London School of Economics and Political Sciences

The Economics of Banking Crisis, Regulation and Deposit Insurance

Nancy Andrea Silva

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy, London, September 2008

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work.

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

Nancy A. Silva

Abstract

This thesis provides an economic analysis of banking crisis, regulation and deposit insurance.
Chapter 1 offers a critical review of the literature, identifying the main determinants of banking crises and their channels of contagion.
Chapter 2 studies the effectiveness of deposit insurance in containing panic runs when depositors have private information. The region of panic runs decreases with the size of the guarantee and the degree of supervisory involvement of the agency in charge of insurance. High levels of insurance tend to increase the equilibrium demand deposit contract and so the probability of runs, but supervision can also limit this effect. Therefore, a scheme with limited insurance and a high degree of supervisory involvement should be preferred.
Chapter 3 evaluates subordinated debt and disclosure requirements as instruments of market discipline. In the presence of deposit insurance, the former can be used to complement the latter, providing a new set of information which is useful to the regulator. If the subordinated bond has a long maturity, the probability of insolvency decreases for any level of noise in the information disclosed by the manager. If the bond can be rolled over, the quality of information improves substantially but the probability of insolvency increases slightly.
Chapter 4 studies the inter-temporal effects of capital adequacy requirements. A bank's risk-taking dynamic depends on critical thresholds of the capital requirements in each period. When the requirement binds in the initial period, risk can be reduced to the social optimum but at the cost of reducing financial intermediation as well. Moral hazard increases because, among the binding banks, the better capitalised ones raise relatively more insured deposits and take on relatively more risk. When the requirement binds in the interim period risk-taking increases, the more so the less capitalised is the bank, making smaller banks weaker.

Thesis Supervisor: Dr. Margaret Bray
Title: Reader in Economics

Acknowledgments

I am much indebted to my supervisor, Dr. Margaret Bray, for her invaluable help. Her questions and suggestions, always important and relevant, restored my hope in this project when I needed it the most.

I am also very grateful to the National Scholarship Program of the Government of Chile, for financial support during the initial years of research; and for financial assistance from the Department of Economics of the London School of Economics, which gave me the opportunity of combining my research with the wonderful experience of teaching. I must also thank my bosses and friends at the Central Bank of Chile, for their constant encouragement for me to finish.

Thanks to God and to my parents, Victor and Marta, for giving me more than life and loving me all the way. And overall, thanks to my husband, Eric, for his support, tenderness and inexhaustible patience and love. This work is dedicated to him.

A mi Eric

Contents

Introduction **12**

1 Crisis, Contagion and Regulation in the Banking Industry. A Review of the Literature **17**

1.1 Introduction . . . 17
1.2 The Nature of Banks . . . 19
1.3 Banking Crises . . . 23
1.3.1 Determinants of Banking Crises . . . 24
1.3.2 Twin Crises . . . 26
1.4 The Safety Net . . . 27
1.4.1 Prudential Regulation . . . 28
1.4.2 Deposit Insurance . . . 32
1.4.3 The Lender of Last Resort . . . 34
1.5 Systemic Risk Shifting or "Too Many To Fail" . . . 37
1.6 Financial Contagion . . . 39
1.6.1 Interbank Funding Markets . . . 40
1.6.2 The Payment System . . . 46
1.6.3 Derivatives Markets . . . 51
1.7 The International Lender of Last Resort . . . 54
1.8 Concluding Remarks . . . 56

2 Deposit Insurance and the Risk of Runs **59**

2.1 Introduction . . . 59
2.2 A Model of Information-based Bank Runs . . . 63
2.2.1 Inefficient Liquidation . . . 68
2.3 Deposit Insurance in a Model of Information-Based Bank Runs . . . 70
2.4 Equilibrium under a Narrow Mandate . . . 73
2.5 Equilibrium under a Broad Mandate . . . 76

2.5.1 DIC's Sub-game 77
2.5.2 Patient Depositors' Sub-game 84
2.6 Demand Deposit Contract with Insurance 86
2.6.1 Narrow Mandate DIC 86
2.6.2 Broad Mandate DIC 89
2.7 Robustness 90
2.7.1 Deposit Insurance Contract (DIC's Planning Period Sub-game) . 90
2.7.2 Credibility and Blanket Guarantees 91
2.7.3 Macroeconomic Shocks 92
2.7.4 Twin crisis 92
2.7.5 Moral Hazard 93
2.7.6 Closure Rule 94
2.7.7 Other Extensions 94
2.8 Concluding Remarks 95
2.9 Appendix 97
2.9.1 Proof of the existence of a unique monotone equilibrium in the case without insurance 97
2.9.2 Proof of proposition 5 98
2.9.3 Proof of proposition 8 99
2.9.4 Proof of proposition 21 100

3 Market Discipline and the Safety Net **107**
3.1 Introduction 107
3.2 The Model 110
3.2.1 Benchmark Model: Perfect Information 111
3.2.2 Imperfect Information 114
3.3 Disclosure 116
3.3.1 Closure Policy 116
3.4 Subordinated Debt 118
3.4.1 Zero Coupon Bond with Maturity at t=2 118
3.4.2 Roll-over at t=1 123
3.5 Discussion 127
3.5.1 Diversifying the bank's technology choice 127
3.5.2 Monitoring costs 128
3.5.3 Other extensions 128
3.6 Conclusions 129

3.7 Appendix 131
3.7.1 Density function of $\widetilde{\theta}$ 131
3.7.2 Interest rate with deposit insurance and no monitoring 132
3.7.3 Proof of proposition 28 132
3.7.4 $b_2(r, s_0)$ increasing on r 134
3.7.5 Proof of proposition 30 135
3.7.6 Proof of proposition 32 137

4 Capital Regulation and Bank Risk-Taking. Completing Blum's Picture 142

4.1 Introduction 142
4.2 The Model 146
4.2.1 First Best 147
4.2.2 Capital Requirements 148
4.3 Unregulated Solution 150
4.4 Regulated Solution 154
4.4.1 Initial Period Binding Requirement 154
4.4.2 Intermediate Period Binding Requirement 156
4.4.3 Binding Requirements in Both Periods 156
4.5 Policy Implications 158
4.6 Concluding Remarks 160
4.7 Appendix 163
4.7.1 The Regulated Bank 163
4.7.2 Comparative Statics 167

Bibliography 182

List of Figures

2-1 Number of total early withdrawals. 102

2-2 Equilibrium and solvency thresholds without insurance. 103

2-3 Extended form of the game in the interim period, when the DIC operates under (a) a narrow mandate and (b) a broad mandate. 103

2-4 Premium to wait to withdraw in the final period (a) without deposit insurance and (b) with deposit insurance. 104

2-5 Expected premium to wait to withdraw in the final period with deposit insurance. 104

2-6 Equilibrium and solvency thresholds with and without insurance. 105

2-7 Equilibrium threshold with insurance, as a function of c_1 and g. 105

2-8 Optimal value of c_1 without insurance and with $g = 1$. 106

2-9 Optimal value of c_1 with partial insurance. 106

3-1 Proof of proposition 28. 138

3-2 Proof of proposition 30. 138

3-3 Bond holders' payoffs. 139

3-4 Number of bond holders redeeming the bond in the interim period. . . . 139

3-5 Bond holders' conditional expected utility: redemption versus rollover. . 140

3-6 Example of two risky technologies with the same mean and different variances. 140

3-7 Density function of $\widetilde{\theta}$: case $\varepsilon_m < \frac{1}{2}$. 141

4-1 Timeline for the optimal decision of the regulated bank. 174

4-2 Example of the equilibrium relationship between risk-taking and the initial level of equity for the unregulated bank. 174

4-3 Example of binding capital requirement thresholds as a function of initial equity. 175

4-4 Risk chosen by the bank facing a binding capital requirement in the initial period. 175

4-5 Optimal value of the capital requirement in the initial period. 176
4-6 Better capitalised banks take on more risk when the capital requirement binds in the initial period only. 176
4-7 Example of the evolution of the equilibrium values of R and D_1, when only the intermediate period requirement binds. 177
4-8 Risk-taking in banks with different levels of equity, under an interim period binding capital requirement. 177
4-9 Binding capital requirements: (a) $\widetilde{c_0} < \widetilde{c_1}$, (b) $\widetilde{c_0} \geq \widetilde{c_1}$. 178
4-10 Possible paths of R^r as a function of c, with constant capital requirements and $\widetilde{c_0} < \widetilde{c_1}$. 178
4-11 Possible path of R^r as a function of c, with constant capital requirements and $\widetilde{c_0} > \widetilde{c_1}$. 179
4-12 Distribution of risks when capital requirements are used to force the smallest bank to the social optimum. 179
4-13 Distribution of risks when capital requirements are used to force the biggest bank to the social optimum. 180
4-14 Distribution of risks when capital regulation is designed to make the smaller banks converge on the unregulated risk choice of the biggest one. 180
4-15 Loss in financial intermediation under (a) a first period binding capital requirement and (b) a binding minimum equity requirement. 181

List of Tables

1.1 Estimation of the cost and length of financial crises. 23

2.1 Patient depositors' payoffs in the game without insurance.. 65

2.2 Patient depositors' payoffs in the game with insurance under a narrow mandate. 73

2.3 Closure and LoLR policies with a broad mandate and perfect information. 82

2.4 Patient depositors' payoffs in the game with insurance under a broad mandate. 85

3.1 Patient deposit holders' payoffs in the game with common knowledge and no insurance. 113

3.2 Patient depositors' payoffs in the game with common absence of knowledge and deposit insurance. 114

Introduction

Since the 1980s many emerging economies have experienced serious financial crises, which costs of recovery have in some cases been as high as 60 percent of GDP. It has been argued that many of theses crises had their roots in the weaknesses of the banking sector, which were exacerbated by insufficient or inefficient supervision and regulation.

Situations of financial distress have not been absent in developed economies either. Some examples are the Savings and Loan crisis in the USA, the financial crisis in Japan in the late 1980s and the ongoing subprime crisis, originated in the USA in 2007 but still having repercussions in other developed countries' financial centres. All these crises have involved hidden losses and increasing moral hazard in financial institutions.

As the soundness of the banking system has proved to have direct effects on macroeconomic stability, poverty and economic growth, the debate on optimal supervisory and regulatory frameworks for the banking industry has increased in the past years. This thesis provides an economic analysis of banking crises and contagion, giving special consideration to some policy issues contained in the design of the so-called "safety net", which includes capital regulation, subordinated debt, systems of deposit insurance and the lender of last resort.

Chapter 1 provides a critical review of the literature on banking, highlighting the sources of vulnerabilities in this market and the grounds for regulation, characterising events of failure and discussing the channels of contagion. Two main channels of contagion are identified. The *informational* channel refers to situations where signals of bad performance of one or more banks are interpreted by depositors as valid information on the solvency of banks with similar characteristics, or for the banking industry as a whole. Contagion through this channel might be controlled by regulatory measures to restore depositors' confidence, such as those included in the safety net. The *credit* channel, on the other hand, explains how the failure of a single bank can spread throughout the web of linkages developed in the interbank funding market, the payment system or derivatives markets. In this case, an analysis of the possible problems in the design of financial systems is developed in order to understand and prevent contagion.

In the same way that the propagation of a crisis among banks in a local economy serves as a justification for the intervention of a local lender of last resort, fears of contagion across economies in a globalised world have raised questions about the need for an international lender of last resort. The most relevant models in the growing literature about international financial architecture are also considered here.

Chapter 1 puts into a homogeneous context the two main models available in the literature, namely sunspot and rational expectation equilibrium models. Sunspot equilibrium models are popular because of their tractability but they have limited capability for use in policy design, as a relevant set of players (depositors) is neglected. Conversely, in rational expectation equilibrium models investors are able to evaluate the probability of a bank run before making their investment decision. Thus, even though they are more complex, they allow for a richer analysis.

In particular, models of deposit insurance have tended to concentrate on the effect that non-risk adjusted premiums have on moral hazard, more than on the direct implications that the introduction of these types of contracts might have on the equilibrium behaviour of depositors and banks. Chapter 2 returns to this problem, introducing deposit insurance in a model of information-based bank runs (Goldstein and Pauzner, 2000). An important feature of this model is that it has a unique equilibrium. This allows for a proper evaluation of the effects of insurance on the behaviour of the different players. I show that while consumers achieve better risk-sharing in a competitive banking system than in autarky, more solvent projects are liquidated. This is because uninsured depositors fail to coordinate in a subset of fundamentals and run on banks they know to be solvent. When introducing deposit insurance, I show that its effectiveness in eliminating panic runs varies with the size of coverage and the degree of supervisory involvement of the agency in charge of insurance. When the agency is not involved in the supervision of banks (narrow mandate), a deposit insurance contract preserving the monitoring role of depositors involves offering less than full protection. The trade-off is that panic runs cannot be completely eliminated with a partial guarantee, although it does reduce the region of fundamentals for which that occurs. With a high degree of supervisory involvement (broad mandate), I show that panic runs tend to disappear for any level of insurance as the regulator's signal becomes more precise, given that liquidity assistance is committed to solvent but illiquid banks. Moreover, it is cost efficient never to provide liquidity to insolvent banks. However, only extremely insolvent banks are closed, and those with enough funds to cover the payment of the final period guarantee are allowed to continue in operation. Therefore, the smaller the protection offered to depositors, the higher is forbearance. All these results hold, irrespective of the specific values of the guarantee, which in particular might imply that the social cost of deposit insurance is lower under a broad mandate.

