Growth and Welfare Effects of Macroprudential Regulation

Pierre-Richard Agénor*

Preliminary draft: October 29, 2015
This version: January 18, 2016

Abstract

This paper studies the growth and welfare effects of macroprudential regulation in an overlapping generations model of endogenous growth with banking and agency costs. Indivisible investment projects combine with informational imperfections to create a double moral hazard problem à la Holmström-Tirole and a role for bank monitoring. When the optimal monitoring intensity is endogenously determined, an increase in the reserve requirement rate (motivated by systemic risk considerations) has ambiguous effects on investment and growth. The trade-off between ensuring financial stability and promoting economic growth can be internalized by choosing the reserve requirement rate that maximizes growth and welfare. However, the risk of disintermediation means that financial supervision may also need to be strengthened, and the perimeter of regulation broadened, if the optimal required reserve ratio is too high.

JEL Classification Numbers: E44, G28, O41.

---

*School of Social Sciences, University of Manchester, United Kingdom, and Centre for Growth and Business Cycle Research. I am grateful to King Yoong Lim for able research assistance and participants at various seminars for helpful comments. Financial support from the DFID-ESRC Growth Research Programme, under Grant No. ES/L012022/1, is gratefully acknowledged. The views expressed in this paper are my own.

1 Introduction

The growth effects of financial volatility, and ways to mitigate them, have been largely absent from recent discussions about the implications of the global financial crisis for financial reform. Indeed, much of the recent debate has focused almost exclusively on the implications of financial volatility for *short-term* economic stability and on the short-run benefits of financial regulation—especially macroprudential policies, which take a systemic approach in addressing financial vulnerabilities—in terms of mitigating procyclicality of the financial system and dampening short-run fluctuations in credit and output.

However, understanding the longer run effects of financial regulation is essential because of the potential dynamic trade-off associated with the fact that regulatory policies, designed to reduce procyclicality and the risk of financial crises, could well be detrimental to economic growth, due to their effect on risk taking and incentives to borrow and lend—despite contributing to a more stable environment in which agents can assess risks and returns associated with their investment decisions.

In low-income countries, where sustaining high growth rates is essential to increase standards of living and escape poverty, understanding the terms of this trade-off is particularly important. These countries are often characterized by an underdeveloped formal financial system, and thus limited opportunities to borrow and smooth shocks. The real effects of financial volatility on firms and individuals can therefore be not only large but also highly persistent, thereby translating into not only transitory drops in output but also adverse effects on growth.[1] In such conditions, the benefits of

[1] These adverse growth effects are consistent with the evidence showing that financial liberalization (to the extent that it is accompanied by greater financial volatility) may not

regulatory measures aimed at promoting financial stability could be fairly substantial. Yet, if regulatory constraints have a persistent effect on the risk-taking incentives of financial intermediaries—because, for instance, they induce structural shifts in banks' portfolio composition, in the form of a move away from risky assets toward safe investments—or more generally if they constrain their capacity to lend, they may translate into high interest rate spreads, suboptimal levels of borrowing by entrepreneurs to finance investment, and shifts of activity to less-regulated financial intermediaries, which could affect negatively growth and welfare. A key question therefore is to determine the *optimal* degree of financial regulation that internalizes this trade-off. Moreover, because the institutional environment in low-income countries is often weak, a related issue is what *type* of financial regulatory instruments should be implemented.

The literature on these issues, however, remains scant. One of the first analytical contributions in this area is Van den Heuvel (2008), who studied the welfare effects of bank capital requirements in a standard growth setting.[2] In line with the foregoing discussion, he argues that capital adequacy requirements may have conflicting effects on welfare. On the one hand, by inducing banks to hold less risky portfolios, they mitigate the probability of a financial crisis, which enhances welfare.[3] On the other, by inducing a shift in

contribute much to promoting growth; see for instance Misati and Nyamongo (2012) and the overview by Fowowe (2013). The latter study, in particular, highlights the need to strengthen prudential regulation to enhance the benefits of financial liberalisation. However, the potential adverse effects of prudential regulation itself are not discussed.

[2]A recent contribution by Barnea et al. (2015) also focuses on capital requirements. However, their focus is on the interactions between monetary policy and macroprudential regulation, rather than growth.

[3]Note that, as argued for instance by Dewatripont and Tirole (2012), equity capital may be equally effective in reducing incentives for excessive risk taking.Thus, capital requirements and portfolio restrictions may end up having the same effect of inducing banks to hold less risky portfolios.

banks' portfolios away from risky, but more productive, investment projects, toward safer, but less productive, projects, it may hamper economic growth and have an adverse effect on welfare. Capital requirements entail therefore a trade-off between banking efficiency and financial safety; as capital levels rise, there are costs, in terms of increased lending spreads or reduced loan volumes. However, a crucial limitation of the paper is that, because growth is exogenous, the implications of this trade-off for long-run growth cannot be fully explored.

This paper contributes to the literature on the growth and welfare effects of macroprudential regulation in several important ways. It uses an overlapping generations (OLG) endogenous growth model where financial intermediation is carried out only by banks. In contrast to existing studies, it focuses on reserve requirements—a prudential instrument that has been used extensively in both low- and middle-income developing countries (often as a substitute to monetary policy, as discussed by Agénor and Pereira da Silva (2015)) and has recently been made part of the liquidity requirement guidelines under the new Basel arrangement (see Basel Committee on Banking Supervision (2013)). In the model, the production of capital is subject to a dual moral hazard problem in the sense of Holmström and Tirole (1997): first, entrepreneurs, who need external funds to finance their investment projects, may be tempted to choose less productive projects with higher non-verifiable returns. Second, although bank monitoring mitigates the moral hazard problem associated with the behavior of entrepreneurs, the fact that banks use deposits from households to fund their loans creates an incentive to shirk when monitoring is costly. However, the model presented here departs from the Holmström-Tirole paradigm in two important ways. First, households cannot lend directly to producers; there is therefore only intermediated finance through banks. This assumption is more appropriate for a low-income environment, where capital markets are underdeveloped—if not entirely absent. Second, the intensity of monitoring, which affects private returns from shirking, is endogenously determined. Both assumptions turn out to be crucial for the results.

The key insights from the analysis are as follows. When the monitoring costs that financial intermediaries face are exogenous, an increase in the reserve requirement rate—motivated by the desire to constrain banks' capacity to lend, reduce the private sector leverage ratio, and mitigate systemic risk—has unambiguously negative effects on investment and economic growth. Making banks safer by requiring them to put away a fraction of the deposits that they receive reduces the supply of loanable funds. However, when optimal monitoring intensity is endogenously determined, an increase in the reserve requirement rate has ambiguous effects on investment, growth and welfare. The trade-off between ensuring financial stability and promoting economic growth can be internalized by choosing optimally the reserve requirement rate. Nevertheless, if reserve requirements are (optimally) set at prohibitive levels, they may foster disintermediation away from the banking sector and toward less regulated channels, which in turn may distort markets, weaken financial stability, and reduce investment and growth. The risk of disintermediation means therefore that financial supervision may also need to be strengthened, and the perimeter of regulation broadened, when more aggressive macroprudential policies are implemented. The analysis also shows that it is optimal to impose high required reserve ratios on banks when their ability to monitor borrowers is weak. This result is consistent with the evidence suggesting that reserve requirements continue to be used extensively in developing countries, and much less so in industrial countries.

The remainder of the paper is organized as follows. Section 2 describes the model, taking the intensity of monitoring as well as the reserve requirement rate as given. The model dwells on Holmström and Tirole (1997) and its extension to a growth and financial development setting by Chakraborty and Ray (2006), but incorporates the two major differences mentioned earlier.[4] In addition, and in contrast to these contributions, monitoring outlays are pecuniary in nature and their opportunity cost affects incentive constraints. The optimal financial contract is characterized in Section 3. The equilibrium level of investment is determined in Section 4, whereas the balanced growth equilibrium path is characterized in Section 5. Autonomous changes in monitoring intensity and the reserve requirement rate are studied in Section 6. In Section 7, optimal monitoring is analyzed and the growth- and welfare-maximizing values of the reserve requirement rate are solved for numerically. The last section provides some concluding remarks and discusses perspectives for further research.

2 Economic Environment

The economy consists of a continuum of risk-neutral agents who live for two periods, adulthood and old age. These agents are of two types: an exogenous fraction $n \in (0,1)$ are *workers*, the remaining are *entrepreneurs*. Without loss of generality, n is normalized to 0.5 and the measure of each type to one. Population is constant. Domestic agents can access the international capital market but access is asymmetric: agents can lend (deposit funds) but they cannot borrow. Foreign intermediaries do not lend because they cannot

[4] Chakraborty and Ray (2006) focus on the link between growth and the evolution of market-based and bank-based financial systems. They do not discuss financial regulation or the optimal setting of policy instruments.

legally enforce in domestic courts the terms of loan contracts.[5] There are three production sectors, all of them producing perishable goods, and a bank-dominated financial sector, which channels funds from savers to borrowers. There is also a financial regulator.

2.1 Workers and Entrepreneurs

A worker (or saver) is born with one unit of time in adulthood, which it supplies inelastically to the labor market. A generation-t worker's lifetime utility depends only upon second period consumption so that the entire wage income, w_t, is saved in adulthood. Workers do not have access to a monitoring technology and therefore do not lend directly to producers; they invest all their savings (or first-period income) either in bank deposits, d_t, or abroad. Arbitrage implies that both investments yield the same (gross) return, $R^D > 1$, which is set exogenously.

