deal with credit booms’ *IMF Staff Discussion Note* 12/06.

- [16] Dell’Ariccia, G. , L. Laeven and R. Marquez (2010), ‘Monetary policy, leverage and bank risk-taking’ *IMF Working Paper* no 10/276.
- [17] Diamond, D.W. (1984), ‘Financial intermediation and delegated monitoring’ *Review of Economic Studies*, vol. 51(3), pp. 393-414.
- [18] Djankov, S., R. La Porta, F. Lopez-de-Silanes and A. Shleifer (2002), ‘The regulation of entry’, *The Quarterly Journal of Economics*, vol.117, no.1 pp.1-37.
- [19] Fariñas, J. and S. Ruano (2005) ‘Firm productivity, heterogeneity, sunk costs and market selection’, *International Journal of Industrial Organization*, vol. 23, pp. 505-534.
- [20] Foster, L., J. Haltiwanger and C.J. Krizan (2006), ‘Market selection, reallocation, and restructuring in the US retail trade sector in the 1990s’, *The Review of Economics and Statistics*, vol. 88(4), pp. 748-758.
- [21] Freixas, X. and J.-C. Rochet (2008), **Microeconomics of Banking**, MIT Press, 2nd edition.
- [22] Gambacorta, L. (2009), ‘Monetary policy and the risk-taking channel’, *Bank for International Settlements Quarterly Review*, December.
- [23] Gale, D. and M. Hellwig (1985), ‘Incentive-compatible debt contracts: the one-period problem’ *Review of Economic Studies*, vol. 52, 4, pp. 647-663.
- [24] Gorton, G.B. and P. He (2008), ‘Bank credit cycles’, *Review of Economic Studies*, vol. 75, no. 4, pp. 1181-1214.
- [25] Hellwig, M. (1991), ‘Banking, financial intermediation and corporate finance’ in **European Financial Integration**, ed. A Giovannini and C. Mayer, Cambridge University Press.
- [26] Holmstrom, B. and J. Tirole (1997), ‘Financial intermediation, loanable funds and the real sector’, the *Quarterly Journal of Economics*, vol 112(3), pp. 663-691.
- [27] Hopenhayn, H. (1992), ‘Entry, exit and firm dynamics in long run equilibrium’, *Econometrica*, vol. 60(5), pp. 1127-1150.

- [28] Hopenhayn, H. and E. Prescott (1992), ‘Monotonicity and stationary distribution for dynamic economies’, *Econometrica*, vol. 60(6), pp. 1387-1406.
- [29] Jiménez, G., S. Ongena, J. L. Peydró and J. Saurina (2008) ‘Hazardous times for monetary policy: what do twenty-three million bank loans say about the effects of monetary policy on credit risk?’ *Documentos de Trabajo*, no. 0833, Banco de España.
- [30] Ioannidou, V., S. Ongena, and J.L. Peydró (2008), ‘Monetary policy, risk-taking and pricing: evidence from a quasi-natural experiment’ paper presented at the 9th Jacques Polak Annual Research Conference hosted by the International Monetary Fund, Washington DC 13-14 November 2008.
- [31] Lown, C. and D.P. Morgan (2008), ‘The credit cycle and the business cycle: new findings using the loan officer opinion survey’, *Journal of Money, Credit and Banking*, vol 38(6), pp.1575-1597.
- [32] King, M.A. (2010) speech at the University of Exeter.
- [33] McCall, J. (1970), ‘Economics of information and job search’ *Quarterly Journal of Economics*, vol 84, pp. 113-126.
- [34] Maddaloni, A. and J.-L. Peydró (2010), ‘Bank risk-taking, securitization, supervision and low interest rates: evidence from the Euro area and the US lending standards’, *European Central Bank Working Paper Series* no 1248.
- [35] Moore, G.H. (1956), ‘Changes in the quality of credit: the quality of credit in booms and recessions’, *Journal of Finance*, vol. 11(2), pp.288-300.
- [36] Portes, R. (2009), ‘Global imbalances’ in M. Dewatripont, X. Freixas and R. Portes (eds), **Macroeconomic stability and financial regulation: key issues for the G20**, Centre for Economic Policy Research, London.
- [37] Ramey, V. and M. Shapiro (2001), ‘Displaced capital: a study of aerospace plant closings’, *Journal of Political Economy*, vol. 109(5), pp. 958-992.
- [38] Ruckes, M. (2004), ‘Bank competition and credit standards’, *Review of Financial Studies*, vol. 17, no. 4, pp. 1073-1102.
- [39] Stokey, N. and R. Lucas (1989), **Recursive Methods in Economic Dynamics**, Harvard University Press.