Finally, I show that deposit insurance increases the equilibrium value of the demand deposit contract in the interim period and so the probability of runs, at least for high levels of the guarantee, but this effect seems also to be smaller under a broad mandate. Limited insurance can contain this externality to some extent, justifying the observed conduct of governments across the world in normal times.

Given the combination of these results and the empirical evidence provided by other authors, this chapter concludes that a preferable scheme would be one in which the agency in charge of insurance has more supervisory involvement (broad mandate) or a high degree of coordination with the supervisory authority.

Chapter 3 uses a variation on the three-period, imperfect information model introduced in Chapter 2, where uninformed insured depositors, informed uninsured sophisticated investors (equity holders and bond holders) and the regulator interact in the market, in order to study and compare the effects that disclosure requirements alone and in combination with a subordinated debt requirement, have on the equilibrium probability of insolvency of the bank and on the regulator's closure policy. I also study how the maturity of the subordinated bond modifies these results.

With a disclosure requirement, I prove that the regulator uses a fully financed closure rule that avoids inefficient survival but involves an inefficient liquidation of assets in a region of the bank's fundamentals. Although the probability of insolvency is higher than the first best, it converges on it as the manager's signal noise goes to zero. When, on top of the disclosure requirement, the bank is asked to issue a zero coupon bond with long maturity, the probability of insolvency is further reduced. The regulator's closure policy in this case forces fewer banks into liquidation, given the subsidy from bond holders to deposit holders when the bank becomes fundamentally unprofitable. Finally, when debt can be rolled over in the interim period and so bond holders do monitor the bank, this period sub-game equilibrium provides very useful information that, in some cases, can completely eliminate inefficient liquidation of assets because the regulator can perfectly observe the bank's fundamentals. However, this will come at the cost of a higher probability of insolvency than in the non-rollover case.

The results in this chapter allow us to conclude that a subordinated debt requirement, when issued with a long maturity, is able to reduce the probability of insolvency for any size of noise. Indeed, intuition suggests that when the manager's signal noise is expected to be high, an appropriately high subordinated debt requirement could restore the first best. On the other hand, a subordinated bond issued with a short maturity can substantially improve on the quality of information (by reducing the noise). Therefore, a subordinated debt requirement can be used to complement disclosure requirements, providing a new set of information which is useful to the regulator.

Capital regulation is a clear counterpart of deposits insurance, as banks receiving a subsidy from the financial authority must agree to be regulated. Going back to sunspot equilibrium models and assuming full deposit insurance, to rule-out panic runs, Chapter 4 studies the intertemporal effects that capital regulation has on curbing a bank risk-taking. In order to do that, in this chapter I build on Blum's (1999) seminal paper in order to obtain some important lessons neglected in his original work. The reason for changing the Diamond and Dybig (1983) setup used in the previous chapters is that Blum's model allows for the study of asset substitution in a much simpler fashion: the bank manager can choose among different returns of the portfolio, which have associated a unique probability of default. Moreover, common knowledge and the assumption of universal risk neutrality allow for the risk choice of the bank to be completely isolated from any other decision in the game.

In this setup, I show that for unregulated banks risk-taking is decreasing on the level of initial equity and converges on the social optimum when equity is sufficiently high. When introducing capital requirements, I show there exist critical threshold values in each period for which regulation starts binding. When the requirement binds in the initial period only, risk can be reduced below the unregulated solution – even to the social optimum for sufficiently tight regulation – but fewer deposits are taken, which reduces financial intermediation. Moreover, capital requirements are not sufficient to control moral hazard because, among the binding banks, the better capitalised ones raise relatively more insured deposits and take on relatively more risk. When the requirement binds only in the interim period, bank risk-taking increases, most likely above the unregulated solution for all values of the requirement. In that case, risk would be decreasing on equity, making risk-taking even more aggressive for poorly capitalised banks. Therefore, interim period binding capital requirements will not only worsen the risk choice of banks but make smaller banks weaker. When a constant capital requirement binds in both periods, the tighter the regulation the fewer deposits are taken from the public, though better capitalised banks raise relatively more deposits. The dynamic of risk in this case depends strongly on the relationship between the threshold values of the requirement in each period, which in turn depends on the level of initial equity of the bank.

These results are of extreme interest in the current situation, where the crisis in the international financial markets has called for tighter regulation in the banking industry. The discussion in this chapter shows that an anticipated rise in capital requirements in the next period, combined with a shock reducing the expected return of the risky technology, might increase the likelihood of a more aggressive risk-taking response by banks. In the light of these results, any amendment to the current regulatory framework should be carefully analysed.

Chapter 1

Crisis, Contagion and Regulation in the Banking Industry. A Review of the Literature

1.1 Introduction

Episodes of financial distress have been neither uncommon nor isolated during the past four decades, when more than 100 serious banking crises have occurred in different economies, especially in emerging markets.[1] The crises in Latin America in the 1980s, in the South Asian economies in the late 1980s and 1990s and, more recently, in Turkey, Argentina and Uruguay are a few among many examples.

Financial crises can occur in the form of stock market crashes, currency crises, external debt defaults or banking crises. Different types of crises can happen at the same time or precede new turbulences. A stock market crash can affect the solvency of an otherwise sound banking system; severe bank failures can precipitate a flight to currency, eventually forcing devaluation or the abandonment of a currency system, etc.

Indeed, banking and currency crises have tended to occur simultaneously in emerging markets.[2] In Argentina (2001-02), for example, the collapse of the currency peg regime and the subsequent devaluation of the Argentine peso led many banks into insolvency. In Ecuador (2000), the weakness of the banking system and the failure of the authorities to act promptly led to a flight to currency that culminated in a strong devaluation, followed by an abandonment of the currency (dollarisation). Empirical evidence, however,

[1] According to IMF estimates, between 1980 and 1996 alone, nearly 130 countries experienced serious banking crises (Goodhart et al, 1997).

[2] This has been known as a *twin crisis*, a concept first introduced by Kaminsky and Reinhart (1996).

suggests that the frequency of banking crises has increased over time while the frequency of other types of crises has remained relatively constant (Glick and Hutchinson, 1999).

The costs of the resolution and recovery from crises have been high, especially when the banking sector is affected, varying from 2-3 percent of GDP in the Savings and Loan crisis in the USA (1986-1995) to as much as 50-58 percent of GDP in Kuwait (1990-91) and Indonesia (1997).[3] These estimates are just one part of the real net costs to society, as general activity substantially declines when the payment system is suspended, international investment falls and jobs are destroyed.

These experiences have underlined the importance of a sound banking system because of its direct effect on macroeconomic stability, poverty and economic growth. Indeed, a robust banking sector can help a country which is facing difficulties to avoid major problems, as in Chile and Argentina after the Tequila crisis; or to recover faster if it is affected, as in Singapore, Taiwan and Hong Kong after the Asian crisis. Conversely, a weak or unsound banking sector may generate or exacerbate a crisis, as happened in Thailand in 1997 or in Turkey in 2000.

For efficiency reasons, insolvent banks should be allowed to fail, in order to contain moral hazard and exercise effective market discipline. The objective of regulators should be to "*minimize the cost of a crisis rather than to avoid it*" (Freixas and Rochet, 1997), which in particular implies implementing appropriate policies to control the failure of one single institution from spreading to the rest of the financial system.

The proposal to have internationally accepted standards promoting a safe banking system is part of the regulatory framework contained in the Basel Agreements. The traditional justification for regulation has been based on the protection of uninformed dispersed depositors against the risk of failure and the guarantee of a safe and efficient payment system, arguments that involve a strong commitment to financial stability. However, as with many other topics in economics, there is no clear consensus on the definitions of financial contagion or financial stability, let alone when this is narrowed down to banking. What is clear, though, is that by reducing the risk of contagion a higher level of stability can be expected.

This chapter surveys the literature on banking and is organised as follows. Section 1.2 describes the nature of banks, presenting theories that explain both their existence and the reasons for their failure. Section 1.3 provides evidence on the characteristics and determinants of banking crises. Section 1.4 introduces the main aspects of the safety net, including prudential regulation, deposit insurance and the lender of last

[3] Cooper and Ross (1998), Curry and Shibut (2000), Calvo (2000).

resort. Section 1.5 discusses the systemic risk shifting problem, which arises when banks invest in correlated portfolios, leading to a "too many to fail" phenomenon. Section 1.6 and the subsections therein describe contagion through the credit channel, which includes the interbank funding markets, the payment system and instruments for credit transfer such as financial derivatives. Section 1.7 presents a discussion of the benefits and costs associated with the existence of an international lender of last resort. Concluding remarks are given in section 1.8.

1.2 The Nature of Banks

The origin of banks goes back to the ancient world and probably predated the invention of money. The Hammurabi's Code, in the eighteenth century BC, already included some laws regulating banking related activities. Although in more developed markets non-bank financial institutions and capital markets have progressively come to overshadow banks as the main providers of liquidity, in many emerging economies banks remain the main or only financing source to households and the corporate sector, as alternatives remain underdeveloped.[4]

Banks have historically been considered *special institutions*, not only because they are financed by a large number of atomised, mostly uninformed depositors but also because they provide the essential service of transforming the maturity of short-term liabilities into long-term assets, allowing for a smooth functioning of the payment system.

The first theories explaining the existence of banks as financial intermediaries emphasised their role in converting securities issued by firms into securities demanded by investors. The existence of transaction costs gives them an advantage over individual investors and, where indivisibilities in investment are present, they can also achieve a better diversification of risk (Gurley and Shaw, 1960; Fama, 1980). Other models explain the emergence of banks as a consequence of the existence of imperfect information in financial markets. Banks can act as providers of liquidity insurance and monitoring services for depositors, which are subject to asymmetric information due to idiosyncratic liquidity shocks on consumption (Bryant, 1980; Diamond and Dybvig, 1983). They can also act as providers of liquidity insurance to investors and firms facing liquidity needs with different maturities (Diamond and Rajan, 1998). Finally, banks can avoid the duplication of monitoring costs (Diamond, 1984) as investors delegate their monito-

[4] According to figures from the People's Bank of China, as of 2006, bank loans represented 99 percent of corporate funding in that country. For a study of the differences between bank and market based systems, see Allen and Gale (1995, 2001). For a model explaining the simultaneous existence of both markets, see Bolton and Freixas (1997).

ring power to them (Campbell and Kracaw, 1980), possibly because they are better at screening potential borrowers (Grossman and Stiglitz, 1980) or at supervising the use of funds (Boot and Thakor, 1993), making resources accessible to firms at a lower price.

The prominent role of banks in financial intermediation, as providers of information and liquidity insurance, exposes them to the continuous threat of runs, given their ability to transform short-term riskless liabilities into long-term risky loans, which is known as the *fractional reserve system.*

Models of bank failure can be broadly classified into sunspot equilibrium models, which are analytically more tractable, and models of equilibrium in rational expectations, which are more complex but which provide a richer framework for policy analysis. Amongst the first class of models, the seminal work of Diamond and Dybvig (1983) provides a very useful framework for understanding how banks' demand deposit contracts can provide allocations which are Pareto superior to those offered by exchange markets, while making them susceptible to runs. Their model considers an economy with three periods, a short-term storage technology in which net return is normalised to 1, a partially illiquid long-term technology returning $r \leq 1$ if liquidated at $t = 1$ and $R > 1$ if liquidated at $t = 2$, and a continuum of mass one consumers, all endowed with one unit of an homogeneous good at $t = 0$, who are uncertain about their future time of consumption. This is because an exogenous fraction π of consumers are impatient, meaning that they must consume in the interim period. Consumers' types are realised in the interim period, this being private information.

If π were known, the optimal allocation could always be achieved in a decentralised market, with banks offering the same contract as a social planner maximising the expected utility of a representative consumer. When types are not verifiable, however, two equilibrium outcomes are possible: a Pareto efficient equilibrium, where all patient consumers wait to withdraw in the final period, and an inefficient equilibrium, where if some patient depositors decided to withdraw early, the rest – because of a first-come-first-served assumption – would fear being left with nothing and would precipitate a run.

Bankruptcy in these types of models arises as a sunspot phenomenon, which is not related to a bank's fundamentals but to exogenous stochastic liquidity shocks on demand. Indeed, as neither depositors nor banks consider the probability of runs when making their investment decisions or computing the optimal level of inter-temporal insurance, runs are not equilibriums in rational expectations, making policy evaluation extremely difficult. A possible solution might be to introduce incentive compatibility constraints, inducing patient depositors to wait until the final period. Von Thadden (1998) explores this idea in a continuous time extension of Diamond and Dybvig (1983), but he concludes that such constraints always bind in equilibrium, severely limiting the provision of liquidity insurance.

Because both investment technologies are risk-free in Diamond and Dybvig (1983), runs cannot be regarded as a desirable equilibrium outcome. However, most banking problems in recent years have been associated with the deterioration of asset quality. Gorton (1988) provides empirical evidence suggesting that even the so-called panic runs in the USA at the beginning of the 20th century were related to fundamentals, as they followed when a leading economic indicator, the liabilities of failed businesses, reached a certain threshold.