Each entrepreneur j, with $j \in [0, 1]$, is also born with one unit of labor time in adulthood, which is used to operate one of two types of technologies: a *modern technology*, which can be used to convert units of the final good into a marketable capital good; or a *traditional technology*, which can be used to produce only nonmarketed consumption goods. Whatever the technology chosen, operating it generates no income in the first period. Entrepreneurs therefore do not consume in that period either. They are altruists and derive utility from their old-age consumption, $c_{t+1}^{j,E}$, and bequests made to their offspring, b_{t+1}^{j}. Specifically, a typical generation-t entrepreneur's preferences are given by the 'warm-glow' utility function:

$$U_t^{j,E} = (c_{t+1}^{j,E})^{\beta}(b_{t+1}^{j})^{1-\beta}, \tag{1}$$

[5] This assumption is consistent with the evidence on lack of access to world capital markets for poor countries (see for instance Agénor (2012)) and helps to simplify the presentation.

where $\beta \in (0,1)$.[6]

Let z_{t+1}^j denote entrepreneur j's realized income in old age, which is derived later. Given Cobb-Douglas preferences in (1), optimal decision rules are linear in z_{t+1}^j. Thus, entrepreneur j leaves to his offspring a constant proportion of his realized income in old age:

$$b_{t+1}^j = (1-\beta)z_{t+1}^j, \tag{2}$$

the remaining fraction being consumed, so that $c_{t+1}^{j,E} = \beta z_{t+1}^j$.

Let $G_t(b)$ denote the cumulative distribution function of wealth distributed among generation-t entrepreneurs, that is, the proportion of them with wealth less than b. Given G_0 and $\{z_{t+1}^j\}_{t=0}^{\infty}$, equation (2) tracks therefore the wealth distribution through time.[7]

2.2 Production Sectors

The production sectors in the economy consist of a *final goods sector*, which produces a unique consumption good, a *home good sector*, which produces (using the traditional technology) the same consumption good but for own use only, and a *capital goods sector*, which supplies (using the modern technology) inputs to firms producing final goods.

2.2.1 Final Goods Sector

Competitive firms produce the final good (which can be either consumed or used as a production input) by combining raw labor and capital goods. The underlying private technology exhibits constant returns in capital and labor

[6] Altruism among workers can be readily incorporated in the model without qualitatively altering any of the basic results.

[7] The initial distribution G_0 is assumed to be continuous and differentiable.

inputs:

$$Y_t = A_t N_t^{1-\alpha} K_t^{\alpha}, \tag{3}$$

where $\alpha \in (0,1)$, N_t is the number of workers, $K_t = \int_{j \in E_t} K_t^j dG$ is the aggregate capital stock, with $E_t \in (0,1)$ denoting the set of entrepreneurs who supply capital goods at date t, and A_t a productivity parameter.

There is an Arrow-Romer type externality associated with the capital-labor ratio $k_t = K_t/N_t$, so that

$$A_t = A k_t^{1-\alpha}. \tag{4}$$

Combining (3) and (4) yields, in standard fashion, a linear relationship between (aggregate) production per worker, y_t, and capital per worker:

$$y_t = A k_t. \tag{5}$$

Final goods producers operate in competitive output and input markets so that equilibrium capital rental and wage rates, R_t^K and w_t, are determined by their marginal product:

$$R^K = \alpha A, \qquad w_t = (1-\alpha) A k_t. \tag{6}$$

To ensure that the gross return $R^K > 1$, the restriction $A > 1/\alpha$ is imposed.

2.2.2 Capital Goods Sector

Each capital good j is produced by a single entrepreneur j. Because (as noted earlier) generations of entrepreneurs are interconnected through a bequest motive, firm j is effectively infinitely lived. At any given period in time, the adult member of entrepreneurial family j is the owner-manager of that firm, converting units of the final good into capital with a one-period lag.

Each entrepreneur invests an indivisible amount q_t^j, which is taken as given for the moment. The investment project is risky. When the project fails, it yields nothing. When it succeeds, it produces a verifiable amount of capital equal to

$$K_{t+1}^j = q_t^j, \tag{7}$$

where for simplicity a one-to-one relationship between flows and stocks is assumed.

But as long as $q_t^j > b_t^j$, he has to raise the difference $l_t^j = q_t^j - b_t^j$ from banks. All entrepreneurs produce the same type of capital good and are price takers. The common return they earn from renting out their capital is $R^K > 1$, the (constant) marginal product of capital in a competitive equilibrium, given by (6). For simplicity, capital goods fully depreciate upon use.

2.2.3 Home Production

Entrepreneurs also have access to a traditional technology whose output is not marketed and entirely self-consumed. This technology enables an entrepreneur j to produce, with a one period lag, the same consumption good (in quantity x_{t+1}^j) that the final goods sector produces, using the unit time available and initial wealth at date t (the bequest obtained from generation $t-1$), b_t^j:

$$x_{t+1}^j = a_t (b_t^j)^{\delta} (1)^{1-\delta}, \tag{8}$$

where $\delta \in (0,1)$ and a_t is a productivity parameter. Thus, if entrepreneurs cannot borrow they can still produce consumption goods, albeit with diminishing marginal returns to initial wealth. The process $\{a_{t+1}\}_{t=0}^{\infty}$ is a weakly increasing sequence of positive numbers with $\lim_{t\to\infty} a_t = \bar{a}$. In effect, a_t improves exogenously through time due, for instance, to some learning-by-doing effect. At the same time, productivity improvements are bounded from above under the (plausible) assumption that the traditional technology can be improved only up to a certain point.

The entrepreneur's choice of technology depends upon which one gives him a higher income and whether or not he is able to obtain external finance to operate the modern technology.

2.3 The Financial Sector

Financial intermediaries consist of banks, which obtain their supply of loanable funds from workers' deposits and use them to lend to entrepreneurs for the purpose of building capital. However, these deposits are subject to a reserve requirement imposed by the regulator. For ease of exposition, each bank is assumed to lend to one entrepreneur only.

Banks are endowed with a technology (specialized skills) that allows them to inspect a borrowing entrepreneur's cash flows and balance sheet, observe the owner-manager's activities, and ensure that the entrepreneur conforms to the terms agreed upon in the financial contract.[8] Monitoring, although imperfect, helps to address a standard agency problem that banks face in lending to entrepreneurs.

Specifically, as in Holmström and Tirole (1997), suppose that each entrepreneur is allowed to choose between three types of investment projects, which differ in their success probability and the nonverifiable private benefits that they bring.[9] Suppose also that the entrepreneur must raise funds

[8] Households do not possess this technology, or even if they do, are too disparate to effectively use it. Thus, in standard fashion (see Diamond (1984)), banks act as delegated monitors. Note that the monitoring activities considered here differ from *ex post* monitoring in the costly state verification literature, where lenders monitor when the project outcome is realized and only when the borrower defaults on repayments. Accordingly, the cost of monitoring in that literature is more akin to a bankruptcy cost.

[9] Private benefits are nontransferable and capture the idea that the entrepreneur gets some kind of non-monetary return from some projects. A common interpretation is that

amounting to $q_t^j - b_t^j > 0$ for his investment. When the project succeeds, it realizes the amount of capital given in (7), with a verifiable return R^K. But when the project fails, it produces nothing; there is no remaining liquidation value.[10] The moral hazard problem arises from the fact that the probability of success depends on an unobserved action taken by the entrepreneur. The unobserved action can be interpreted as his choice of how to spend q_t^j. He can spend it on an *efficient* (good) project that results in success with probability $\pi^H < 1$ (and returning therefore $R^K q_t^j$), but uses up all of q_t^j. Or, he can spend it on one of two *inefficient* projects that may not succeed. One of these alternatives, a low-moral hazard project, costs $q_t^j - \upsilon q_t^j$, where $\upsilon \in (0,1)$, leaving υq_t^j for the entrepreneur to appropriate. The other inefficient choice, a high-moral hazard project, costs $q_t^j - V q_t^j$, where $V \in (0,1)$, thereby leaving $V q_t^j$ in private benefits. Inefficient projects carry both the same probability of success, $\pi^L < \pi^H$, but it is assumed that $0 < \upsilon < V < 1$. Hence, the entrepreneur will always prefer the high-moral hazard project over the low-moral hazard one.[11] Only the efficient project is, however, economically viable and thus socially valuable; to ensure that's the case, the following standard conditions are imposed:

Assumption 1. $\pi^H \alpha A - R^D > 0 > \pi^L \alpha A + V - R^D$.

Intuitively, this condition states that the expected net surplus per unit invested in a good project is positive, while that of a high-moral hazard

they capture effort. Lower effort is clearly a benefit to the entrepreneur, but (as discussed next) it also leads to a lower probability of success.

[10] Returns in this framework are verifiable at no cost.

[11] While entrepreneurs consume in the second period of life only, they invest in the first. As in Chakraborty and Ray (2006), it is assumed that the private benefits generated by negligent behavior cannot be invested. Instead, they have to be stored away for one period in the form of goods. Such storage yields zero net return but is unobservable and stored goods cannot be seized by lenders if the investment project fails.

project is negative—even after the private benefit is accounted for.[12]

Monitoring partially resolves the agency problem and reduces the entrepreneur's opportunity cost of being diligent. By monitoring borrowers, banks eliminate the high-moral hazard project but not the low-moral hazard one (Holmström and Tirole (1997)). Thus, because misrepresenting high levels of effort can be achieved only by performing some level of effort, an entrepreneur is left with two choices under monitoring: selecting the efficient project or the low-moral hazard project. At the same time, monitoring involves a non-pecuniary cost for the bank, representing a nonverifiable amount $\gamma \in (0, 1)$, in terms of goods, per unit invested. Hence, bank monitoring will be an optimal arrangement only if the gains from resolving agency problems outweigh the monitoring costs.