- [40] Tauchen, G. (1986) ‘Finite state Markov chain approximations to univariate and vector autoregressions’, *Economic Letters*, vol. 20, 2, pp. 177-181.
- [41] Taylor, J. (2009), ‘The financial crisis and the policy response: an empirical analysis of what went wrong’, *NBER working papers*, no. 14631.
- [42] Topkis, D., (1998), **Supermodularity and Complementarity**, Princeton University Press.
- [43] Williamson, S.D. (1986), ‘Costly monitoring, financial intermediation and equilibrium credit rationing’ *Journal of Monetary Economics*, vol.18, pp.159-179.

9. Appendix

9.1. Proofs of Propositions

Proposition 1 *Given Assumptions A and B and a banking contract ψ , unique, bounded and mutually consistent functions $V_E(a; \psi, V_I)$ and $V_I(a; \psi, V_E)$ exist. These value functions yield unique and continuous functions in ψ for the entry and exit thresholds a_E and a_X respectively. $a_E(\psi)$ and $a_X(\psi)$ are both strictly increasing in r .*

Proof. It greatly simplifies the presentation of the proof (and with no loss of generality) to ignore the presence of the default option and focus on the choices of voluntary exit or continuation.

Let $\mathcal{B}(A)$ denote the set of all bounded functions on the set A . And for $v \in \mathcal{B}(A)$, let $\|v\| = \sup_a |v(a)|$ be the usual sup norm. Re-write equation (4) as

$$V_E(a) = q(a) - r + \beta \max \{ \phi_1(V_E, V_I, a), \omega_1(V_I, a') \}$$

where

$$\phi_1(V_E, V_I, a) = \varphi I_{a_T}(a) \left(\int_A V_I(a'; \cdot) G(da') - L \right) + (1 - \varphi I_{a_T}(a)) \int_A V_E(a'; \cdot) F(da', a)$$

$I_{a_T}(a)$ is an indicator function with value 1 if $a < a_T$ and zero otherwise and

$$\omega_1(V_I, a') = \int_A V_I(a'; \cdot) G(da') - L$$

Define T_1 to be the operator $(V_E, V_I) \mapsto T_1(V_E, V_I)$:

$$T_1(V_E, V_I)(a) = q(a) - r + \beta \max \{ \phi_1(V_E, V_I, a), \omega_1(V_I, a') \}$$

For any $V_I \in \mathcal{B}(A)$, $V_E \mapsto T_1(V_E, V_I)$ satisfies both of Blackwell's sufficient conditions for a contraction on $\mathcal{B}(A)$ with modulus β . Thus a unique V_E exists for a given V_I with

$$V_E = T_1(V_E, V_I) \tag{13}$$

Similarly re-write equation (7) as

$$V_I(a) = r_d + \max \beta \{ \phi_2(V_E, a), \omega_2(V_I, a') \}$$

where

$$\phi_2(V_E, a) = \int_A V_E(a'; \cdot) F(da', a) - S$$

and

$$\omega_2(V_I, a') = \int_A V_I(a'; \cdot) G(da')$$

Define T_2 to be the operator $(V_E, V_I) \mapsto T_2(V_E, V_I)$:

$$T_2(V_E, V_I)(a) = r_d + \max \beta \{ \phi_2(V_E, a), \omega_2(V_I, a') \}$$

Again, for any $V_E \in \mathcal{B}(A)$, $V_I \mapsto T_2(V_E, V_I)$ satisfies both of Blackwell's sufficient conditions for a contraction on $\mathcal{B}(A)$ and a unique V_I exists for a given V_E with

$$V_I = T_2(V_E, V_I) \tag{14}$$

The existence of each value function individually does not, however, imply the existence or uniqueness of any pair of functions (V_E, V_I) satisfying both conditions (13) and (14) simultaneously.