In an economy à la Diamond and Dybvig (1983), Allen and Gale (1998) use the idea that an economic downturn may reduce the value of banks' assets, increasing their probability of insolvency. When depositors anticipate this, they rationally precipitate a run. They model a decline in aggregate output as a shock to assets' returns, perfectly correlated across banks. The first-come-first-served assumption is dropped and the long-term technology is assumed to be completely illiquid ($r = 0$), so that something is always left for late withdrawal. The first best allocation can be achieved when depositors are offered a deposit contract contingent on the return of the risky asset but not on their type. If patient depositors observe a low signal about the return of the risky asset, they will precipitate a run.

Other models explaining runs as a rational response to negative private information include Postlewaite and Vives (1987), where a fraction of depositors observe signals on the probability of a run, and Cooper and Ross (1998), where depositors consider the probability of runs before making their deposits. In both models, however, the probability of runs is an exogenous parameter, which therefore is unaffected by the form of the deposit contract.

Some models have allowed for the probability of runs being endogenised but have excluded panic based runs as an equilibrium phenomenon (Jacklin and Bhattacharya, 1988; Alonso, 1996). When introducing asymmetric information, panic runs can be recovered as an equilibrium solution, either because private signals are misinterpreted (Chari and Jagannathan, 1988) or because signals are noisy and depositors fail to co-ordinate on a small region of the fundamentals (Goldstein and Pauzner, 2000). In an economy à la Diamond and Dybvig (1983), Chari and Jagannathan (1988) introduce a long-term technology with stochastic returns (with probability $p > 0$ the return equals

$R > 1$ and with complementary probability equals zero) and asymmetric information. A fraction β of all patient depositors are informed and receive a credible signal about the return of the risky asset. Uninformed depositors can only observe the size of the queue to withdraw, sometimes confusing states of high liquidity demand with insolvency. Here the proportion of impatient consumers, π, is a random variable that can take values π_L or π_H, where $\pi_H = \pi_L + \beta(1 - \pi_L)$. In a state of low liquidity demand, if the signal on future returns is high, there are no runs. If, instead, the signal is low, informed patient consumers withdraw. Uninformed patient depositors, observing π_H, decide to withdraw as well and an efficient run occurs. However, in a state of high liquidity demand, an efficient run occurs only when a low signal on future returns is received in the interim period. If the signal is high, uninformed depositors still observe a long line and prefer to withdraw, forcing informed consumers to run on a solvent bank if what is left for the second period is insufficient to guarantee at least the same level of consumption they would get by withdrawing early.

Goldstein and Pauzner (2000) modify Diamond and Dybvig's (1983) model using global games techniques and show that a unique equilibrium in rational expectations arises where both the probability of bank runs and the optimal demand deposit contract are endogenous. This time the long-term technology's return, $R = R(\theta)$, depends continuously and increasingly on a uniform random variable, θ, with support in $[0, 1]$, which represents the underlying fundamentals of the project financed by the bank. At the beginning of period 1, each depositor receives a private, non-verifiable signal about the true value of the fundamentals, $\theta_i = \theta + \varepsilon_i$, where ε_i are i.i.d. random variables, uniformly distributed in the interval $[-\varepsilon, \varepsilon]$. Notice that, in this model, all patient depositors receive equally informative signals, though they are not common knowledge. After receiving a signal θ_i, depositor i knows the true value of θ lies in the interval $[\theta_i - \varepsilon, \theta_i + \varepsilon]$. With incomplete information, depositors must condition their beliefs upon their private signals, which are positively correlated with the signals of others. The effect of private signals is then twofold. The higher the signal, the higher the posterior distribution attributed to the realisation of $R(\theta)$, which increases a patient depositor's willingness to wait. Also, as signals are positively correlated, a high value signal makes a depositor believe that other players are receiving high value signals as well, assigning a lower probability to the event that other patient depositors will run on the bank. Goldstein and Pauzner (2000) prove that a unique equilibrium *"on switching strategies"* exists, that is, an equilibrium in monotone strategies with threshold θ^*, such that if $\theta_i < \theta^*$ a patient depositor withdraws and remains if $\theta_i > \theta^*$. This equilibrium turns out to be the only equilibrium of the game. The equilibrium threshold θ^* is increasing on c_1, a measure of intertemporal insurance, as explained in Chapter 2.

A Remark about Demand Deposit Contracts

"Bank runs are an inevitable consequence of the standard [demand] deposit contract in a world with aggregate uncertainty about assets return. But these contracts allow the banking system to share the risk among depositors". Gale (2000)

The natural question is whether a different type of contract could provide a more stable banking system. Alonso (1996) demonstrates that contracts where runs occur may be better than those which ensure no runs because they improve on risk-sharing. Jacklin and Bhattacharya (1988) compare equity contracts, whose prices are fully revealing, with demand deposit contracts with a suspension of convertibility and conclude that, when the dispersion of returns is low, demand deposit contracts are preferred. Gorton and Pennacchi (1990), in a similar exercise, allow for equity prices not to be fully revealing and conclude that less informed traders prefer safer deposits to equity contracts.

1.3 Banking Crises

The soundness of the banking system can have a direct effect on macroeconomic stability, poverty and economic growth.[5] Banking intermediation promotes the growth of an economy by providing an efficient allocation of savings to more productive technologies and by promoting capital accumulation through the attraction of local and foreign capital (Beck, Levine and Loayza, 1999; Beck and Levine, 2001). Conversely, banking crises can cause severe damage, as they tend to be long and very costly, both in terms of the direct fiscal costs involved in their resolution and in terms of net output losses (table 1.1).

	IMF (1998)		Bordo et al. (2000)	
Crisis type	Cost	Length	Cost	Length
Currency	4.3	1.6	5.9	2.1
Banking	11.6	3.1	6.2	2.6
Twin	14.4	3.2	18.6	3.8

TABLE 1.1: Estimation of the cost and length of financial crises. Average costs measured as a percentage of GDP. Average length in years. The IMF (1998) study includes 53 countries over the period 1975-1997. Bordo et al. (2000) includes 56 countries over the period 1973-1997.

[5] See Levine (1997) for a survey.

One of the main roles performed by banks is the monitoring of investment projects, through which they learn important information about the payment ability of their borrowers. This lender-borrower relationship is valuable to both banks and customers because it enables reputation building (Diamond, 1991), reduces agency problems (Rajan, 1992) and allows for more flexible contracts (Von Thadden, 1995). When a bank fails, however, this relationship is destroyed and valuable information is lost.

Indeed, as it is difficult and costly in practice for firms to obtain quick access to new credit, especially when bank failures are systemic, a financial crisis can propagate to the real sector. Bernanke (1983) argues that, during the Great Depression, the cut-off of financing to borrowers strongly dependent on banking lending exacerbated the decline in the US economy. Slovin et al. (1993) study the effect of the collapse of Continental Illinois, in 1983, on the stock price of 29 publicly traded firms that had maintained a lending relationship with the bank until its *de facto* failure. They interpret the average 4 percent decline in their market value – prior to the announcement by the Federal Deposit Insurance Corporation (FDIC) that the bank would be bailed out – as evidence that the costs associated with the loss of lender-borrower relationships can be substantial. They argue that financial markets may effectively punish firms related to distressed banks, if failure is interpreted as an indicator of bad monitoring or poor choice of loan customers. Kang and Stulz (2000) reach similar conclusions when studying the deflation of the Japanese stock market in the period 1990-93. In particular, they find that the performance of firms with close banking relationships was lower when their main banks were also experiencing financial problems.

1.3.1 Determinants of Banking Crises

Among the typical causes of banking crises are macroeconomic instability, deficient supervision, inadequate management and operational risks (Latter, 1997). All these elements may appear to be strongly correlated in practice, as in the processes of financial liberalisation in the 1980s. Inefficient regulation and supervision, implicit bailout guarantees and lack of competition, led to weak management and moral hazard. Financial liberalisation allowed for foreign over-borrowing by banks, high real exchange rate appreciation, sudden rises in real interest rates and large fiscal and current account deficits. The result was a long list of banking crises in emerging markets during that decade.

Macroeconomic shocks can affect the stability of the banking system due to the deterioration of asset quality. The collapse of asset prices, a sharp increase in interest rates, a fall in exchange rates, a slowdown in the pace of general inflation, or the onset of a recession, can all create pressure on sectors to which banks are exposed, forcing otherwise profitable projects into insolvency. Macroeconomic shocks, however, are not always an exogenous phenomenon but can also be generated by mistaken macroeconomic policies, agency problems or asset mismanagement. A remarkable example of a macroeconomic shock is asset price bubbles. Allen and Gale (2000a) show that asset prices are related to the available and expected amount of credit in an economy. When credit expands, investors borrow from banks to finance the purchasing of pre-existing assets. Risk shifting, resulting from the inability of lenders to observe how borrowers invest, may increase the prices of these assets, creating a bubble. There are many examples of asset price bubbles, with repercussions of varied intensity through the banking sector. Some of these are the South Sea Bubble in 1722, the Wall Street Market Crash in 1929, the real estate bubble in Japan in the 1990s, the Dotcom Bubble in 2001 and the subprime crisis in the USA in 2007.

Prudential supervision is intended to prevent and cope with financial crises. Empirical evidence indicates that it was precisely the lack of effective supervision in the financial liberalisation process conducted during the 1980s, notably in the former Soviet Union countries, which led to many banking and currency crises (Latter, 1997; Stiglitz, 2002). In fact, *"government interventions can be destabilizing as well as stabilizing"* (Calomiris, 1997). A regulation so strong that it eliminates the probability of failure could constrain the competitiveness and efficient operation of the banking system. On the other hand, lenient supervisory policies, such as bailouts without restrictions, tend to increase moral hazard and so encourage risk-taking. Government interference, as with the obligation to fund government deficits on non-market terms or the pressure to lend to particular customers, could also precipitate liquidity or solvency crises.

Imprudent managerial decisions can clearly affect the solvency of a bank and the stability of the banking system as a whole. Taking on too much risk in a rush to expand or to *"gamble for resurrection"*, poor credit assessment and adverse credit selection, high credit exposures, concentration of lending and connected lending, taking on new areas without the needed expertise (e.g. derivatives trading), unauthorised trading or position-taking associated with corruption and fraud, and failures associated with human-capital (non-enforceable managerial effort, high staff turnover, etc.) all have a direct effect on the solvency and profitability of a financial institution. Some models have been developed to study mechanisms inducing proper behaviour by managers, for example via the transfer of control to more intrusive supervisory institutions, as in Dewatripont and Tirole (1993) and Repullo (2000). At a more practical level, the IMF's code of best practice establishes that a bank's owners and managers should be subject to *"fit and proper"* tests before being allowed to operate a banking licence (Hoelscher and Quintyn, 2003).

1.3.2 Twin Crises

During the 1980s and 1990s, a high correlation was observed between exchange rate collapses and banking crises. Banking and currency crises appeared to arise virtually at the same time in Latin America in the early and mid-1980s, in Scandinavia in the early 1990s, in Mexico in 1995, and in Thailand, Indonesia, Malaysia and Korea in 1997-98. The term *"twin crisis"*, understood as a banking crisis accompanied by a currency crisis (or vice versa) in a period of 24 months, was originally coined by Kaminsky and Reinhart (1996) when studying the experience of the financial liberalisation processes conducted in the 1980s, where this phenomenon became more frequent.[6]

In an empirical study, Gonzalez-Hermosillo (1996) shows that twin crises are more common in developing than in industrial countries, and particularly so in financially liberalised emerging economies. In a database of banking and currency crises in 90 countries over the period 1975-1997, using a signal-to-noise ratio methodology, she concludes that, in a financial system where depositors may substitute domestic assets for foreign ones, a banking crisis may increase the relative risk of holding bank deposits, leading to a flight to currency. Indeed, she shows that banking crises provide a good leading indicator of currency crises (as in Kaminsky and Reinhart, 1996).

Other models supporting the connection between a weak banking sector and the emergence of currency crises include Obstfeld (1994), where rational speculators anticipate that policymakers will commit themselves to maintain inflation targets rather than exchange rate stability in order to avoid bankruptcies; Velasco (1987) and Calvo (1998), where a currency attack is prompted by an increase in liquidity inconsistent with a stable exchange rate, resulting from a government bailout of the banking system; and Miller (1999), who explicitly considers currency devaluation as an option for a government confronted with bank runs and a fixed exchange rate regime.

Theories supporting this same connection but in the opposite direction – that is, of banking crises caused by currency crises – include Miller (1996), who shows that, when deposits are used to speculate in the foreign exchange market and banks are 'loaned up', the attack on the currency can lead to a banking crisis; and Rojas-Suarez and

[6] It is argued that in the 1970's, when financial systems were highly regulated, currency crises were rarely accompanied by banking crises (Stiglitz, 2002).

Weisbrod (1995), where a currency crisis can directly weaken a bank's balance sheet via depreciation, or indirectly, by forcing the central bank to raise interest rates in an attempt to defend the currency.

Finally, some authors explain currency and banking crises as a simultaneous phenomenon. In Chang and Velasco (1998), both crises are modelled in a Diamond and Dyvbig (1983) set up and, therefore, are explained as sunspot equilibriums. They regard banking and currency crises as simultaneous manifestations of a phenomenon of *international illiquidity*. Goldstein (2005) shows how strategic complementarities between depositors and speculators generate a vicious cycle between banking and currency crises, which magnifies with the correlation between the two crises.