3 Optimal Financial Contract

In this setting, there are three parties to the (one period) financial contract: the entrepreneur, the bank and workers. Whether or not an entrepreneur prefers to be diligent depends, as noted earlier, upon appropriate incentives and outside monitoring by the bank. For its part, the bank chooses either to lend the full amount needed to invest in the efficient technology (net of the borrower's initial wealth) or not at all. Because workers delegate to the bank the task of monitoring entrepreneurs, banks must ensure that the return that savers obtain is sufficiently high for them to deposit their funds. This section characterizes the optimal contract when entrepreneurs behave diligently and choose only good projects.

[12]Because $\upsilon < V$, this restriction implies that the condition $0 > \pi^L \alpha A + \upsilon - R^D$ also holds.

3.1 Basic Structure

The optimal contract is such that no party (due to limited liability) earns anything when the project fails, whereas when it succeeds the gross return, R^K, is distributed so that

$$R^B_{t+1} + R^E_{t+1} + R^W_{t+1} = R^K, \tag{9}$$

where R^B_{t+1}, R^E_{t+1} and R^W_{t+1} denote the gross returns to the bank, the entrepreneur, and the savers, respectively.

Entrepreneur j invests q^j_t in the good project (using the modern technology) as long as it yields an incentive compatible return. As noted earlier, given that the banker always monitors if it lends, an entrepreneur will not choose the high-moral hazard project. The good project returns $R^E_{t+1}q^j_t$ with probability π^H, whereas the expected return to the low-moral hazard project (if it succeeds) is $\pi^L R^E_{t+1}q^j_t + \upsilon q^j_t$, that is, the sum of the (expected) market return plus the private return. The incentive compatibility constraint for the entrepreneur is thus $\pi^H R^E_{t+1}q^j_t \geq \pi^L R^E_{t+1}q^j_t + \upsilon q^j_t$, or equivalently

$$R^E_{t+1} \geq \frac{\upsilon}{\Delta\pi}, \tag{10}$$

where $\Delta\pi = \pi^H - \pi^L$.

The incentive compatibility constraint for the bank depends on the fact that it engages in monitoring. The monitoring cost is proportional (at the rate γ, as noted earlier) to the size of the project. Thus, the bank's incentive constraint for monitoring, and thus to engage in lending, requires that its expected return on a good project, net of monitoring outlays (valued at their opportunity cost), be greater than or equal to the expected return of a low-moral hazard project without monitoring, that is, $\pi^H R^B_{t+1}q^j_t - \gamma R^D q^j_t \geq$

$\pi^L R^B_{t+1} q^j_t$, or equivalently

$$R^B_{t+1} \geq \frac{\gamma R^D}{\Delta\pi}. \tag{11}$$

The contract's objective is to maximize the representative entrepreneur's expected share of the return, $\pi^H R^E_{t+1} q^j_t$, subject to the incentive compatibility constraints (10) and (11), as well as the participation constraint for workers

$$\pi^H R^W_{t+1} q^j_t \geq R^D d_t, \tag{12}$$

and the bank's resource constraint,

$$l^j_t = (1-\mu)d_t - \gamma q^j_t, \tag{13}$$

where $\mu \in (0,1)$ is a reserve requirement rate set by the financial regulator, and non-negativity constraints $R^i_{t+1} \geq 0$, where $i = E, B, W$.[13] Equation (12) indicates that the expected return from the project for workers must be at least equal to the return on deposits, whereas equation (13) indicates that the loan cannot exceed deposits (adjusted for required reserves) net of monitoring costs.

With $\pi^H R^E_{t+1} q^j_t$ and $\pi^H R^B_{t+1} q^j_t$ denoting the entrepreneur's and the banker's expected return, respectively, when the good project succeeds, equation (9) implies that the maximum income that savers can be expected to earn, or what Holmström and Tirole (1997) define as the *pledgeable expected (gross) income* that the borrower can credibly commit while still preserving incentives, is given by $\pi^H(R^K - R^E_{t+1} - R^B_{t+1})q^j_t$. The participation constraint for workers, equation (12), must therefore also satisfy $\pi^H R^W_{t+1} q^j_t \leq \pi^H(R^K - R^E_{t+1} - R^B_{t+1})q^j_t$, or equivalently, using (10) and (11),

$$\pi^H R^W_{t+1} q^j_t \leq \pi^H[R^K - (\frac{\upsilon + \gamma R^D}{\Delta\pi})]q^j_t. \tag{14}$$

[13] Because banks behave competitively, all of them offer the same contract that would be offered by a single bank that maximizes the entrepreneur's expected profits.

Combining (12) with (14) yields therefore

$$R^D d_t^j \leq \pi^H R_{t+1}^W q_t^j \leq \pi^H [R^K - (\frac{\upsilon + \gamma R^D}{\Delta\pi})] q_t^j.$$

Using the bank's resource constraint (13) and $l_t^j = q_t^j - b_t^j$ to eliminate d_t^j in this expression yields

$$\frac{R^D}{1-\mu}[(1+\gamma)q_t^j - b_t^j] \leq \pi^H [R^K - (\frac{\upsilon + \gamma R^D}{\Delta\pi})] q_t^j,$$

which can be rewritten as

$$b_t^j \geq b_m(q_t^j) = (1+\gamma)q_t^j - \frac{(1-\mu)\pi^H}{R^D}[R^K - (\frac{\upsilon + \gamma R^D}{\Delta\pi})] q_t^j. \qquad (15)$$

Thus, entrepreneurs with internal funds lower than the minimum level $b_m(q_t^j)$ cannot borrow because workers cannot be convinced, in the first place, to deposit the funds that banks need to lend.[14] To ensure that $b_m(q_t^j)$ is increasing in q_t^j the following condition is imposed:

Assumption 2. $\pi^H R^K - (1+\gamma)R^D/(1-\mu) < \pi^H(\upsilon + \gamma R^D)/\Delta\pi.$

This condition means that the expected per unit surplus of a good project, net of the cost of banking activity—which consists of interest paid to depositors and the opportunity cost of monitoring, with the deposit rate adjusted for the tax on financial intermediation imposed by reserve requirements—is less than the expected sum of the shares paid to the entrepreneur and the banker to ensure that they behave diligently. Thus $b_m(q_t^j)$ can interpreted as the expected per unit return used for deterring moral hazard incentives, net of the expected per unit surplus of the project.[15]

From (15), the following proposition can be established:

[14] Note that, if banks could borrow on world capital markets, foreign lenders would not provide funding either if (15) is not satisfied.

[15] Put differently, if $b_m(q_t^j)$ is positive, the incentive to engage in moral hazard behavior is eliminated.

Proposition 1. *The threshold level of wealth below which an entrepreneur cannot borrow is increasing in the reserve requirement rate, μ.*

This is fairly intuitive; because higher required reserves reduce the bank's loanable funds, and thus the income generated through lending, the incentive compatible constraint for savers requires more self financing by borrowers.

Assuming that condition (15) holds, and given perfect competition, in equilibrium the entrepreneur earns just enough to choose the efficient project, and each bank is paid just enough to have an incentive to monitor. The incentive constraints (10) and (11) therefore hold with equality, and using (9), yields

$$R^E_{t+1} = \frac{\upsilon}{\Delta\pi}, \qquad R^B_{t+1} = \frac{\gamma R^D}{\Delta\pi}, \tag{16}$$

$$R^W_{t+1} = R^K - (\frac{\upsilon + \gamma R^D}{\Delta\pi}) > 0. \tag{17}$$

Thus, when there is a greater incentive to divert funds (υ is larger), or when the monitoring activity is more costly (γ is higher), the payment share of the entrepreneur or the bank must be larger (and that of workers correspondingly smaller), to be incentive compatible.

With $\pi^H R^W_{t+1} q^j_t = R^D d^j_t$, and with (14) holding with equality, the value of deposits in bank j is given as

$$d^j_t = \frac{\pi^H}{R^D}[R^K - (\frac{\upsilon + \gamma R^D}{\Delta\pi})]q^j_t, \tag{18}$$

with the difference in a worker's savings in adulthood, $w_t - d_t \geq 0$, invested abroad at the same rate R^D.[16]

Let R^L_{t+1} denote the (gross) loan rate charged by the bank if the project succeeds. By definition, the payoff to lending l^j_t to entrepreneur j, should

[16] Given (18), the right-hand side of (15) can be written as $b_m(q^j_t) = q^j_t - [(1-\mu)d_t - \gamma q^j_t]$, which states that, given the size of investment and related monitoring costs, the initial wealth must equal the difference between investment and available resources, defined as the difference between deposits (net of required reserves) and monitoring costs.

the project succeed, must be such that $R^L_{t+1} l^j_t = R^B_{t+1} q^j_t$. Using (16), this expression yields therefore

$$l^j_t = \frac{\gamma R^D q^j_t}{\Delta\pi R^L_{t+1}}, \tag{19}$$

which determines how much the bank is willing to lend, given the loan rate and the size of the project. From this expression, the bank's total monitoring cost is

$$\gamma q^j_t = \frac{\Delta\pi R^L_{t+1} l^j_t}{R^D}. \tag{20}$$

Using (20), the bank's resource constraint (13) can be rewritten as

$$l^j_t = (1-\mu) d^j_t - \frac{\Delta\pi R^L_{t+1} l^j_t}{R^D}. \tag{21}$$

In equilibrium, only good projects are selected and banks make zero (expected) profits, so that $\pi^H R^L_{t+1} l^j_t = R^D d^j_t$. Substituting (21) for l^j_t in the zero-profit condition and rearranging yields[17]

$$R^L_{t+1} = \frac{R^D}{(1-\mu)\pi^H - \Delta\pi}, \tag{22}$$

which implies that the loan rate is increasing in both the deposit rate and the required reserve ratio. Because $\mu, \pi^H \in (0,1)$ and $\Delta\pi > 0$, the condition $(1-\mu)\pi^H < 1 + \Delta\pi$ always holds, implying that $R^L > R^D$. To ensure that R^L is also positive, the following condition must be imposed:

Assumption 3. $(1-\mu)\pi^H > \Delta\pi$.