As a preliminary step towards the proof of the existence of a unique pair note that for $V_I, V'_I \in \mathcal{B}(A)$

$$|\omega_1(V_I, a') - \omega_1(V'_I, a')| = \left| \int_A [V_I(a'; \cdot) - V'_I(a'; \cdot)] G(da') \right| \leq \|V_I - V'_I\|$$

The absolute value of the difference in expected value between any V_I and V'_I must be less than the largest absolute difference. By similar arguments:

$$|\omega_2(V_I, a') - \omega_2(V'_I, a')| = \left| \int_A [V_I(a'; \cdot) - V'_I(a'; \cdot)] G(da') \right| \leq \|V_I - V'_I\|$$

$$|\phi_2(V_E, a) - \phi_2(V'_E, a)| = \left| \int_A [V_E(at; \cdot) - V'_E(at; \cdot)] F(dat, a) \right| \leq \|V_E - V'_E\|$$

Function $\phi_1(V_E, V_I, a)$ is a bit more tricky because of the presence of both value functions but note that for $V_E, V'_E, V_I, V'_I \in \mathcal{B}(A)$

$$\begin{aligned} |\phi_1(V_E, V_I, a) - \phi_1(V'_E, V'_I, a)| &= \left| \begin{aligned} &\varphi I_{a_T}(a) \left[\int_A [V_I(a'; \cdot) - V'_I(a'; \cdot)] G(da') \right] \\ &+ (1 - \varphi I_{a_T}(a)) \left[\int_A [V_E(at; \cdot) - V'_E(at; \cdot)] F(dat, a) \right] \end{aligned} \right| \\ &\leq \|V_E - V'_E\| \vee \|V_I - V'_I\| \\ \text{since } 0 &\leq \varphi I_{a_T}(a) \leq 1 \end{aligned}$$

Now let \mathcal{M} be the set of ordered pairs (V_E, V_I) such that both V_E and V_I are in $\mathcal{B}(A)$. Impose the following metric d on \mathcal{M} :

$$d((V_E, V_I), (V'_E, V'_I)) = \|V_E - V'_E\| \vee \|V_I - V'_I\|$$

where $a \vee b$ is the max of a and b . Now consider the operator $T : \mathcal{M} \rightarrow \mathcal{M}$ defined by

$$T(V_E, V_I) = (T_1(V_E, V_I), T_2(V_E, V_I))$$

A fixed point exists if

$$(V_E, V_I) = (T_1(V_E, V_I), T_2(V_E, V_I))$$

which is equivalent to

$$V_E = T_1(V_E, V_I)$$

and

$$V_I = T_2(V_E, V_I)$$

which are the separate value functions. Fix $a \in A$ and observe that

$$\begin{aligned} |T_1(V_E, V_I)(a) - T_1(V'_E, V'_I)(a)| &= \beta |\phi_1(V_E, V_I, a) \vee \omega_1(V_I, a') - \phi_1(V'_E, V'_I, a) \vee \omega_1(V'_I, a')| \\ &\leq \beta \{ |\phi_1(V_E, V_I, a) - \phi_1(V'_E, V'_I, a)| \vee |\omega_1(V_I, a') - \omega_1(V'_I, a')| \} \\ &\leq \beta \{ \|V_E - V'_E\| \vee \|V_I - V'_I\| \} \end{aligned}$$

where the first inequality is simply an example of the general property that $|\max(a, b) - \max(c, d)| \leq \max(|a - c|, |b - d|)$ for any $a, b, c, d \in \mathbb{R}$ and the second inequality uses the properties of ϕ_1 and ω_1 stated above. Taking the supremum over both sides:

$$\|T_1(V_E, V_I)(a) - T_1(V'_E, V'_I)(a)\| \leq \beta \{ \|V_E - V'_E\| \vee \|V_I - V'_I\| \}$$

Exactly the same arguments give

$$\|T_2(V_E, V_I)(a) - T_2(V'_E, V'_I)(a)\| \leq \beta \{ \|V_E - V'_E\| \vee \|V_I - V'_I\| \}$$

Therefore

$$\|T_1(V_E, V_I)(a) - T_1(V'_E, V'_I)(a)\| \vee \|T_2(V_E, V_I)(a) - T_2(V'_E, V'_I)(a)\| \leq \beta \{ \|V_E - V'_E\| \vee \|V_I - V'_I\| \}$$

This is the same as

$$d(T(V_E, V_I), T(V'_E, V'_I)) \leq \beta d((V_E, V_I), (V'_E, V'_I))$$

Hence T is a contraction mapping on the complete metric space (\mathcal{M}, d) establishing a unique fixed point exists. $V_E(a)$ and $V_I(a)$ are unique continuous functions.