1.4 The Safety Net

The term *"safety net"* refers to the set of instruments that are available to central banks to insure the healthy functioning of the banking sector and a smooth working payment system. It includes tools of prudential regulation and supervision (deposit interest rate ceilings, portfolio restrictions such as reserve requirements and narrow banking, capital requirements, regulatory monitoring, and restrictions on entry, branching, network and merger), deposit insurance and mechanisms for emergency liquidity assistance (lender of last resort (LoLR) and open market operations (OMO)).

Before the 1930s, the involvement of official institutions in issues of financial stability was modest but, after the Great Depression, several economies started to develop systems of protection for depositors and banks. In the USA, in 1933, the Reconstruction Finance Corporation was allowed not only to offer discount window facilities to banks with good collateral but also to purchase the preferred stock of banks and other firms, in order to reduce the probability of failure of distressed institutions. In the same year, the Glass-Steagall Act set up the FDIC for insuring bank deposits.

The safety net that then emerged remained largely untested during decades characterised by tranquil market conditions. However, after the collapse of the Bretton Woods system in the 1970s, a series of banking crises started to be repeated throughout the world, marked by abuses of the safety net by banks and lenient regulation and supervision by governments. The Savings and Loan crisis in the USA (1986-1995) *"was of central importance in galvanizing the debate over safety net reform. It provided clear...evidence of ways in which government protection of financial institutions could be abused...[as] much of the loss experienced within the industry resulted from legal, voluntary risk-taking and fraud (rather than exogenous shocks)"* (Calomiris, 1997).

As a result of this, the perception of the success of the safety net changed substantially. Supervisory intervention has since being criticised for being costly and slow, while regulation has been blamed for introducing distortions into the market, which induce bank managers to take on excessive risk (moral hazard). The cycle seems to spiral, as the distortions created by regulation call for additional regulation, bringing about changes in the informational structure of markets, which make the achievement of Pareto efficient outcomes more difficult (Stiglitz, 1994).

The following years witnessed important steps forward in the task of limiting protection and enforcing market discipline, of which the 1988 Basel Capital Accord and its amendments contained in Basel II; the 1989 Financial Institution Reform, Recovery and Enforcement Act; the 1991 FDICIA[7] in the USA, and similar initiatives by the European Commission, are some examples.

1.4.1 Prudential Regulation

The Basel Committee was established at the end of 1974 by the central bank governors of the Group of Ten countries, with the aim of gathering central bankers and bank supervisors and regulators to discuss issues related to prudential banking supervision. The 1988 Basel Capital Accord introduced a minimum risk-weighted capital to asset ratio (the *Cooke ratio*) of 8 percent, which in the earliest version only considered credit risk. New amendments to include market and operational risks, as well as credit risk, in the computation of weights are part of a more comprehensive revision of the regulatory policy, in the context of Basel II, whose first agreed document was published in 2004 but which only started to be adopted by all credit institutions in the European Union in 2008 and is yet to be implemented in the United States.

Although these guidelines were designed for internationally active banks, they have been widely adopted by banks in developing economies. In emerging economies, efficient supervision has usually been constrained by the limitations in technology, resources and the legal environment for applying effective sanctions to banks that do not comply with regulatory requirements (Laporta et al, 1998). Evidence that these measures are in fact the best ones for countries with different institutional and political frameworks is still limited.[8]

The objective of capital regulation is to strengthen the soundness and stability of

[7] FIDIC improvement act, establishing prompt corrective actions to enforce new standards.

[8] See Barth, J.; G. Jr. Caprio and R. Levine (2002) in *"Bank Regulation and Supervision: What Works Best?"*. Mimeo, and Griffith-Jones, S. and S. Spratt (2001), *"Will the proposed new Basel Capital Accord have a net negative effect on developing economies?"*. Mimeo, Institute of Development Studies, Brighton.

the international banking system and to reduce competitive inequalities across markets. Kahane (1977), Kareken and Wallace (1978) and Sharpe (1978) justify the use of capital requirements to control the solvency of banks whose asset allocations are distorted by deposit insurance, using models which assume that markets are complete and that the economy lasts for just one period. A static framework, though, may fail to capture important intertemporal effects that capital requirements have on the behaviour of banks. One of the first theoretical models for studying these effects was provided by Blum (1999). In a three periods model he studies the incentives for asset substitution coming from the reduction in expected profits imposed by the requirement. In the initial period the manager has two investment options, a safe asset whose rate of return is normalised to 1 and a higher yielding but risky technology. In the interim period, investment technologies are risk free. He shows that when capital regulation just becomes binding in the initial period, a tighter requirement reduces risk-taking. However, when capital requirements become binding in the interim period, tightening the requirement raises the level of risk. This is because the amount of equity in the final period is endogenous; therefore, when raising equity is excessively costly, banks increase their risk in the initial period in order to have a higher amount of final equity.

Under an incomplete market approach, Koehn and Santomero (1980) and Kim and Santomero (1988) use a mean variance portfolio model with fixed liabilities to prove that, in the absence of a solvency requirement and abstracting from the limited liability clause, the probability of failure decreases on the capital ratio, which is independent of the (non-negative) weights used in the computation of the ratio. However, the introduction of capital requirements changes the asset allocation of banks so that, while the share of the risky portfolio decreases because the bank shifts to those assets within a lower weight category, its composition is distorted in the direction of more risk inside the chosen weight category, increasing the probability of failure (Rochet, 1992). An example of why this might be a problem was the collapse of Argentina's banking system in 2001-02, which was closely related to their over-exposure to defaulted government bonds, with zero weight.[9] Thakor (1996) also shows how a bad selection of risk weights can have a negative impact on the real sector, given that asset allocation can be distorted by the difference between market and regulatory assessments of risks. As a way of correcting

[9]During 2001, banks rolled over maturing debt and increased their holdings of public debt as well as their credit to the public sector. The share of public sector assets increased from 17.4 percent of the total in December 2000 to almost 20 percent in September 2001. Public sector exposure was widespread in the banking system, being particularly high in large and primarily publicly-owned banks. For further references, see B. Eichengreen (2001), "*Crisis prevention and management: any new lessons from Argentina and Turkey?*", The World Bank.

this problem, Koehn and Santomero (1980), Kim and Santomero (1988) and Rochet (1999) propose the introduction of regulatory weights proportional to market based measures of the risk of the assets. This proposal has indeed been adopted by the Basel Committee, which in 1995 accepted that commercial banks could use their own Internal Models of risk assessment. The Brussels Commission did the same for European banks in 1997.

Nonetheless, when limited liability is taken into account, Rochet (1992) proves that, even with the correct weights, capital adequacy requirements are not enough to control moral hazard and an additional requirement, as a minimum level of capital independent of the size of the assets, might be needed.

Bhattacharya et al. (2002) model the optimal closure and replenishment rules for a bank facing Poisson distributed audits by the regulator. They consider three simultaneous regulatory policy instruments: a capital replenishment rule, a closure rule and how frequently a bank is audited. They show that a combination of these policies can completely eliminate risk-taking incentives when the bank is solvent. Weights on capital requirement and closure rules can be endogenously determined in this model and, as suggested by VaR models, closure levels increase almost linearly in the risk of the underlying assets. Their assumptions, though, are quite strong because either all agents are risk neutral or agents other than the regulator are assumed to have perfect information. The implicit objective function of the regulator in this model is to completely eliminate risk shifting incentives via the closure rule. However, if a cost-benefit analysis were to be explicitly formulated, other policies could be evaluated and compared. Moreover, the use of VaR models can also have a negative effect on risk shifting, especially in situations of financial distress. Daníelsson (2000) argues that because "*market data is endogenous to market behavior, statistical analysis in times of stability does not provide much guidance in times of crisis*". He finds that VaR models are excessively volatile, sometimes increasing idiosyncratic and systemic risk, and questions how appropriate risk modelling is for regulatory design.

Hellmann et al. (2000) blame the deregulation of interest rates for facilitating the increasing frequency of financial crises in recent years. They show that a Pareto efficient outcome can be implemented by a combination of deposit interest rate controls and capital requirements. The use of deposit rate ceilings can complement capital requirements, creating additional policy flexibility and allowing the government to relax binding constraints on capital, reducing the total costs imposed by them. Deposit rate ceilings create a franchise value for banks, which they lose if a gamble on assets is un-successful, introducing incentives for prudent investment. They recognise, though, that a trade-off exists, as banks with either better investment opportunities or lower costs of intermediation would be limited in their rate of growth.

In the late 1990s a new idea, called the *pre-commitment approach* (Kupiec and O'Brien, 1997), seemed to open the way for more effective regulation, using the private interest of banks in assessing their risk-taking and making appropriate provision for it. This idea basically states that banks, having a better knowledge of the market and the credit risk contained in their portfolios, can self-assess their maximum possible losses, which would determine their capital requirements. If this capital proves to be insufficient for an observed volatility, the regulator can impose an ex-post pecuniary penalty for violation. Nonetheless, and as was noticed by Rochet (1999), the main real difference between the pre-commitment approach and the use of internal models is simply the time of intrusion by the regulator (when the penalty is applied). In fact, the implementation of this method has proved to be no easier than that of internal models and it has practically been abandoned (Freixas and Santomero, 2002).

Instead, interest has shifted to the use of market instruments and prices to increase private discipline (see Freixas and Santomero (2002) for a survey). The new BIS proposal for capital regulation, Basel II, includes three pillars: minimum capital requirements, supervisory review of internal assessments and capital adequacy, and effective disclosure to strengthen market discipline. Decamps et al. (2004) and Rochet (2004) study the combined effects of these three pillars in a continuous time model. Capital requirements are modelled here as an optimal closure threshold for the bank. The bank's asset value follows a stochastic diffusion process, which mean is increasing on the manager's monitoring effort level. They show that, when all agents are risk neutral, a requirement for the bank to issue a security whose payoff is conditional on the value of the bank's asset (a subordinated bond), can be used to reduce the minimum capital requirement needed to prevent moral hazard (the closure threshold). Indeed, when the regulator is subject to political pressure for supporting the banks who hit the closure threshold (forbearance), they prove that market discipline is still effective, the more so the higher the frequency of the rollover of the subordinated bond.

As supervisors have access to imperfect information, limited by the frequency of balance sheet disclosures and onsite inspections, some policymakers have advocated introducing a requirement of a minimum issuance of subordinated debt, as a complement to supervisory efforts to encourage a sound banking system. Efficiency should be enhanced, as uninsured junior bond holders have the correct incentives to monitor the bank's investments and demand an appropriate compensation for risk-taking. Then, supervisors could use these market prices to improve their assessments of a bank's solvency (Decamps et al, 2004). There are, of course, detractors of this idea, who raise questions about the reliability of creditors' signals, given that they are also subject to asymmetric information, or the depth of the ability of secondary markets to provide informative price signals (Kane, 1995).

1.4.2 Deposit Insurance

Situations of panic based runs were fairly common in the USA banking system during the National Banking Era (1864-1913).[10] The introduction of deposit insurance and regulation during the 1930s has been regarded by many authors and policymakers as the main cause of the disappearance of panic runs (Miron, 1986; Williamson, 1995).[11]

Systems of deposit insurance started to appear outside the USA only in the 1960s. However, with the escalating banking crises during the 1980s, financial stability concerns led to the widespread establishment of implicit and explicit deposit insurance systems, which are well documented by the work of Demirguç-Kunt and Sobaci (2000).

At a theoretical level, Diamond and Dybvig (1983) were the first to address deposit insurance as a mechanism to stop inefficient runs on solvent banks. In the model described in section 1.2, they introduce deposit insurance as a promise to pay 1 at $t = 1$ and R at $t = 2$ – the outside option or autarkic solution – if the bank goes bankrupt in the interim period. This guarantee is financed by a tax charged on early withdrawals, in an amount depending on the number of depositors queuing to withdraw their money. Since the payment of the guarantee is triggered only by a run, and given that $R > 1$, waiting until the final period turns out to be the only dominant strategy for patient consumers. Therefore, the guarantee is never used in equilibrium.

An alternative to deposit insurance is the suspension of convertibility, which gives banks the ability to stop paying out deposits when withdrawals reach a pre-determined level. Suspensions of convertibility were an unpopular but common practice in the 1930s, before the introduction of the FDIC (Friedman and Schwartz, 1963). A recent example of its application was the so-called "*corralito*" in Argentina, in 2002.

In the model of Diamond and Dybvig (1983), suspension of convertibility means that when the number of early withdrawals rises above π – the proportion of impatient

[10] Miron (1986) documents that before 1914 banking crises in the USA occurred on average every 5 years.

[11] Chari (1989) disagrees with this judgement and argues that runs in the USA were a particular consequence of local legal restrictions and problems of regulation.

consumers – the conversion of deposits into the consumption good is interrupted. In this framework, both mechanisms (suspension of convertibility and deposit insurance) are equally effective in deterring patient depositors from withdrawing early. However, the authors show that when liquidity shocks are stochastic (π is a binary random variable) suspension of convertibility can create serious inefficiencies, as some impatient consumers could be deprived of consumption in the interim period. Thus, deposit insurance is Pareto dominant in this case.