Substituting (22) in (19) yields the optimal loan size as

$$l^j_t = \frac{\gamma[(1-\mu)\pi^H - \Delta\pi]}{\Delta\pi} q^j_t = \Theta q^j_t. \tag{23}$$

[17] This expression for the competitive loan rate can alternatively be derived from the aggregate equilibrium condition for the loan market, defining total loans as $\int_{j \in E_t} l^j_t dG$ and aggregate monitoring costs as $\gamma q_t \int_{j \in E_t} dG$. However, this does not change the results.

This expression shows that, irrespective of initial wealth b_t^j, banks finance a fixed proportion Θ (which varies inversely with μ) of each entrepreneur j's investment. To ensure that the loan size does not exceed the level of investment, the condition $\Theta < 1$ must be satisfied. In turn, given assumption 3, the following restriction is required:

Assumption 4. $\gamma/(1+\gamma) < \Delta\pi/(1-\mu)\pi^H$.

This assumption imposes an upper bound on the unit monitoring cost. As discussed later, a critical issue is the interplay between (15) and (23).

3.2 Entrepreneurial Income

Let z_{t+1}^j denote entrepreneur j's second period income and consider first the case where his first-period initial wealth is insufficient, $b_t^j < b_m(q_t^j)$, to obtain external financing. He can either deposit his assets abroad (at the same rate R^D as workers) or use them in household production.[18] He will engage in the latter as long as $a_t(b_t^j)^\delta \geq R^D b_t^j$, that is, $b_t^j \leq \hat{b}_t^j = (a_t/R^D)^{1/(1-\delta)}$. This will be true under appropriate restrictions.[19] The entrepreneurs' income in that case is given by

$$z_{t+1}^j = a_t(b_t^j)^\delta. \tag{24}$$

Second, consider the case where entrepreneur j does borrow from the banking system, so that $q_t > b_t^j \geq b_m(q_t^j)$. From (16),

$$z_{t+1}^j = \frac{\upsilon}{\Delta\pi} q_t^j = \frac{\upsilon}{\Delta\pi}(b_t^j + l_t^j). \tag{25}$$

[18] If $b_t^j < b_t^L \; \forall j$, banks would not be able to lend to any entrepreneur and would therefore not accept any deposits.

[19] To ensure that the entrepreneur chooses to invest in the traditional technology, instead of investing in deposits abroad, it must be assumed that $b_m < \hat{b}_0 = (a_0/R^D)^{1/(1-\delta)}$, where b_m, defined formally later, is the minimum level of wealth needed to qualify for bank financing when $q = q_m$.

4 Investment Decision

Having characterized financial contracts and returns from an arbitrary investment level, the analysis now turns to the entrepreneur's investment decision, that is, the optimal choice of q_t^j. It then considers the case where there are minimum capital requirements.

4.1 Optimal Investment and Borrowing

As noted earlier, for an investment of $q_t^j > b_t^j$, the minimum amount of initial wealth required to qualify for bank finance is $b_t^j \geq b_m(q_t^j)$. Given optimal contracts and financing arrangements for any investment q_t^j, an entrepreneur j chooses q_t^j to maximize his income z_{t+1}^j, as defined in (25). From (6) and (15), the maximum level of investment, for a given level of entrepreneurial wealth, must satisfy

$$q_t^j \leq \frac{R^D b_t^j}{(1+\gamma)R^D - (1-\mu)\pi^H[\alpha A - (\upsilon + \gamma R^D)/\Delta\pi]} = \Phi b_t^j, \tag{26}$$

where, in addition to Assumption 2, it must also be that $\Phi > 1$ to ensure positive borrowing, or equivalently

Assumption 5. $\gamma R^D < (1-\mu)\pi^H[\alpha A - (\upsilon + \gamma R^D)/\Delta\pi]$.

This assumption generates another upper bound on γ.

Let $\tilde{q}_t^j$ denote the maximum level of investment, for which (26) holds as an equality. For b_t^j given, the required level of bank loans is thus $\tilde{q}_t^j - b_t^j$. However, from equation (23), at the level of investment $\tilde{q}_t^j$ it is optimal for banks to provide $\Theta\tilde{q}_t^j$. Thus, an equilibrium with maximum investment $\tilde{q}_t^j$ exists if and only if $\tilde{q}_t^j - b_t^j = \Theta\tilde{q}_t^j$ as well, that is,

$$\tilde{q}_t^j = \frac{b_t^j}{1-\Theta}. \tag{27}$$

In turn, for (26), holding as an equality, and (27) to hold at the same time requires

$$\Phi = \frac{1}{1-\Theta}. \tag{28}$$

If this condition is not satisfied, then entrepreneur j is constrained in his optimal investment choice; because banks are always on their supply curve, the *actual* level of investment is in that case equal to $(1-\Theta)^{-1}b_t^j < \Phi b_t^j$, which ensures that the participation constraint (15) is satisfied—otherwise, banks would not be able to generate the resources to lend and therefore (23) would not apply—but entrepreneurs are subject to credit rationing.[20] These results can be summarized in the following proposition:

Proposition 2. *An equilibrium with positive lending and income maximization by entrepreneurs requires* $\Phi(1-\Theta) = 1$, *where* $\Theta < 1$ *and* $\Phi > 1$. *If so the equilibrium level of investment is* Φb_t^j. *If not entrepreneurs are subject to credit rationing; actual (constrained) investment is* $(1-\Theta)^{-1}b_t^j$, *which must be less than* Φb_t^j *to ensure that the participation constraint holds.*

Thus, all entrepreneurs in the range $b_t^j > b_m(q_t^j)$ borrow from the bank, as long as $\tilde{q}_t^j - b_t^j > 0$. By contrast, $b_t^j < b_m(q_t^j)$ for any $q_t^j > \tilde{q}_t^j$. Thus, an entrepreneur wishing to invest more than $\tilde{q}_t^j$ cannot obtain funds for his project, and is rationed out of the credit market; he can only resort to household production in that case and earn an income defined in (24).

Using (25) and the results of Proposition 2, the entrepreneur's income, depending on the level of investment, is thus

$$\tilde{z}_{t+1}^j = \begin{cases} \upsilon\Phi b_t^j/\Delta\pi & \text{if } \Phi(1-\Theta) = 1 \\ \upsilon(1-\Theta)^{-1}b_t^j/\Delta\pi & \text{if } \Phi(1-\Theta) > 1 \end{cases}. \tag{29}$$

[20] Recall that banks cannot borrow abroad; but even if they could, the same participation constraint would hold for foreign savers. Note also that in the standard Holmström-Tirole model rationing does not emerge in equilibrium; in that setting the residual amount of resources needed to achieve the maximum level of investment $\tilde{q}_t^j$ is provided by direct borrowing from savers. In the present setting domestic capital markets do not exist and entrepreneurs cannot borrow abroad.

Note that the case where $\Phi(1-\Theta) < 1$ can be ruled out because in that case one would have $(1-\Theta)^{-1}b_t^j > \Phi b_t^j$, and the participation constraint would be violated.

4.2 Minimum Capital Requirements

Recall that entrepreneurs operate either a modern or a traditional technology. Suppose now that entry into modern-sector activities requires a setup cost, in the form of fixed capital requirements and costs of adapting newer types of technologies. Thus, any entrepreneur wishing to produce capital goods must invest a minimum of q_m. For an equilibrium with maximum investment, condition (28) must again hold.[21] Conversely, credit-rationed entrepreneurs, with $b_t^j < b_m$, operate the traditional technology, and leave bequests according to, given (2) and (24),

$$b_{t+1}^j = (1-\beta)z_{t+1}^j = (1-\beta)a_t(b_t^j)^\delta. \tag{30}$$

The minimum wealth level b_m and the wealth distribution determine the size of the traditional (home production) sector at any point in time. Indeed, $G_0(b_m)$ indicates the fraction of generation-0 entrepreneurs with assets less than b_m, and hence, the initial size of the traditional sector.

As noted earlier, the traditional technology is subject to exogenous productivity improvements. To rule out perpetual stagnation in the traditional sector, $\bar{a}$ must be allowed to be large enough to ensure that $b_{t+1}^j(\bar{a}) > b_m$. This means that entrepreneurial families who do not obtain external financ-

[21] The minimum investment size q_m associated with the use of the modern technology defines the minimum wealth b_m required to secure external finance. From (6) and (15), this constraint is given by $b_m \geq \Phi^{-1}q_m$. For an investment of q_m, the required level of loans is $l_m = q_m - b_m$; but from (23), banks are willing to provide Θq_m. Thus, an equilibrium with maximum investment q_m exists if only if $q_m - b_m = \Theta q_m$, that is, $b_m = (1-\Theta)q_m$, which in turn implies again that $\Phi(1-\Theta) = 1$.

ing initially would ultimately accumulate enough wealth to enter the modern sector anyway. But how long they remain in the traditional sector depends on the efficiency of the banking system and on the process characterizing a_t.