The entry threshold a_E is determined by:

$$\int_A V_E(a'; \cdot) F(da', a_E) - S = \int_A V_I(a'; \cdot) G(da') \quad (15)$$

To show uniqueness, note that $\int_A V_I(a'; \cdot) G(da')$ is constant and independent of

a . Since $q(a) - r$ is increasing in a and $F(da', a)$ is stochastically increasing in a , $V_E(a)$ is increasing in a from Lemma 3.9.4 in Topkis (1998). Therefore

$\int_A V_E(a'; \cdot) F(da', a_E)$ is strictly increasing in a because $V_E(a'; \cdot)$ is an increasing

function and $F(da', a)$ is strictly stochastically increasing in a . By the intermediate value theorem there is a unique value of a_E . By analogous reasoning, there is a unique value of a_X .

The explanation why a_E and a_X are increasing functions of r is intuitive but the formal proof is long, tedious and available on request. The following sketches the argument. First, recall from the text that a_E and a_X are related by the following equation

$$\int_A V_E(a'; \cdot) F(da', a_X) = \int_A V_E(a'; \cdot) F(da', a_E) - S - L$$

Since S and L are constant, $V_E(a; \cdot)$ is an increasing function and $F(a', a)$ stochastically increasing in a , it follows that any increase in a_E has to be associated with an increase in a_X . So it suffices to show that a_E is strictly increasing in r to establish the case for a_X . To do this, consider the equation 15 that determines the entry threshold and fix the value of a_E . The left hand side is the expected value of being an entrepreneur conditional on a_E . There is clearly a direct reduction in $V_E(a'; \cdot)$ from an increase in r through the fall in $q(a) - r$. If the right hand side were unchanged, then a_E would have to rise to re-establish equality. Unfortunately the right hand side is not fixed because $V_I(a'; \cdot)$ is a function of $V_E(a'; \cdot)$ so the right hand side also falls. And $V_E(a'; \cdot)$ is a function of $V_I(a'; \cdot)$ which reduces the left hand side etc. Intuitively the direct effect should dominate the subsequent chain of indirect effects and left hand side should fall by more than the right. The formal proof establishes that this intuition is correct. ■

Proposition 2 *For each ψ there is a unique invariant distribution, $\bar{H}([0, a]; \psi)$ $\forall a \in A$.*

Proof. The transition equation for the end of period distribution of entrepreneurs can be re-written as an operator on probability measures:

$$(T^* H)(A, \psi) = I \int_{a_E}^1 G(a) + \int_{a_X}^1 F(a', a) H(da; \psi) - \varphi \int_{a_X}^{a_T} F(a', a) H(da; \psi) \quad (16)$$

Since $\int_{a_E}^1 G(a)$ and $F(a', a)$ are continuous probability measures and H is continuous by assumption, T^* maps a continuous function into another continuous function and thus has the Feller property. A is compact and therefore the operator function (16) satisfies the requirements for Theorem 12.10 in Stokey and Lucas (89) and an invariant distribution exists. $F(a', a)$ is stochastically increasing and,

since $0 \leq \varphi \leq 1$, the third term never dominates the second. Monotonicity plus the monotone mixing condition, Assumption A (iii), ensure that Theorem 2 of Hopenhayn and Prescott (92) is satisfied and the invariant distribution is unique.

■

Proposition 3 *There is a unique value \tilde{r} that ensures that the balance sheet of the bank is equal on both sides for given values of φ and ξ .*

Proof. For the bank to be able to match deposits with liabilities we require:

$$\frac{1}{2} = \tilde{H}(A, \tilde{r}) = I \tag{17}$$

Substitute $I = \frac{1}{2}$ into equation 16 and observe that all the terms on the right hand side are continuous. Ignore the third term and recall that that a_E and a_X are strictly and continuously increasing in r . Thus for any given processes $G(a)$ and $F(a', a)$ and I are fixed and $\tilde{H}(A; \psi)$ is continuously and strictly decreasing in r through the first two terms. a_T is also strictly increasing in r so the third term is also reducing $\tilde{H}(A; \psi)$ through the upper limit of the integral. The lower limit is also increasing but since $0 \leq \varphi \leq 1$, the third term never dominates the second. Therefore by the intermediate value theorem, there is only one value for r which satisfies equation (17). ■