According to some authors, the main aim of deposit insurance is to provide a risk-free asset to small savers (Folkerts-Landau and Lindgren, 1998), while others argue that the market could provide cheaper assets to satisfy this purpose (Calomiris, 1996 and Stiglitz, 1992). All agree, however, that deposit insurance induces moral hazard. In fact, a guarantee on deposits can be seen as a callable put option on the agency offering insurance (Merton, 1977; Acharya and Dreyfus, 1989), whose value increases monotonically in the volatility of the investment portfolio and then is maximised at the highest possible level of risk.

Despite the limitations of Diamond and Dybvig's (1983) model,[12] research has abstracted from studying the real effects of this policy on depositors and bank managers' behaviour and has concentrated instead on the study of the pricing of the insurance option and its effect on banks' moral hazard.[13] I will come back to this point in the following chapter.

Risk insensitive insurance premiums have been linked to higher bank risk-taking because the introduction of a guarantee on deposits implies inefficient investment, which cannot be compensated for by an increase in capital requirements (Mishkin, 1992). Moreover, full deposit insurance suppresses the incentive of depositors to require their banks to self-protect through capitalisation (Bond and Crocker, 1993). By fixing the banks' future funding costs, risk insensitive insurance premiums also eliminate the funding-related benefits of reputation (Boot and Greenbaum, 1993). This problem is particularly severe when there is higher competition among banks, as the monopolistic rents that would otherwise have encouraged banks to effectively monitor their investments are eliminated.

Acharya and Dreyfus (1989) develop a model in which the authorities receive a signal of the true value of the bank's assets and can close a bank before the end of

[12] Runs are a sunspot phenomenon and, as the possibility of a bank becoming insolvent is excluded, runs are always inefficient.

[13] Notice that as in Diamond and Dybvig (1983) deposit insurance is not used in equilibrium, the insurance premium is nil.

the contract. They determine the optimal closure policy together with the price of deposit insurance, as the minimum cost policy for the insurer. They find that a bank is optimally closed if the net increase in the insurer (discounted) liability exceeds the immediate cost of reorganizing the bank, or if the bank's current asset value is too low for the insurer to be able to charge an actuarially fair premium. However, as was noticed by Chan et al. (1992), when asymmetric information is taken into account, fairly priced deposit insurance may not be possible. Adverse selection and the existence of time lags between banks' portfolio decisions and premium adjustments give managers incentives to gamble for resurrection, due to limited liability. Freixas and Rochet (1997) show that deposit insurance may be feasible under asymmetric information but not desirable from a general welfare point of view because ex-post cross-subsidies between banks may lead to an artificial survival of insolvent banks.

1.4.3 The Lender of Last Resort

The concept of the Lender of Last Resort (LoLR) was first introduced in the eighteenth century, originally by Baring (1797) and later by Thornton (1802) and Bagehot (1873). They identified this as an important role to be performed by central banks, in order to promote financial stability and avoid the consequences of the spread of bank failures for the real sector.

The LoLR facility can be defined as "*a discretionary provision of liquidity to a financial institution (or the market as a whole) by the central bank, in reaction to an adverse shock which causes an abnormal increase in demand for liquidity that cannot be met from an alternative source*" (Freixas et al, 1999). However, as in practice it may be difficult to distinguish whether OMOs have been implemented to rescue troubled banks or with pure monetary policy objectives, Goodhart (1995) argues that the term LoLR should be reserved to mean liquidity support to individual banks.

Bagehot's doctrine is that the LoLR should lend only to solvent banks facing a liquidity shortage, at a high penalty rate (in order to reduce the incentives for using these loans to fund normal businesses) and against good collateral, according to their value at pre-crisis levels. He believed that without good collateral, an institution is plainly insolvent and therefore should be allowed to fail. However, some authors consider that banks with good collateral should be able to borrow from the interbank market and, therefore, any emergency liquidity assistance should be provided only through open market operations, in order to avoid moral hazard (Goodfriend and King, 1988). Put in a different way, with efficient financial markets a solvent institution should never be illiquid.

Interbank markets, though, do not always operate efficiently. A recent example of this fact is the ongoing subprime crisis, where the spread of the 3 month LIBOR over the monetary policy rate kept above 100 basis points on average between September 2007 and June 2008 in the USA and 80 basis points in Europe.[14] This has required central banks to inject liquidity in the market, not only though open market operations but also through direct emergency liquidity assistance.

Due to problems of incomplete or asymmetric information, a solvent illiquid bank may become insolvent if it is unable to obtain liquidity in credit markets. In a model of a bank liquidity crisis with a single Bayesian equilibrium, Rochet and Vives (2002) show that with positive probability a solvent bank cannot find liquidity assistance, even in an otherwise efficient interbank market. Their model considers three periods and a continuum of depositors consuming only in the final period. At $t = 0$ a bank has a positive level of equity and collects 1 unit of uninsured deposits. With these resources, the bank funds its investment in a risky, partially illiquid asset and its holdings of cash reserves. If the return of the risky asset is high, depositors are repaid in the final period and stockholders appropriate the profits. On the other hand, if a bad signal on future returns is received at $t = 1$, depositors can withdraw early, precipitating a run if their withdrawal exceeds the cash reserves.

An interesting feature of this model is that withdrawal decisions are not directly pursued by depositors but are delegated to uninsured, risk neutral fund managers. This assumption aims to capture what the authors call *"the modern form of bank runs"*, where large well-informed investors refuse to rollover their credits to banks. Fund managers obtain a private benefit $b > 0$ if they make a good investment and get the money back at date 2; they are penalised with a reputational cost $C > 0$ if they have to withdraw in the interim period, so they receive $b - C > 0$; and they get nothing if the bank fails and they do not recover the money. As the expected benefit of early withdrawal is $b - C$ while that of waiting is $(1-p)b$, where p is the probability that the bank fails, managers withdraw at $t = 1$ if and only if $p > \frac{C}{b}$. In the interim period, managers receive a noisy signal on the deposits' returns and decide whether to withdraw or to wait. If early withdrawals exceed cash balances, the bank is forced to prematurely liquidate assets. If the needed amount of assets is greater than the total assets available at $t = 1$, the bank fails. If not, but what is left is not enough to cover late withdrawals, the bank

[14] The normal levels of this spread are about 30 basis points in the USA and 20 basis points in Europe, according to pre-crisis estimates.

fails at $t = 2$. Otherwise, the bank is solvent and can continue in operation. Premature liquidation is costly, in the sense that for any unit of the risky asset sold at $t = 1$ the bank receives just a fraction $r < 1$ of its value. The authors show that when a bank is close to insolvency (modelled by small values of R) or when there is a liquidity shortage (modelled by small values of r), the interbank market is not able to prevent the early closure of the bank, due to coordination problems among fund managers. They also show that coordination failure can be alleviated through a combination of appropriate solvency and liquidity requirements and liquidity assistance through discount window loans.

Although this model may justify the intervention of central banks, it does not resolve the moral hazard issue. As was noticed by Goodhart (1995), the time available to decide whether to lend to a bank is often not sufficient to assess its solvency and, even more complex, ex-ante solvent banks may become ex-post insolvent if general economic conditions worsen. It is also true that, on some occasions, central banks have rescued clearly insolvent banks because of stability concerns (*"too-big-to-fail"*), when they have assessed that their failure could generate a knock-on effect on the rest of the banking system, so the cost of restructuring is smaller than the cost of failure. Indeed, allowing banks to fail seems to be not the rule but the exception. In a study of a sample of 104 failing banks, Goodhart and Schoenmaker (1995) show that less than a third were liquidated and all the rest were rescued.

A possible answer to the moral hazard problem is the so-called principle of *"constructive ambiguity"*. The basic idea behind this is that central banks can rescue some banks and liquidate others, without committing ex-ante to a specific policy. In that way, bank managers and shareholders are uncertain about the cost they will have to bear, giving them an incentive not to take on too much risk. This principle is supported by Freixas (1999), who studies the optimal bailout policy that should be applied by a LoLR using a cost-benefit analysis. By introducing a social cost of liquidation – increasing on the value of the bank's asset – and a continuation cost, defined as a function of the expected subsidy that uninsured depositors receive through the bailout of the bank, he shows that banks with large asset holdings (equal to the sum of equity, insured and uninsured deposits) are *"too big to fail"* and are always rescued (also see Rochet and Tirole, 1996). For smaller banks, he finds that the optimal policy is to use a mixed strategy equilibrium, by which banks satisfying certain regulatory requirements are rescued with positive probability (but less than one), and computes a threshold for the amount of uninsured deposits over which every bank must be liquidated.

Rochet and Vives (2002) also address this issue. In their model, the commitment of liquidity assistance can be seen as a signal whose level of disclosure is decided by the central bank. They show that, if this signal becomes common knowledge, its effects can be destabilising and so it may be optimal for regulators to send an oblique statement to the market or to add enough noise to the signal to limit moral hazard.

While constructive ambiguity may limit moral hazard, it also gives large discretionary power to the agency responsible for crisis management. Freixas et al. (1999) propose a solution, by giving central banks a role as mediators to organize private sector liquidity assistance. Although coordination problems may be difficult to solve in the short-term, there are many examples where the concept of being "too big to fail" makes sense even from a private perspective, so that creditors cooperate and invest in new equity, as in Barings in 1890, the Long-Term Capital Management (LTCM) in 1998 and the rescue of IKB in Germany in 2007.

1.5 Systemic Risk Shifting or "Too Many To Fail"

Regulation should seek the stability of the financial system as a whole and not of individual institutions, whose failure might be an efficient exercise of market discipline. However, in the same way that a LoLR implicitly guarantees the rescue of banks considered "too-big-to-fail", a partial equilibrium approach in the design of regulation that does not take into account the emergence of correlated exposures to specific sectors of the real economy, could lead to the equivalent of a "*too-many-to-fail*" outcome. Acharya (2001) explains this phenomenon as a consequence of *systemic risk-shifting.* In his paper, he shows that regulatory mechanisms such as bank closure policies and capital adequacy requirements, based exclusively on the assessments of a bank's own risk, fail to mitigate aggregate risk-shifting incentives and can accentuate systemic risk.

Formally, he considers an economy with three periods, two banks geographically separated, and a continuum of depositors holding standard demand deposit contracts. Banks can decide whether to invest in a safe asset, whose return equals the marginal product of capital, or a risky technology consisting of loans extended to entrepreneurs over which banks have monopolistic power, so that the supply of this asset is determined by the amount of the risky investment in each bank.

A bank can choose both the riskiness of the project and the industry in which the project will be developed. Information costs are assumed to be so high that it is not feasible to invest in more than one industry. For simplicity, it is assumed that there are just two industries in the economy. The choice of industry by banks determines the correlation of their investment portfolios. If both banks choose the same industry, the correlation is high. In this way, systemic risk is an endogenous outcome of the banks' investment decisions. Other interbank relationships are excluded from this analysis.

When a bank fails, two types of externalities affect the surviving bank. There is a (negative) "*recessional spillover*", reducing the profitability of the surviving banks through the reduction in aggregate investment. There is also a (positive) "*strategic benefit*", resulting from the migration of depositors from the failed bank to the surviving institution, and the reduction in operational costs due to the acquisition of the failed bank's lending facilities. If the negative effect dominates the positive one, banks may prefer to increase their correlation, thus increasing the probability of their joint survival or failure, a phenomenon that the author calls "*systemic risk-shifting*".

Such an outcome is not optimal from a social welfare perspective because the cost of joint failure is clearly higher. Effective regulation should consider a closure policy, consisting of a bailout of the failing bank together with dilution of the bank owners' equity, where greater dilution implies less forbearance. However, while on the one hand a bailout eliminates the externalities, on the other hand it induces moral hazard because of the resulting "*too-many-to-fail*" guarantee. If bank owners can anticipate greater forbearance upon joint failure, this policy will increase the systemic risk by inducing banks to make correlated investments in order to extract greater regulatory subsidies. The author proposes a "*collective regulation*" to counteract systemic moral hazard. By conducting sales of failed banks' assets, the surviving banks' charter value will increase (positive externality) inducing them to prefer less correlated portfolios.

There is some empirical evidence that banks may find optimal an industrial organisation comprising several specialised banks instead of a large number of diversified banks, when diversification does not guarantee higher profits or greater safety for them (Acharya et al., 2006). Indeed, the contagious effects of the failure of one bank on the others, understood in the model as depositors demanding higher interest rates from the surviving bank when another has failed, could increase the incentives for joint survival and joint failure if the competitive lending margins of diversification are not so high (Acharya and Yorulmazer, 2008), this effect being specially stronger for smaller banks (Acharya and Yorulmazer, 2007b).

Acharya's (2001) work also has implications for the new Basel regulation on capital requirements. He recommends that capital requirements should be "*correlated based*", not only on each individual's bank risks but also increasing on the correlation with other banks' portfolio returns, in order to take account of systemic issues. The implementation of such a policy may be difficult, though, as these correlations are themselves time varying.

1.6 Financial Contagion

Given that the failure of a large bank or several small banks can have systemic consequences, no analysis of the banking system would be complete without considering the effects that contagion has on financial stability.