4.3 Graphical Illustration

The determination of the equilibrium levels of investment and loans is illustrated in Figures 1, 2 and 3. For an investment of q_t^j, the minimum amount of initial wealth required to qualify for a bank loan is $b_m(q_t^j) = \Phi^{-1}q_t^j$, which is linear and increasing in the amount invested, given Assumption 1 (see (15)). In all three figures, the point of intersection between this cut-off level, which passes through the origin, and the vertical line at q_m (point A) corresponds to b_m. Points located in the hatched area below $b_m(q_t^j)$ and those located above the 45-degree line cannot be an equilibrium; only points located between (or on) these curves, which imply $q_t^j \leq b_t^j$ (and thus $q_t^j \leq \Phi b_t^j$, given that $\Phi > 1$) are feasible. From (23), the supply of loans is also increasing in investment, $l_t^j = \Theta q_t^j$. In general, $\Theta \lesseqgtr \Phi^{-1}$; in all three figures, and without loss of generality, it is assumed that $\Theta > \Phi^{-1}$.[22]

Now, consider a generation-t entrepreneur with inherited bequest level $b_t^j > b_m$. Assume first that condition (28) is satisfied. In all three figures, $\tilde{q}_t^j = \Phi b_t^j = (1-\Theta)^{-1}b_t^j$ is given by the point of intersection of $b_m(q_t^j)$ with b_t^j, that is, point B, or equivalently point D, given that $q_t^j = b_t^j + l_t^j$ holds continuously. The distance BC represents self financing (initial wealth), whereas the distance $B'C$, which is also equal to BD by construction, represents bank borrowing. Because $b_t^j < b_m(q_t^j)$ for any $q_t^j > \tilde{q}_t^j$ (which corresponds to the hatched area under the curve $b_m(q_t^j)$), an entrepreneur desiring to invest more

[22]Note that by construction $q_t^j > b_t^j$ for both curves, given that they are located below the 45-degree line.

than $\tilde{q}_t^j$ is completely rationed from the credit market. He can only resort to home production in that case and earn an income $z_{t+1}^j = a_t(b_t^j)^\delta > R^D b_t^j$ (see equation (24)). This earning is given by the horizontal line EE', given that it does not depend on q_t^j.[23]

To illustrate whether an equilibrium exists or not when condition (28) is not satisfied, Figures 1, 2 and 3 distinguish three cases: $\Theta > \max(\Phi^{-1}, 1 - \Phi^{-1})$, $\Phi^{-1} < \Theta < 1 - \Phi^{-1}$, and $\Phi^{-1} < \Theta = 1 - \Phi^{-1}$. In all three figures, the line $b_t^j + \Theta q_t^j$, which corresponds to actual investment, is also represented. Points on that line located above the 45-degree line cannot be an equilibrium, because they imply that $q_t^j < b_t^j + l_t^j$, and neither can points located below that line, given that they correspond to $q_t^j > b_t^j + l_t^j$. The only possible equilibrium corresponds to the point of intersection between the two lines, that is, point F. However, in Figure 1, the slope of the loan supply curve is too steep and $\Phi(1 - \Theta) < 1$; the participation constraint is violated. On the horizontal axis, the intersection point with the line corresponding to b_t^j is at B'', located in the hatched area; as noted earlier, there is no equilibrium.

By contrast, in Figure 2 the loan supply curve is flatter and the condition $\Phi(1 - \Theta) > 1$ holds. The equilibrium point at F and B'' implies that the actual level of investment is lower than the maximum level that entrepreneur j would prefer, that is, point B, where the participation constraint holds exactly. The supply of loans is again given by $B'C$ (or equivalently now $B''F$), but there is credit rationing. Finally, in Figure 3, condition (28) holds, and there is no credit rationing in equilibrium.

[23] Note that for $b_t^j < b_m$, the entrepreneur would not invest either given the minimum size q_m; if so it would also invest in home production, assuming that initial wealth is such that $a_0(b_0^j)^\delta > R^D b_0^j$. This case is omitted to avoid cluttering the figure. Note also that there is always a discontinuity at point B, given that $a_t(b_t^j)^\delta > R^D b_t^j > b_t^j$.

5 Balanced Growth Equilibrium

From (2) and (29), the wealth of an entrepreneur j who is not credit constrained ($b_t^j \geq b_m(q_t^j)$) evolves according to

$$b_{t+1}^j = (1+g)b_t^j, \tag{31}$$

where the constant growth rate $1+g$ is defined as

$$1+g = \begin{cases} (1-\beta)\upsilon\Phi/\Delta\pi & \text{if } \Phi(1-\Theta) = 1 \\ (1-\beta)\upsilon(1-\Theta)^{-1}/\Delta\pi & \text{if } \Phi(1-\Theta) > 1 \end{cases}. \tag{32}$$

As implied by (26), the optimal investment choice is linear in entrepreneurial wealth. The aggregate stock of capital in $t+1$ depends on all investments undertaken in t. Define

$$Q_t = \int_{b_m}^{\infty} \tilde{q}_t^j dG_t, \qquad B_t = \int_{b_m}^{\infty} b_t^j dG_t.$$

Because optimal loan contracts ensure that all entrepreneurs behave diligently, aggregate (and per capita) capital produced is, using (7),

$$k_{t+1} = Q_t = \Theta B_t. \tag{33}$$

Using (31) yields $k_{t+1} = (1+g)k_t$, so that aggregate capital per capita grows at the same rate as entrepreneurial wealth.

From (5), because the aggregate production function is linear in capital, the growth of output mimics that of capital. Similarly, given that from (6) the equilibrium wage rate is linear in capital, and that second-period consumption of workers is equal to $R^D w_t$, worker consumption grows at the same rate as the rate of growth of the capital stock. For entrepreneurs, in equilibrium all of them access credit markets and borrow from banks. From (29), and because $c_{t+1}^E = \beta z_{t+1}^j$, their consumption is linear in wealth and hence grows at the same rate as workers' consumption and output.

In what follows, the focus will be on the case where condition (28) holds; the alternative case (credit rationing) can be studied in a similar manner and is left to the reader.

6 Autonomous Policy Changes

Consider first the effect of a permanent reduction in the unit monitoring cost, γ, perhaps through better contract enforcement, and suppose that the private benefit of the low-moral hazard project is decreasing and convex in monitoring intensity, so that $\upsilon = \upsilon(\gamma)$, with $\upsilon' < 0$, $\upsilon'' \geq 0$, and $\lim_{\gamma\to\infty}\upsilon'(\gamma) = 0$. Thus, monitoring not only helps to eliminate the high-moral hazard project, it also mitigates the private benefits that can be derived from (and thus the incentives to engage in) low-moral hazard projects. The following proposition can then be established:

Proposition 3. *A reduction in monitoring intensity, γ, when the private benefit of the low-moral hazard project is decreasing and convex in that variable, has ambiguous effects on investment and the steady-state growth rate.*

Equations (16) and (17) help to illustrate the direct impact of this change. A reduction in γ raises $\upsilon(\gamma)$ and the per-unit project return R^E_{t+1} that must be promised to entrepreneurs, because it increases their ability to divert resources. It also lowers R^B_{t+1}, the per-unit share of the project's return that must be allocated to bankers in order for them to find it worth monitoring as intensively as promised. From (23), and for a given level of investment, the supply of loans tends to fall. At the same time, equation (17) shows that the per-unit share of project return that can be credibly promised to workers supplying loanable funds is ambiguous in general, given that γ and $\upsilon(\gamma)$ move in opposite directions. The net effect depends on the efficiency of the monitoring technology. From (26) and (32), the effect on the optimal level of investment and output growth is thus also ambiguous. If υ' is relatively small, the impact on the banker's return will dominate. This will lead to greater borrowing and higher investment by all entrepreneurs. A lower unit monitoring cost therefore also increases the growth rate of output.

Consider now a permanent increase in the reserve requirement rate, μ, motivated by a desire to reduce across the board the leverage ratio of each private borrower, l_t^j/b_t^j, by constraining the capacity of banks to lend, and thereby increase the resilience of the financial system.[24] The following result can be established:

Proposition 4. *An increase in the reserve requirement rate, μ, with constant monitoring intensity, unambiguously lowers investment and the steady-state growth rate.*

Intuitively, the policy raises the cost of borrowing (see (22)), which leads to lower optimal investment (see (26)) and a lower growth rate of output. There is also an adverse level effect, because the policy tends now to raise the minimum wealth needed to borrow and enter the modern sector (see (15)). It thus also worsens income distribution among entrepreneurs. As discussed next, however, once monitoring intensity is endogenized, these results are no longer unambiguous when $\upsilon' < 0$.

7 Optimal Policy

In the foregoing analysis, both the intensity of monitoring and the reserve requirement rate were taken as given. In the first part of this section the in-

[24] Alternatively this policy can be viewed as an attempt by the financial regulator to mitigate the risk of bank runs or to provide partial deposit insurance to savers, by forcing banks to hold higher liquid reserves than they would otherwise. Modeling these alternative interprations, however, is beyond the scope of the present paper.

tensity of monitoring is endogenized, as part of the representative bank's optimization problem. In the second part the growth- and welfare-maximizing values of the reserve requirement rate, taking into account its impact on the optimal monitoring intensity, are derived.

7.1 Optimal Monitoring Intensity

The optimal choice of γ, with $\upsilon = \upsilon(\gamma)$, must also maximize the entrepreneur's expected payoff, $\pi^H R^E_{t+1} q^j_t$, which, using (16) and (26), can be written as

$$\pi^H R^E_{t+1} q^j_t = \frac{\pi^H \upsilon(\gamma)}{\Delta\pi} \Phi b^j_t. \tag{34}$$

To derive a tractable analytical solution to the problem, suppose as in Haavio et al. (2014) for instance, that $\upsilon(\gamma)$ takes the following functional form:

$$\upsilon(\gamma) = \begin{cases} \Gamma\gamma^{-\varepsilon/(1-\varepsilon)} & \text{if } \gamma > \gamma_m \\ \upsilon_m & \text{if } \gamma \leq \gamma_m \end{cases}, \tag{35}$$

where $\Gamma > 0$, $\upsilon_m > 0$, $\varepsilon \in (0, 1)$, and $\gamma_m \geq 0$. The first row of equation (35) shows that $\upsilon(\gamma)$ is differentiable and strictly convex for $\gamma > \gamma_m$ and that the monitoring technology is the more efficient the larger Γ is or the smaller ε is. The second row implies that there is a minimum efficient scale for monitoring investment projects or an upper bound for the private revenues. This upper bound ensures that the net rate of return on a bad project is negative even for low levels of γ.