The empirical literature has provided evidence of contagion in the banking industry. For example, Schoenmaker (1996) applies an auto-regressive Poisson model to a data set of monthly bank failures during the USA National Banking System – before the establishment of the Federal Reserve, in order to abstract from distortions introduced by central bank interventions. He shows that the risk of contagion among banks is significant, as their failures appear to be dependent even after controlling for macroeconomic fundamentals.

Schoenmaker (1996) identifies two main channels of contagion: the *informational* channel and the *credit* channel. The informational channel relates to situations where signals of bad performance of one or more banks are interpreted by depositors as valid information on the solvency of banks with similar characteristics (*firm-specific contagion*), or for the whole banking industry (*industry specific or pure contagion*). Pure contagion can be associated with sunspot equilibriums, unrelated to economic fundamentals and affecting depositors' beliefs in a way that turns out to be self-fulfilling (as in Diamond and Dybvig, 1983). Firm specific contagion is also known as information-based contagion. In this case, when depositors receive a bad signal on the performance of a bank, they run on banks with similar characteristic (similar types of business or the same geographical area). This type of contagion may be related to macroeconomic shocks, for example, a downturn in the business cycle. When depositors receive information about an economic downturn affecting a specific sector, they withdraw their deposits from those banks that are perceived to have a higher exposure, thus precipitating a crisis. The credit channel, on the other hand, explains how the failure of a bank can spread through the web of linkages between banks in the interbank funding market, the payment system and markets for risk transfer.

In principle, contagion through the informational channel could be controlled via regulatory measures that restore depositors' confidence (deposit insurance, regulation, or a lender of last resort). Contagion through the credit channel, on the other hand, requires the analysis of problems in the design of financial systems, which could be improved in order to prevent – or ameliorate the consequences of – a crisis. Because the safety net has already been discussed in section 1.4, I will concentrate here on contagion through the credit channel.

1.6.1 Interbank Funding Markets

Based on the model of Diamond and Dybvig (1983), Bhattacharya and Gale (1987) observe that without aggregate uncertainty and if each bank's investment in the safe asset is publicly observable (capital requirements), an interbank market for loans can insure depositors against idiosyncratic liquidity shocks. Allen and Gale (2000b) build on this idea and develop a simple model of contagion through an interbank market for deposits in a multi-region economy (more specifically, four economic regions à la Diamond and Dybvig). Consumer types continue to be private information but this time π, the fraction of impatient consumers, is modelled as a binomial random variable that takes values π_H or π_L ($\pi_H > \pi_L$) with equal probability. Without aggregate uncertainty, and if the representative bank in each region has access to the same technologies, banks get full insurance against liquidity risk through cross deposit holdings.

The authors consider two perfectly correlated scenarios: if economy j has a high demand – π_H – in the first scenario, it has a low demand – π_L – in the second one and vice versa. They show that the optimal solution is achieved when every bank commits itself to satisfying average demand across scenarios, $\pi = \frac{\pi_H + \pi_L}{2}$, and deposits the difference between the average and actual demand ($\pi_H - \pi = \pi - \pi_L$) in the bank in the adjacent region.

They later introduce aggregate uncertainty, through a zero probability scenario in which every region but one faces the average demand. The remaining region, let say A, has an excess liquidity demand of size $\varepsilon > 0$ (i.e. total demand in A equals $\pi + \varepsilon$). For any positive value of ε, the whole economy faces a liquidity shortage and the redemption of cross deposit holdings only cancel each other, reducing the value of deposits in each cancellation. If ε were large enough for A not to be able to meet its obligations, even when liquidating all of its long-term assets, when the adjacent bank claimed its deposits it would receive less money than it needed and it would face a liquidity shortage as well. Hence, a second bank would be forced to liquidate part or all of its long-term assets, even though it was perfectly solvent in isolation. In this way, a liquidity shock in one region would generate a chain reaction that would extend from region to region. The authors conclude that connectedness is directly proportional to the risk of contagion, while completeness (understood as each bank holding cross deposits with all the other banks in the economy) is inversely proportional to this risk. When each bank is connected to all the others in the economy, all the economy has to cope with the excess demand shock but in a smaller proportion than a single bank would do. That is, each bank liquidates just a small fraction of its long-term assets, and therefore bankruptcy does not occur.

Although appealing, this model has a couple of shortcomings. For example, as all banks are essentially solvent, without an aggregate liquidity shock they would never fail. Moreover, if interbank deposits were senior to other claims, even under an aggregate liquidity shock there would be no contagion, no matter whether a bank in zone A defaults or not.

These problems are a consequence of the common information assumption. Hence, as in section 1.2, a natural answer would be to introduce private information. Dasgupta (2002) extends Goldstein and Pauzner's (2000) model to a two region economy and explains contagion as a rational expectation equilibrium in a dynamic setting that, contrary to Allen and Gale's (2000b) prediction, complete markets for interbank deposits cannot prevent. He considers two ex-ante identical economies, that in the interim period experience a negatively correlated, non-aggregate liquidity shock ($\pi_H = \pi + \varepsilon$, $\pi_L = \pi - \varepsilon$). The representative bank in each location insures against this shock by issuing cross deposits of size $\pi - \pi_L$ at $t = 0$. Importantly, interbank deposits are assumed to be senior to depositors' redemptions.

Two technologies are available during the planning period: a safe technology whose interest rate is normalised to zero; and a long-term, risky technology, whose return depends on a bank-specific fundamental value, θ_j, uniformly distributed in the interval $[0, 1]$. Early liquidation is costly, as the technology returns only $r < 1$ if it is liquidated in the interim period.

In the interim period, nature decides which country's patient depositors play first. After interbank claims settle, depositors in the selected bank, j, receive a private non-verifiable signal $\theta_j^i = \theta_j + \varepsilon_j^i$ (where i stands for depositors, and $\varepsilon_j^i \backsim U[-\varepsilon, \varepsilon]$ i.i.d.) and choose their optimal strategy. Depositors demanding early withdrawals are paid out in this period, provided there are sufficient available resources. After observing region $j's$ outcome, depositors in the other region, $-j$, also receive private signals and choose their actions. In the final period, interbank claims settle and then residual depositors' claims in the two banks are paid out. Upper and lower dominance regions in each stage of the game ensure the existence of a monotonic equilibrium for each sub-game. The author proves that a unique monotone equilibrium exists for the complete game and, interestingly, contagion flows only from the debtor to the creditor bank. Formally, he proves that when a debtor bank fails, the region of fundamentals for which the creditor bank survives shrinks. Moreover, when cross deposit holdings are complete, contagion, far from being ruled out, increases with the size of the local liquidity shock (ε).

A different approach is taken by Huang and Xu (2000). They describe an economy where a financial crisis can develop endogenously in an interbank market with informational asymmetries, where both liquidity and technological shocks are present. When financial institutions cannot commit themselves to the liquidation of bad projects, a negative externality affects the interbank market due to the reduction of the average portfolio quality of the participating banks. If the interbank market cannot distinguish between solvent and insolvent institutions, solvent banks may find it too expensive to borrow, so they prefer to liquidate their assets. Having withdrawn from the interbank market, the average portfolio quality is further reduced until the market finally collapses.

These authors consider an economy with four periods and M ex-ante identical banks, comprising an interbank market for the trading of liquidity. Banks are risk neutral profit maximisers and they have no equity of their own. Instead, they choose to invest consumers' deposits in projects offered to them by entrepreneurs, who have no capital to finance themselves. Liquidity risk is modelled, as in Diamond and Dybvig (1983), by the existence of two types of depositors. Impatient depositors consume at $t = 1$ and patient depositors consume at $t = 3$, the final period. They are all ex-ante identical and realise their type at $t = 1$. Depositors' types are private information, while the realisation of the random proportion of early consumers, π, is common knowledge.

The projects exhibit constant returns to scale, last for three periods and require a positive investment, I_t, in every period. Projects can be good or bad. A good project has an ex-ante profitable return Y in the final period and a bad project returns 0. Neither the banks nor the entrepreneurs know the type of the project when the investment decision is made. All that is known is that a fraction λ of them are good, the remaining fraction $(1 - \lambda)$ are bad and the expected return of the pool of projects is positive: $\lambda Y + (1 - \lambda)X - I_1 - I_2 - I_3 > 0$.

By running a project, an entrepreneur learns its type at $t = 1$. If the project is good and it is financed, she receives a private benefit $\overline{b}_3$ at time $t = 3$, while the financing banks receive Y. If the project is bad, she can either quit immediately, receiving b_1, or she can hide this information until period 2, when the project's type becomes public information. At $t = 2$ banks have to consider whether to liquidate or to reorganise bad projects. If a bad project is liquidated, the entrepreneur receives b_2 and the banks receive nothing. Reorganisation has a cost F. If a bad project is reorganised, the entrepreneur receives $\underline{b}_3$ and banks receive a return $X > 0$, which is ex-post profitable if F is small enough, i.e. if $I_3+F < X < I_2+I_3$. Private benefits are such that $0 \leq b_2 < b_1 < \underline{b}_3 < \overline{b}_3$.

The commitment problem is introduced through the financing mechanism in operation. With a *single-bank financing scheme* (each project is financed by only one bank), reorganisation costs are negligible and bad projects are never stopped. The interbank market is then unable to distinguish among banks, as the only available information is the average quality of the pool of projects and they charge the same price for liquidity to every bank. Defining by λ_m the probability that an illiquid bank m is financing a good project, and assuming $\lambda_m \geq \lambda_{m-1}$ for all $m = 1, ..., M$, the average quality of the interbank market is $\overline{\lambda} = \sum_{m=1}^{M} \lambda_m$. Let π_m be the proportion of impatient consumers in bank m. If $\pi_m > \overline{\pi}$ (the average), the bank is illiquid and issues a bond contingent on the realisation of its project at date 3; returning 1 if the project is good and 0 if it is bad. The equilibrium bond price must equal the expected probability that the bond will pay out, that is $\overline{\lambda}$. Banks, having private information about their liquidity, withdraw from the interbank market if they face borrowing costs higher than premature liquidation costs (assumed to be fixed and exogenous). The authors show that the marginal borrowing cost is equal to $\frac{\lambda_m}{\overline{\lambda}^2}$, so the higher the quality of the bank, the higher its borrowing cost. As the more solvent banks are precisely those which first quit the market, the average quality of the pooled portfolio is reduced ($\overline{\lambda}_{m-1} < \overline{\lambda}_m$) and borrowing costs increase. When more solvent illiquid banks quit, runs spread, eventually causing the collapse of the interbank market for liquidity.

With a *multi-bank financing scheme* (each project is financed by two banks) and if reorganisation costs are so high that bad projects are always liquidated, illiquid but solvent banks can always be differentiated from insolvent ones. Hence, solvent banks can obtain credit from the interbank market at a price lower than the cost of asset liquidation. The authors prove that, in this scenario, a bank run only occurs when a bank faces both severe technological and liquidity shocks and, more interestingly, that contagion is always ruled out.

Giannetti (2001) provides another example of how contagion can arise in an economy with close bank-firm relationships. Under incomplete information, international investors cannot distinguish among banks' types, raising interest rates and precipitating the default of otherwise solvent banks. As borrowing costs become progressively higher, contagion spreads throughout the affected country and possibly to countries equally rated by international investors. His model considers three types of agents: a continuum of mass one of project managers, who have private information about the quality of their projects at $t = 0$; domestic banks, which at $t = 1$ realise the quality of the projects they finance; and international investors, who provide capital by making deposits in domestic banks. Projects can be of two types, fast or slow. Fast projects are perfectly solvent and are able to repay their loans in every period. Slow projects can be illiquid but solvent with positive probability; or insolvent, also with positive probability. At time t, international investors announce the minimum interest rate that they require to make deposits in domestic banks. Given this cost of funds, banks decide whether or not to renew the loans of the projects that they finance. If previous loans were repaid and a project is refinanced, the project manager appropriates net profits at $t + 1$. If loans have not been repaid, the bank must decide between refinancing and appropriating the realised profits.

In this framework, the author shows that, in a small open economy with close bank-firm relationships, even if banks commit themselves to stopping bad projects (slow insolvent projects), the financial system may be subject to contagion. Incomplete information generates uniformly low interest rates when a lending boom starts. A later increase in interest rates makes insolvent banks default. Illiquid banks (those financing temporarily illiquid projects) do not default immediately – sending a signal of their solvency to the market – but the temporary increase in interest rates may drive them into insolvency, causing their default after a few periods.

The two papers described above emphasise – although in different fashions – that the exposure of the banking sector to given sectors of the real economy, via close bank-firm relationships, creates informational problems that make interbank markets for loans operate inefficiently. In the first model, it was precisely the more solvent banks that were forced to quit the interbank market first. In that case, a discount window facility would be desirable to avoid the inefficient liquidation of assets. In the second model, however, it is not the premature liquidation of assets which makes banks insolvent but it is the rise in interest rates which turns illiquid banks into insolvent ones. Hence, the failure of the first banks is efficient – because they are insolvent – but, after they have left the market, an appropriate policy will be needed in order to protect solvent banks.

The gridlock of international interbank markets is explored by Freixas and Holthausen (2001), who show that with unsecured lending and if cross-country information is noisy, an equilibrium with integrated markets need not always exist; a phenomenon that could explain contagion. Their model involves an economy with two countries, each of them modelled à la Diamond and Dybvig (1983), but with full deposit insurance, so that runs are not caused by liquidity shocks alone. Banks can invest in a safe storage asset, in government bills or in a long-term risky asset, paying $R > 1$ with probability $p \geq \frac{1}{2}$ (common knowledge at $t = 0$) and nothing otherwise. This technology is ex-ante efficient ($pR > 1$) and it can be prematurely liquidated. In the interim period, each bank faces a *specific* shock to its demand deposits, represented by the proportion of impatient depositors, π, a binomial random variable taking values π_H with probability q and π_L with probability $1 - q$. Additionally, there can be *country-wide* aggregate shocks to demand deposits, depending on the value of q, which is also a binomial random variable taking equiprobable values $q_L < q_H$. Solvency and liquidity shocks are assumed to be uncorrelated.

At $t = 1$, each bank is characterised by a pair of common, non-verifiable domestic and foreign signals, (s_D, s_F), which can either be good or bad. There is no moral hazard because for simplicity the model assumes that banks cannot observe their own solvency but only their signals. Banks within a country receive an informative but non-verifiable common signal about the performance of their partners.[15] Cross-country signals are assumed to be noisy.[16] Then, when a bank seeks liquidity in the foreign market this could be interpreted either as showing that the bank belongs to a liquidity short country, or that the bank has a bad signal at a domestic level and cannot borrow locally. Depending on the probability of these two events, an integrated interbank market may or may not exist. The authors show that integrated markets exist only for good levels of cross-country information and significant liquidity differentials among regions. If not, due to persistent interest rate differentials, solvent banks in the liquidity short country may be excluded from the market and be forced into liquidation. They also observe that, while markets in government bills are efficient in channelling liquidity and reducing interest rate spreads, as loans in these instruments are collateralised, monitoring plays no role, allowing for inefficient survival.

Among the empirical work showing that interbank markets are an effective channel of contagion is Van Rijckeghem and Weder (2001), who test the importance of financial contagion through the bank lending channel relative to trade contagion, using data on recent emerging market currency crises.[17] They test the hypothesis that when banks compete intensively for funds from a "*common lender*" – a bank highly exposed to a crisis country –, adjustments to correct exposure or capital adequacy ratios can lead to a reduction in the credit lines offered to other countries. They construct a comprehensive

[15] In the sense that $prob(s_D = good/\widetilde{R} = R) = prob(s_D = bad/\widetilde{R} = 0) > \frac{1}{2}$.

[16] In the sense that $prob(s_F = good/s_D = good) = prob(s_F = bad/s_D = bad) > \frac{1}{2}$. Thus, cross country signals are informative conditional on the domestic signal but not directly on the return of the technology.

[17] See section 1.3.2 for the connection between banking and currency crises.

indicator of competition for bank funds, which proves to be robust to the inclusion of trade linkages in the regressions for the Mexican, Asian and Russian crises. This result is robust to the introduction of Markov switching regimes, which control for pure contagion (Fratzscher, 2000).

Goldstein and Pauzner (2004) provide a theoretical model explaining contagion due to the existence of common lenders. They consider two regions and a continuum of mass one depositors, each endowed with 2 units of consumption in the planning period, who invest one unit in each location. The rest of the assumptions are the same as in the model described in section 1.2 (Goldstein and Pauzner, 2000). In the interim period, a signal in location j is received first and depositors decide their actions. Knowing the outcome of their investment in that region, they receive a second signal, now on the bank in location $-j$, and decide their strategies in that location. The authors prove that there exists a unique range of the fundamentals for the two regions, such that if a run occurs in the first location the other region fails too; and if the first location survives so does the second one. The occurrence of a banking crisis in one region raises the probability of a crisis unfolding in another investor-related country, as investors' welfare is significantly reduced after the first run. This happens because, when two regions share the same international investors, correlations in investments emerge after one region has failed, despite the fact that their fundamentals are independent. Hence, a negative externality is generated: once an investor diversifies her portfolio, the benefits from diversification to other agents are reduced.

1.6.2 The Payment System

In order to make a payment, a payer needs to issue a paper-based (cheque) or electronic (plastic card) instruction to the bank where the money to be transferred is held, and then that bank proceeds to transfer the money to the bank where the payee's account is. A payment system is an arrangement to facilitate these transfers (Latter, 1997).

Payment systems are essential for the efficient functioning of financial markets because they allow transactions to be completed safely and on time. They have become increasingly important during recent decades. For example, in 1990 Fedwire in the US processed 63 million transactions valued at $200 trillion, against 134 million transactions in 2006 valued at $573 trillion. Indeed, both the Fedwire in the US and the Sterling CHAPS in the UK, represented more than 4,000 times the countries' GDP in 2006.[18]

Nonetheless, as they build on interbank relationships, payment systems can involve

[18]BIS, Payment System Statistics.

significant exposure to risk for their members. Indeed, as transactions tend to have time-critical settlement deadlines, because they are part of a chain of transactions, a domino effect following the failure of a large bank might trigger a crisis.

The BIS[19] classifies the risks in payment systems into four main categories:

1. Credit risk: the risk that a counterparty will not meet an obligation in full value, either when it is due or at any time thereafter.

2. Liquidity risk: the risk that a counterparty will not settle an obligation in full value when it is due but at some time thereafter.

3. Operational risk: the risk that hardware or software problems, human error or malicious attack will cause a system to break down or malfunction, giving room for financial exposures and possible losses.

4. Legal risk: the risk that unexpected interpretations of the law or legal uncertainty will leave the payment system, or some of its members, with unforeseen financial exposures and possible losses.

The first two types of risks are studied in a paper by Freixas and Parigi (1998). Using a cost-benefit analysis, they compare the two main types of large interbank payments systems, *net* and *gross*, with the aim of determining which one is better in terms of a safe and efficient use of liquidity. In gross systems, transactions are settled on a one-to-one basis in central bank money. Banks are not linked through intra-day credit but they need to hold large reserve balances in order to execute their payments. In net systems, intra-day credits are extended among banks, as positions are settled only at the end of the day. This exposes them to contagion risk, while significantly economising on liquidity.

Their model considers two island-economies à la Diamond and Dybvig (1983), each with one risk neutral and perfectly competitive representative bank, with access to the same two technologies. The storage technology returns 1 unit of consumption per unit invested in the previous period. A riskless, perfectly liquid long-term technology returns $R > 1$ at $t = 2$ per unit invested at $t = 0$. If a fraction α is liquidated in the interim period, this technology returns αR at $t = 1$ and the remaining $(1 - \alpha)R$ at $t = 2$. As this model does not consider the costs of early liquidation, the only possible runs are speculative.

[19]See *Core Principles for Systematically Important Payments Systems,* BIS (2000).

Consumers cannot directly invest in these technologies, but they can invest their initial wealth in their own island's bank. As in Diamond and Dybvig (1983), consumers are of two types, patient or impatient, and this is private information. The number of impatient consumers, π, is common knowledge. An additional source of uncertainty is introduced in this model, by allowing patient consumers to decide where (in which island) to consume in the final period. Payments across locations can be made by direct transfers of liquidity (patient consumers withdraw early and take their deposits with them to the bank in the other island), or through claims against the bank in the other location. The first option models a gross payment system and the second a net payment system. A fraction $1-\beta$ of patient consumers are "*compulsive travellers*", meaning that they have to consume in the other island in the final period. The remaining fraction β are "*strategic travellers*" and they decide at $t = 1$ where to consume at $t = 2$. As in this model the return of the long-term asset is deterministic, strategic travellers prefer to stay in their own island because by doing so they obtain a greater return ($c_2 \geq c_1$).

In a gross payment system, banks always need to liquidate a fraction of their investment to satisfy compulsive travellers' excess demand. The implicit cost of this system is the foregone investment return, as liquidation occurs before the arrival of incoming travellers and then their deposits cannot be used to match those of leaving customers. Conversely, with a net payment system, banks are linked by a contract and they extend credit lines to each other to finance the future consumption of travellers, without the need to liquidate assets. The obvious conclusion is that when returns are certain, a net settlement dominates a gross settlement. This result, though, is not robust to the introduction of uncertain returns.

Consider the case where the long-term return is random, being high ($\widetilde{R} = R_H$) with probability p_H and low ($\widetilde{R} = R_L$) with probability p_L. Assume that these values satisfy $R_L < 1 < R_H$ and $p_H R_H + p_L R_L > 1$, hence the expected return is ex-ante profitable. Also assume that at $t = 1$ patient consumers receive fully revealing signals on the return of the risky asset, which are uncorrelated across islands. Then, as in the asymmetric information models in section 1.2, bank runs have a disciplinary role in closing insolvent institutions.

Ex-ante, both banks offer the same demand deposit contract. Strategic travellers have three possible actions: to wait (W) and consume in the same location, to travel (T) and consume in the other island, and to run on the bank (R) and take the money with them to whatever destination they choose. Clearly, compulsive travellers will always choose to run on the bank if strategic travellers decide to do so (as their set of strategies is a subset of the set of strategies of strategic travellers). When a high signal is observed, the optimal action for strategic travellers is to wait and for compulsive travellers to travel: (W,T). Conversely, if a low signal is observed in the native island, optimal strategies depend on the signal received about the other island bank's return. If it is high, all the patient depositors will choose to travel (T,T); and if it is low, the best they can do will be to run on the bank (R,R).

As in a gross payment system banks are not linked, they are not exposed to contagion but have to make intensive use of liquidity. In a net system, on the other hand, banks are exposed to contagion through intra-day credits. The authors prove that under the latter scheme two equilibriums are possible:

i. A *potential contagion* equilibrium ((W,T) or (T,T)), which occurs if and only if the equilibrium expected payoff for strategic travellers in the low-signal bank exceeds that from running. Hence, although there are no runs, banks are exposed to contagion.

ii. A *contagion-triggered bank run* equilibrium (R,R), which occurs if consumers rationally anticipate the potential effect of contagion on future consumption and optimally decide to run on their banks.

The benefits from netting come from the possibility of raising investment in the risky technology, allowing travellers to share the higher expected return in the final period. Its costs, on the other hand, come from the inefficient survival of low signal banks. By comparing the costs and benefits of both systems, the authors conclude that a gross system should be preferred if the probability of low returns is high, the opportunity cost of holding reserves is low and the proportion of consumers that will go to consume in another allocation is low. Otherwise, a net system is dominant.

While a real time gross settlement system (RTGS) eliminates the risk of contagion, when the central bank provides uncollateralised intraday liquidity to facilitate the process of settlement (as the Fed does) it also assumes credit risk. In the European Union, as a measure to control this source of risk, central banks are advised to take collateral from members of the system in case of overdraft. A paper by Kahn and Roberds (2001), though, advises against this measure. They study the welfare cost of a RTGS system using a neoclassical monetary model and derive the cost of this system from the cash-in-advance constraints imposed by gross settlements. They show that the effect of these constraints can be undone if the central bank makes intra-day credit freely available and that, if collateral is required against credit, a RTGS will continue to impose a liquidity cost.

Freixas, Parigi and Rochet (2000) extend their previous model to N different locations, each with a representative bank and a continuum of consumers of mass one. The focus of this paper, though, is not on the comparison of different payment systems but on the study of the *potential contagion equilibrium* under different configurations of interbank markets for liquidity. All consumers are now assumed to be patient, consequently uncertainty is not about when they consume but about *where* they consume. Let $\beta \in [0,1]$ be the proportion of traveller depositors (the same for every region), that is, those depositors in location j who need to consume in location $k \neq j$ in the second period. This time the remaining fraction of depositors, non-travellers, always consume in their home location. To be able to consume at location k, a traveller in j can withdraw her money at $t=1$, carrying the cash to the following bank (hence losing higher returns in the long-term technology if the home bank were solvent), or transfer her deposits to region k, for which at $t=0$ bank j extended credit lines to other banks in the system. Deposits have a fixed contracted value in the first period (if banks are solvent) equal to c_1 but their value in the second period is endogenously determined in each location as an equal share of the remaining assets, assuming that all liabilities have the same priority in the final period. Finally, the risky technology is only available to banks,[20] it has a random return of R_j in location j if held until period 2, and pays $r \leq 1$ (the same for every region) if liquidated at $t=1$. This assumption implies that the long-term technology is now partially illiquid. R_j are publicly observable at $t=1$ but not verifiable until $t=2$ and can take only two values, $R \geq c_1$ or 0, both with positive probability.

The authors compare the stability of the payment system under two configurations, namely a "*credit chain interbank funding*", where banks extend clockwise credit lines to the bank in the adjacent region;[21] and "*diversified lending*", where each bank extends credit lines of equal size to every other bank in the system. They find that when all banks are solvent ($R_j \geq c_1$ for all j), coordination failure leads to the existence of at least two equilibriums in pure strategies: a *speculative gridlock equilibrium*, where all banks are inefficiently liquidated; and a *credit line equilibrium*, for which the payment system works smoothly. The possibility of a speculative gridlock equilibrium justifies central bank intervention, as in that case all banks are solvent but illiquid. Indeed, credit lines guaranteed by the central authority need not be used in equilibrium, implying no cost

[20] Which is effectively a participation constraint for consumers.