Assuming that a unique interior solution exists (so that $\gamma > \gamma_m$), the first-order condition for this problem is, after substituting equations (35) in (34) and taking initial wealth as given,

$$-\left\{\gamma R^D - \frac{(1-\mu)\pi^H}{\Delta\pi}[\frac{\varepsilon}{1-\varepsilon}\upsilon(\gamma) - \gamma R^D]\right\}$$

$$\frac{-\varepsilon}{1-\varepsilon}\left\{(1+\gamma)R^D - (1-\mu)\pi^H[\alpha A - \frac{\upsilon(\gamma)+\gamma R^D}{\Delta\pi}]\right\} = 0,$$

from which it can be shown that

$$\gamma^* = \frac{\varepsilon[(1-\mu)\pi^H \alpha A - R^D]}{R^D + (1-\mu)\pi^H R^D/\Delta\pi}. \tag{36}$$

This solution value is feasible if it satisfies Assumptions 2, 4 and 5. If so equation (36) yields the following proposition:

Proposition 5. *The optimal level of monitoring intensity, when the private benefit of the low-moral hazard project is decreasing and convex in that variable, is decreasing in the reserve requirement rate, μ, and increasing in the elasticity of the monitoring technology, ε.*

Intuitively, a more efficient monitoring technology magnifies the benefit of monitoring and raises the optimal intensity of monitoring; this lowers the private benefit of the low-moral hazard project. However, as implied by Proposition 3, this does not necessarily promote investment and growth. More importantly for the issue at stake, a higher reserve requirement rate reduces the optimal intensity of monitoring because it reduces the bank's income if the project succeeds. Put differently, prudential regulation distorts the incentives of banks to monitor and lend.

From equations (26), (32) and (36) it can also be established that:

Proposition 6. *An increase in the reserve requirement rate, μ, with monitoring intensity set optimally and with $\upsilon' < 0$, has ambiguous effects on investment and the steady-state growth rate.*

Thus, in contrast to the results reported in Proposition 4, when monitoring intensity is endogenous and set optimally, it is possible for an increase in the reserve requirement rate to have a *positive* effect on investment and growth. Intuitively, as noted earlier, when υ is endogenous a higher monitoring intensity affects incentives for the entrepreneur and the banker in opposite directions; for the entrepreneur, it tends initially to increase the optimal level of investment (as implied by (26)), whereas for the banker, it increases monitoring costs and thus the loan rate. This tends in turn to mitigate borrowing and the increase in investment, which translates into ambiguous effects on growth. As discussed next, these results are important to study the determination of the optimal reserve requirement rate.

7.2 Optimal Reserve Requirement Rate

A natural benchmark for the optimal reserve requirement rate set by the financial regulator is the solution that maximizes either growth or welfare. In the first case, it is obvious that, given Proposition 4, the growth-maximizing value of μ, given by $d\ln(1+g)/d\mu = 0$, is 0 if υ is constant, and regardless of whether γ is set optimally or not. By contrast, when $\upsilon' < 0$ and γ is set optimally, a non-trivial solution may exist. This solution is obtained by solving the condition

$$\frac{d\ln(1+g)}{d\mu} = \frac{d\ln\upsilon(\gamma^*)}{d\mu} + \frac{d\ln\Phi(\mu)}{d\mu} = 0, \tag{37}$$

where $\upsilon(\gamma^*)$ and γ^* are given by (35) and (36), and with

$$\ln\Phi(\mu) = \ln R^D - \ln\left\{(1+\gamma^*)R^D - (1-\mu)\pi^H[\alpha A - \frac{\upsilon(\gamma^*)+\gamma^* R^D}{\Delta\pi}]\right\}.$$

However, the resulting expression is too complex to be solved analytically.

Consider next the welfare-maximizing solution. Suppose that the financial regulator is far sighted and benevolent, in the sense that it takes into account the welfare of all future generations of entrepreneurs and workers. To calculate the welfare for each generation, recall that there is no consumption in adulthood and that given (1), the indirect utility function of entrepreneurs in old age is linear in income, so that $U_{t+1}^E = \beta^\beta(1-\beta)^{1-\beta}\beta z_{t+1}$. Similarly, for workers, the indirect utility function is $c_{t+1}^W = R^D w_t$, given that all income (including interest) is consumed in adulthood. Recall also that each group represents half of the population. Thus, the welfare criterion is the equally weighted within each generation, but discounted sum of utility across an infinite sequence of generations (see De la Croix and Michel, 2002, p. 91):

$$\mathfrak{W}_t = \sum_{h=0}^{\infty} \Lambda^h \left\{ 0.5 \ln \beta^\beta (1-\beta)^{1-\beta} z_{t+h+1} + 0.5 \ln R^D w_{t+h} \right\}, \tag{38}$$

where $\Lambda \in (0,1)$ is the regulator's discount factor. From (6) and (29), along the balanced growth path,

$$\mathfrak{W}_t = \sum_{h=0}^{\infty} \Lambda^h \left\{ 0.5 \ln[\frac{\beta^\beta (1-\beta)^{1-\beta} \upsilon \Phi \tilde{b}_{t+h}}{\Delta\pi}] + 0.5 \ln R^D (1-\alpha) A \tilde{k}_{t+h} \right\}. \tag{39}$$

From (31) and (33), b_t and k_t grow at the same rate along the balanced growth path. Thus, along the steady-state equilibrium path, $\tilde{x}_{t+h} = (1+g)^{t+h} x_0$, where $x = b, k$. Substituting these results in (39), and ignoring constant terms, yields

$$\mathfrak{W}_t = \sum_{h=0}^{\infty} \Lambda^h \left\{ 0.5[\ln \upsilon(\gamma^*) + \ln \Phi(\mu)] + (t+h) \ln(1+g) \right\}. \tag{40}$$

Given that $\Lambda < 1$, $\mathfrak{W}_t$ is strictly concave and bounded, and the choice set is convex and compact. Thus, the optimization problem $\max_\mu \mathfrak{W}_t$ has a single solution. Solving (40) gives[25]

$$\mathfrak{W} \simeq 0.5 \frac{\ln \upsilon(\gamma^*)}{1-\Lambda} + \frac{\Lambda}{(\Lambda-1)^2} \ln(1+g) \tag{41}$$

$$-\frac{0.5}{1-\Lambda} \left\{ \ln[(1+\gamma^*)R^D - (1-\mu)\pi^H[\alpha A - \frac{\upsilon(\gamma^*) + \gamma^* R^D}{\Delta\pi}] \right\},$$

with $1+g$, $\upsilon(\gamma^*)$, and γ^* calculated from the solution of (32), (35), and (36). The optimal value of μ is the one for which $d\mathfrak{W}/d\mu = 0$ is obtained.

[25] This derivation uses the standard results $\ln(1+x) \simeq x$, and $\sum_{n=0}^{\infty} n x^n = x/(x-1)^2$ when $|x| < 1$.

However, once again the resulting expression is too complex to allow an explicit analytical solution for μ.

Accordingly, a numerical evaluation in performed. This is done by noting that equations (32), (35) with $\gamma > \gamma_m$, and (36) consist of a recursive, static system. This system can be solved, subject to Assumptions 2 to 5, for values of $\mu \in (0,1)$. The optimal value of μ is thus the one that maximize either growth (equation (37)) or welfare (equation (41)).

To perform the simulations, the following initial values are used: $\Lambda = 0.96$, $\alpha = 0.3$ and $\alpha = 0.32$, $A = 5$, $\beta = 0.55$, $\pi^H = 0.95$, $\pi^L = 0.2$, $\Gamma = 0.1$, $\varepsilon = 0.05$, and $R^D = 1.02$. Thus, the base scenario assumes that the elasticity of the monitoring technology is fairly low. These values are fairly standard and reflect a high sensitivity of the private benefit to monitoring. For this set of values, the restrictions $A > 1/\alpha$, $\gamma/\Delta\pi < 1$, and Assumptions 2 to 5 are all satisfied.

The results are shown in Figure 4 for ε varying over the interval $(0.05, 0.2)$; for values higher than 0.22, one (or several) of the Assumptions 2 to 5 is (are) not satisfied because the resulting optimal values of γ or μ are too high. There are two results. First, regardless of the value of ε, the growth- and welfare-maximizing solutions are (almost) the same, because growth effects dominate changes in welfare. Thus, only a single optimal value is shown in the figure. Second, there is an inverse relationship between the optimal value of μ and ε. For $\varepsilon = 0.1$ and 0.15 for instance, the optimal values of μ are 0.143 and 0.093, respectively. Intuitively, a higher elasticity of the monitoring technology raises the optimal monitoring intensity, as implied by Proposition 5 and as shown in the figure. In turn, the increase in γ reduces the private benefit of the low-moral hazard project, $\upsilon(\gamma)$, which helps to promote for investment and growth. At the same time, however, higher monitoring costs reduce the bank's capacity to lend and increases the threshold level of wealth below which an entrepreneur cannot borrow, which tends to reduce investment and growth. To maximize growth and welfare, it is thus optimal to reduce μ (which weakens the borrowing constraint, as implied by Proposition 1) as ε increases. The idea that it is optimal to impose high required reserve ratios on banks when their ability to monitor borrowers is weak is consistent with the evidence which suggests that the use of reserve erquirements remains widespread in developing countries, compared to developed countries.

8 Conclusion

Using an OLG model with banking, this paper examined the growth and welfare effects of macroprudential regulation. In the model, the production of capital is subject to a dual moral hazard problem in the tradition of Holmström and Tirole (1997). This problem, which arises from the information asymmetry between borrowers and lenders, makes it impossible to write incentive contracts contingent on the production technology chosen by borrowers. To ensure prudent investment of the funds that they lend, banks must monitor borrowers. A monitoring bank is always successful in preventing entrepreneurs from using some of the proceeds of their loans to pay for unproductive activities. However, given that it is costly to monitor, banks must be provided with adequate incentives to do so. In addition, however, it was assumed that households cannot lend directly to producers and that the intensity of monitoring, which affects private returns from shirking, is endogenously determined.

The analysis focused on the impact of reserve requirements, a prudential instrument that has been used extensively in both low- and middle-income developing countries, and has recently been made part of the liquidity requirement guidelines under the new Basel arrangement. It was first shown that the direct effect on investment and economic growth of higher reserve requirements, aimed at reducing banks' capacity to lend, private sector leverage, and mitigating systemic financial risks, is negative when monitoring intensity is taken as given. However, the same policy has ambiguous effects on growth when monitoring intensity is endogenized because it also affects banks' incentives to monitor. Simple numerical experiments showed that, depending on the parameters characterizing the economy, macroprudential policy can be both growth- and welfare enhancing. Put differently, a trade-off does not necessarily exist between ensuring financial stability and promoting economic growth; both concerns can be balanced by setting the reserve requirement rate at his optimal value. The analysis also showed that there is an inverse relationship between the optimal reserve requirement rate and the efficiency of monitoring.

An important caveat to the analysis relates to the fact that the model did not account explicitly for the possibility that even though reserve requirements are set optimally, their level may be so high that they may foster disintermediation away from the formal banking system and toward less regulated channels. Even though the impact of this shift on investment and growth may be muted, it may exacerbate systemic risks. The possibility of leakages means therefore that financial sector supervision may also need to be strengthened, and the perimeter of regulation broadened, when aggressive macroprudential policy reforms are implemented. This is an important message for policymakers.

The analysis presented here can be extended in a number of directions. First, a better integration of short-run stabilization issues and longer-term considerations would be useful. In particular, if macroprudential regulation can mitigate business cycle volatility and reduce uncertainty about future economic conditions, it may have a permanent, positive effect on private investment and economic growth. These effects should be internalized in setting required reserve ratios. Second, the focus of the present analysis was on a single macroprudential instrument. In addition to the fact that reserve requirements have been used extensively (as noted earlier), this focus is justified by the fact that in a weak institutional environment (as is often the case in developing countries, especially the poorest ones), macroprudential rules aimed at preserving financial stability should not be overly complicated. However, it would be useful to consider other instruments in this setting, such as capital requirements and a leverage ratio, and solve *jointly* for the optimal levels of these instruments.

Finally, in the model, as in the original Holmström-Tirole tradition, monitoring reduces entrepreneurial moral hazard (which increases pledgeable income and facilitates access to credit) but it has not impact on profitability. However, as discussed in Favara (2012), monitoring may also affect the quality (or value) of the projects that entrepreneurs choose to implement, by interfering in the *ex ante* selection of investments. Integrating this mechanism would provide an additional channel through which macroprudential policy could affect growth and welfare.

References

Agénor, Pierre-Richard, and Luiz Pereira da Silva, "Reserve Requirements and Loan Loss Provisions as Countercyclical Macroprudential Instruments," unpublished, Inter-American Development Bank (November 2015).

Barnea, Emanuel, Yoram Landskroner, and Meir Sokoler, "Monetary Policy and Financial Stability in a Banking Economy: Transmission Mechanism and Policy Tradeoffs," *Journal of Financial Stability*, 18 (June 2015), 78-90.

Basel Committee on Banking Supervision, *Basel III: The Liquidity Coverage Ratio and Liquidity Risk Monitoring Tools*, Report No. 238 (January 2013).

Chakraborty, Shankha, and Tridip Ray, "Bank-Based versus Market-Based Financial Systems: A Growth-Theoretic Analysis," *Journal of Monetary Economics*, 53 (March 2006), 329-50.

De la Croix, David, and Philippe Michel, *A Theory of Economic Growth: Dynamics and Policy in Overlapping Generations*, Cambridge University Press (Cambridge: 2002).

Dewatripont, Mathias, and Jean Tirole, "Macroeconomic Shocks and Banking Regulation," *Journal of Money, Credit and Banking*, 44 (December 2012), 237-54.

Diamond, Douglas W., "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51 (July 1984), 393-414.

Favara, Giovanni, "Agency Problems and Endogenous Investment Fluctuations," *Review of Financial Studies*, 25 (July 2012), 2301-42.

Fowowe, Babajide, "Financial Liberalization in Sub-Saharan Africa: What do we Know?," *Journal of Economic Surveys*, 27 (March 2013), 1-37.

Haavio, Markus, Antti Ripatti, and Tuomas Takalo, "Macroeconomic Effects of Bank Recapitalizations," unpublished, Bank of Finland (February 2014)

Holmström, Bengt, and Jean Tirole, "Financial Intermediation, Loanable Funds, and the Real Sector," *Quarterly Journal of Economics*, 112 (September 1997), 663-91.

Misati, Roseline N., and Esman M. Nyamongo, "Financial Liberalization, Financial Fragility and Economic Growth in Sub-Saharan Africa," *Journal of Financial Stability*, 8 (September 2012), 150-60.

Van den Heuvel, Skander J., "The Welfare Cost of Bank Capital Requirements," *Journal of Monetary Economics*, 55 (March 2008), 298-320.

Figure 1

Determination of Optimal Bank Borrowing and Optimal Investment

Case I: Θ > max(1/Φ, 1 - 1/Φ)

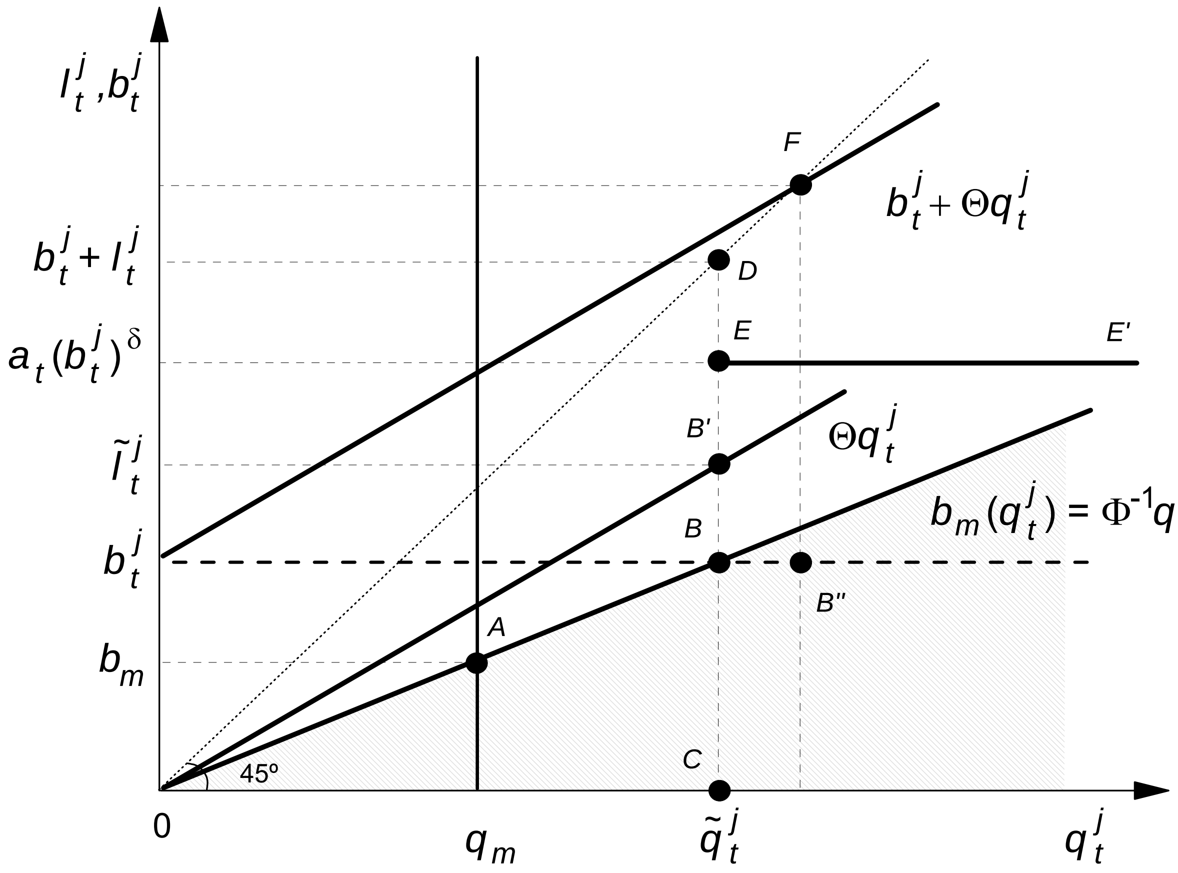


Figure 2
Determination of Optimal Bank Borrowing and Optimal Investment
Case II: $1/\Phi < \Theta < 1 - 1/\Phi$

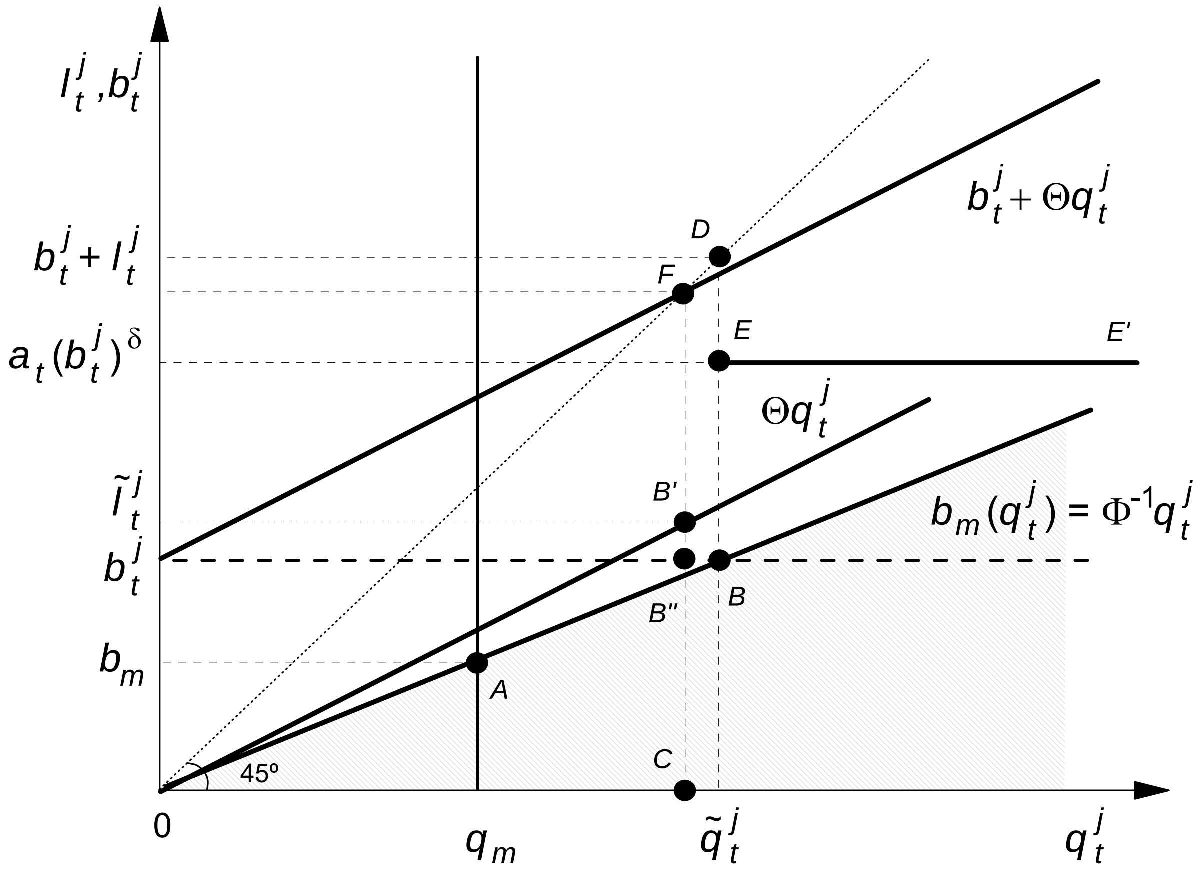


Figure 3
Determination of Optimal Bank Borrowing and Optimal Investment
Case III: $1/\Phi < \Theta = 1 - 1/\Phi$

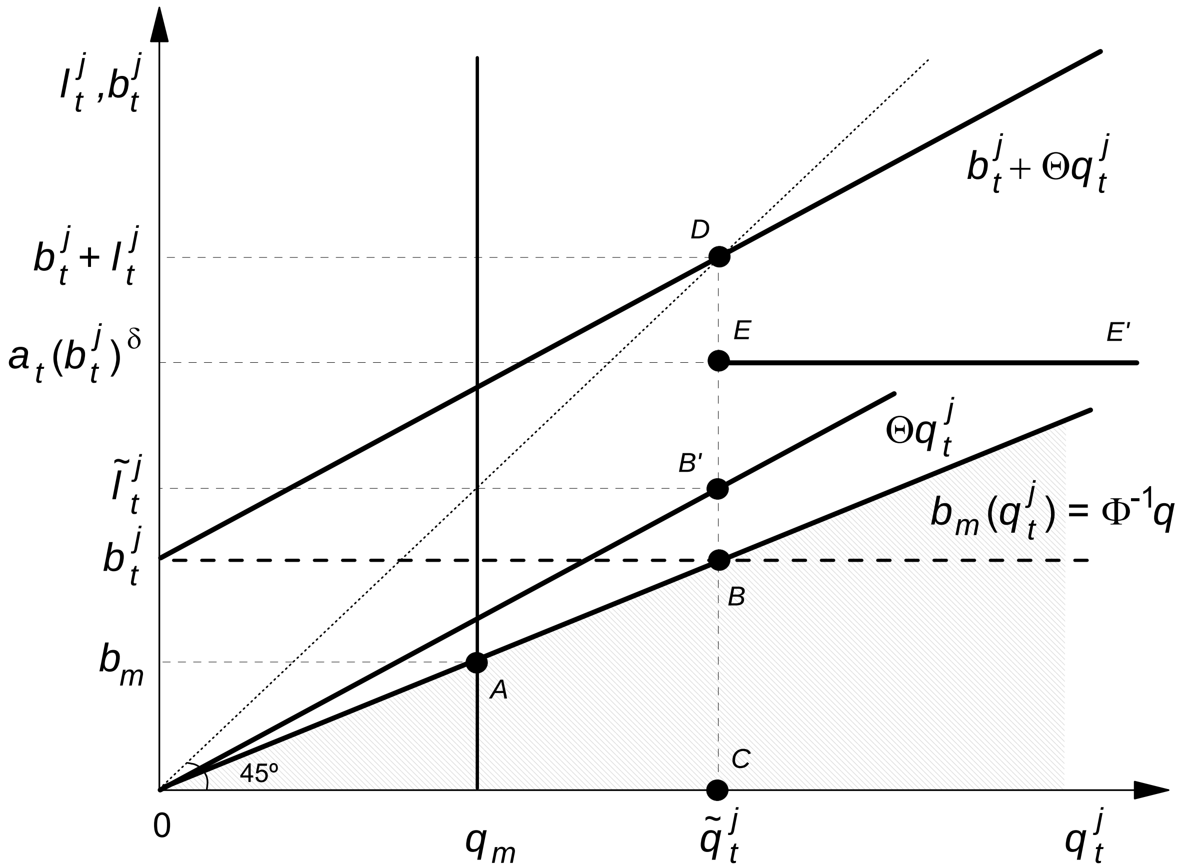
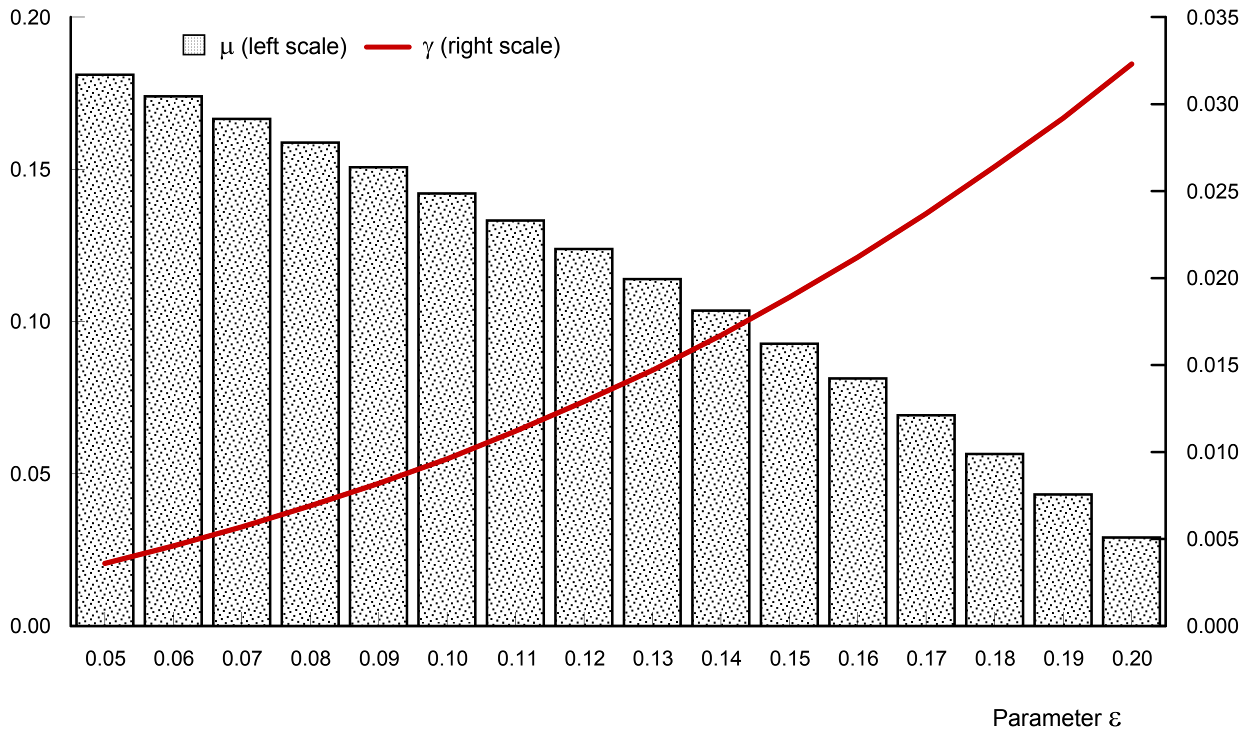


Figure 4
Optimal Monitoring Intensity and Reserve Requirement Rate
Monitoring Parameter ε varying between 0.05 and 0.2



Source: Author's calculations.