[21] Thinking of the N locations ordered as in Salop's (1979) circular city model. (*"Monopolistic Competition with Outside Goods."* Bell Journal of Economics 10, 141-156).

for taxpayers.[22] Then they move to consider the case where only one bank is insolvent and use the minimum value of deposits in the final period as a measure of the exposure of the interbank system to market discipline. Using comparative statics analysis, they show that the system is more exposed to market discipline under diversified lending than under credit chains.

However, as in the model of Allen and Gale (2000b), these results might depend on the structure of information that depositors have. Indeed, if interbank claims were senior to deposits, it is not clear that contagion would ever be an equilibrium. If private signals on the returns of banks were introduced, while the speculative gridlock equilibrium might disappear (or at least shrink), the power of diversified lending to enforce market discipline and prevent contagion might also be reduced (see Dasgupta's (2002) model in section 1.6.1 and the discussion therein).

1.6.3 Derivatives Markets

Markets in risk transfer have been developed as a way of improving the stability and efficiency of the financial system. By reducing the concentration of exposure of banks and diversifying risk beyond their customer base these instruments should, at least in theory, make them less vulnerable to regional, industry specific or market shocks. Among the main risk transfer techniques which are used by banks are loan trading, portfolio securitisation and derivatives.

Derivatives are instruments which are designed to reduce costs, enhance returns and allow investors to hedge positions, exchanging future payments contingent upon the future behaviour of a specified variable. In derivatives markets, the taker of credit (or protection seller) compensates the counterparty (protection buyer) only when a particular "*credit event*" occurs.

Financial derivatives contracts can be divided into exchange traded and over-the-counter (OTC) contracts. Exchange traded contracts are highly standardised and therefore easy to trade with any counterparty. They have a low counterparty risk because every transaction is cleared via a clearing house (or special purpose vehicle), which takes collateral from the counterparties in order to insure its position (daily margining) and it is also responsible for the administration of closing out contracts and delivery procedures. However, this comes at the cost of a higher exposure to market risk (compared

[22]This property, that the LoLR needs not to be used in equilibrium, is a clear consequence of the assumptions of common knowledge and perfect information. As in the model of deposit insurance proposed by Diamond and Dybvig (1983), emergency liquidity assistance acts as a mechanism to coordinate depositors' withdrawals, preventing the 'bad' equilibrium from occurring (see section 1.4.2).

to OTC's), as they do not always allow for a perfect hedge. OTCs, being mainly non-standardised, better meet the needs of counterparties, but at the same time they expose them to higher liquidity and credit risks.

There are three main types of derivatives products, all of which can be standardised or non-standardised. These are *futures* contracts, under which both parties are obliged to conduct a transaction at a specified price and on an agreed date; *swaps* contracts, almost exclusively OTC's, which can be seen as a subset of futures contracts involving the exchange of an asset or liability against another at a specified future date;[23] and *option* contracts, where the holder has the right but not the obligation to require the other party to buy or to sell an underlying asset at the specified price and the agreed date. The typical underlying assets are short and long-term loans, foreign currencies and equities but more recently they have also included credit risk instruments. Although still small compared to other markets,[24] credit derivatives have grown exponentially, both in size and complexity, during recent years.[25]

However attractive, derivatives contracts in general and credit derivatives in particular open up possibilities for new and unexplored channels of contagion. Although under normal circumstances derivatives are more liquid than the underlying cash market, without a market maker liquidity is often more easily lost in times of crisis. Indeed, the inherent leverage associated with derivative creation may encourage systemic risk when information is not available or is delayed. One recent example is the subprime crisis, which started in the USA in 2007, but which spread liquidity and solvency concerns over financial institutions throughout the globe; not necessarily because of the direct exposure of these institutions to mortgages in this segment of credit, but because they held derivates and other complex financial instruments which were structured over these credits, whose total exposure to losses could not be easily valued.[26]

The degree of counterparty risk to which holders of derivatives instruments are exposed to depends on the size of the exposure, the probability of the counterparty defaulting and the recovery value in the event of default. This risk could be reduced through bilateral netting, collateralised margining, guarantees or letters of credit. However, all

[23] More precisely, a swap requires exchanging cash flows, based on a notional principal, for a given period of time.

[24] Outstanding notional amounts of credit derivatives in 2006 were close to US$30 trillion, compared to US$292 trillion in interest rate linked instruments.

[25] Credit derivatives include swaptions, credit default swaps (CDS), CDS index, synthetic collateralised debt obligations (CDO) and credit linked notes (CLN), among others. Between 2004 and 2006 alone, the notional amounts in CDS contracts were raised almost 5 times (from $6,396 billion to $28,828 billion), according to BIS figures.

[26] The European Central Bank, the New York Federal Reserve, the Bank of Canada, the Bank of England and the Bank of Japan, are all known to have provided massive direct and indirect liquidity assistance to banks in the second half of 2007.

of these procedures need to be legally enforceable. In the case of OTC derivatives, the legal and operational risks are higher, because usually there are no clearing houses for these instruments and legal documentation is not standardised. Indeed, when used for speculative purposes, derivatives can be very risky because of their high leverage ratio and also because their volatility is usually higher than that of the underlying instruments. Examples of important losses due to fraud or the inappropriate management of derivatives are those experienced by the Orange County in the USA in 1994, and Barings Plc. and Sumitomo in 1995.[27]

According to figures from the British Bankers' Association (2006),[28] banks are the main players in credit derivatives markets, with a global participation close to 50 percent. More important, there is a high concentration of outstanding contracts in a few institutions. In the US, for example, at the end of 2006, one single bank held 52 percent of those contracts (Echeverria and Opazo, 2007).

New regulation has been sought to control for the risks imposed by derivatives contracts. The 1995 *Windsor Declaration* (a joint CFTC[29]/BIS effort, following the failure of Barings Plc.) called for greater transparency and cross border information sharing, especially regarding large exposures. The new European and BIS capital requirements have also tried to take into account market and credit risks, coming from on- and off-balance sheet operations, including derivatives.

However, not much work has been done, either theoretical or empirical, to model these specific markets. A step in that direction might be identifying net settlement systems with OTC contracts and gross systems with exchange traded derivatives, in the model of Freixas and Parigi (1998) (section 1.6.2), given their similarities in terms of their exposure to counterparty risk. In the case of derivatives, though, it is not clear how to define (and so compare) the costs of both contracts, an essential step in evaluating their relative level of stability. Another possibility would be to include derivatives contracts in the model of Freixas and Holthausen (2001) (described in section 1.6.1), similarly to the way they introduced repo markets, in order to evaluate their effect on liquidity provision and market discipline.

[27] The Orange County, a district of California, reported losses of $1.6 billion on interest rate derivatives trading and declared bankruptcy in December 1994. Barings, a 223-year-old British merchant bank, reported $1 billion losses on equity futures trading on the Simex and Osaka exchanges. In March 1995, the bank was purchased by ING Bank for just £1. By the end of 1995, Sumitomo Corporation reported about $1.8 billion losses in commodity (cooper) derivatives trading, threatening the stability of US banks that had been involved in financing commodity activities.

[28] British Bankers' Association (2006). "*BBA Credit Derivatives Report 2006*", BBA Enterprises Ltd. 2006.

[29] Commodity Futures Trading Commission, in the USA.

1.7 The International Lender of Last Resort

In the same way that the risk of contagion in a local economy serves as a justification for central banks' intervention as a LoLR, fears of contagion across economies have given support to the existence of an international lender of last resort.

The World Bank and the International Monetary Fund (IMF) were "*as a result of the UN Monetary and Financial Conference...in July 1944, part of a concerted effort to finance the rebuilding of Europe after the devastation of the World War II and to save the world from economic depression*" (Stiglitz, 2002). The establishment of these institutions involved a high degree of confidence in the effectiveness that the provision of international financial assistance to governments could have on the management of losses and risks. The proliferation of similar institutions, such as the Inter-American Development Bank (IADB), the European Bank for Reconstruction and Development (EBRD), the Asian Development Bank and many other UN related organisations, are evidence of this.

Some authors observe such a close a relationship between international and local LoLRs, that policy analysis for the former is seen as a direct extension of the powers of local central banks. Jeanne and Wyplosz (2001), for example, compare the effectiveness of an international LoLR intervening in international financial markets to that of a local LoLR that uses its resources to back up a domestic safety net. An obvious difference between local central banks and an international lender of last resort, though, is that the latter has no real access to inflationary money creation and, indeed, the amount of liquidity assistance that it can provide is usually ex-ante limited by charter. For this reason, institutions like the IMF have chosen to act instead as a partial LoLR, in the hope that their intervention would alleviate the immediate liquidity pressures on borrower countries. This would be done by restoring the confidence of international investors, whose incentives to run would be reduced; and also through loan conditionality, which facilitates the implementation of domestic policies that private investors would favour, therefore encouraging greater private involvement.[30]

The idea that partial official assistance would promote a complementary involvement by the private sector has become known as the "*catalytic effect*" of an international LoLR funding (most of the time this is associated specifically with IMF lending). However, recent empirical evidence (particularly after the collapse of Argentina in 2001-02) suggests a limited effectiveness of the IMF's catalytic effect. Indeed, rather the opposite effect has been observed, as emerging economies increasingly tend to depend on official

[30] These are what Hovaguimian (2003) calls the "*lending*" and "*policy*" channels, respectively.

lending when market conditions turn adverse (see Hovaguimian (2003) and Morris and Shin (2003) for a survey).

An empirical assessment of the effectiveness of catalytic finance is problematic, though, because of the endogeneity of the data. In particular, as was noticed by Morris and Shin (2003), comparisons should not focus on net outflows of capital after intervention but on the effect that no intervention might have had. Using global games techniques, Morris and Shin (2003) build a model of currency crises caused by creditors' coordination failure, to evaluate the effect of an official bailout on moral hazard and policy reform. In every period, a debtor country needs to raise funds in order to pay outstanding interest on long-term loans (normalised to zero) and principal on maturing short-term debt (normalized to 1), unless short-term creditors decide to rollover. The debtor country can raise funds in an amount θ, a normally distributed random variable, whose mean depends linearly on the value of the economy's fundamentals and the level of adjustment effort (policy reform) by the government. A country is said to be fundamentally sound if it has enough funds to pay outstanding long-term debt (i.e., if $\theta \geq 0$). If short-term creditors decline to rollover, the country could be forced to default when it has no access to additional liquidity assistance.

After observing the debtor economy's fundamentals and its level of effort, an international LoLR publicly announce a liquidity assistance package for an amount m. Calling ℓ the proportion of short-term creditors who decline to rollover, the debtor country defaults if and only if $\theta + m < \ell$. Defining correlated payoffs for the three players of this game (debtor country, short-term creditors and the LoLR) and if the LoLR intervenes only when the fundamentals of the debtor country are sound ($\theta \geq 0$), Morris and Shin (2003) prove that the success of catalytic finance depends strongly on the spillover effects of the LoLR assistance on the decisions of other players. In particular, catalytic finance succeeds only when the LoLR decision is a strategic complement both to policy reform in the debtor country and to rollover decisions by private creditors. Nonetheless, the "*window of effectiveness*" may be narrow. Over a range of the fundamentals where the LoLR lacks the commitment to tough intervention, debtors' moral hazard (understood as a poor commitment to reform by borrower governments) may prevail.

A similar model, also making use of global games techniques, is presented by Corsetti et al. (2003). When evaluating the trade-off between official liquidity provision and debtors' moral hazard, they also conclude that by introducing incentives for coordination among creditors' expectations and, therefore, raising the number of creditors who are willing to extend new loans (or rollover maturing ones), an international LoLR can help to prevent liquidity runs for any given level of fundamentals. In particular, when the LoLR becomes better informed, its signal to the market reduces the aggressiveness of private speculators, lowering the likelihood of a crisis in a way that reinforces government incentives to implement "desirable but costly" policies. Conversely, the inability to obtain these funds may discourage governments from undertaking reform.

The desirability of the conditionality (policy reform) imposed by the international LoLR is another point of conflict.[31] The reader may notice the resemblance to the case of a local central bank imposing regulatory constraints in the banking system. At an international level, though, the problem goes far beyond the simple intrusion into the management of a private bank. Fiscal austerity measures are politically unpopular, seriously undermining the commitment of governments to implement them, and even raising problems of "ownership" in policy reform. Indeed, it is not even clear that international creditors always support this conditionality, particularly when it involves rising interest rates, which weakens the banking sector and lowers the economy's expected growth.

1.8 Concluding Remarks

Financial crises, whether stock market crashes, currency, or banking crises, have a negative impact on the growth and stability of the economies that experience them. Banking crises are particularly important because of the close links between the banking system and the productive sector of the real economy. This is especially relevant in emerging economies, where alternative markets are less developed, and recent episodes of financial distress have been more common.

Throughout this chapter I have analysed the causes of banking crises and the channels by which they can spread to the rest of the banking system. Four main groups of variables explaining the causes of banking crisis have been identified: macroeconomic instability, deficient supervision, inadequate management and operational risks. All of these may interact together in igniting or aggravating a crisis.

Models explaining the emergence of banks are intrinsically related to those explaining their failure, as it is precisely their ability to transform short-term safe deposits into long-term risky loans which justify both their existence and their continuous exposure to runs. Models of bank failure can be classified into sunspot and rational expectation equilibrium models. In the latter, depositors are able to evaluate the probability of a run before making their deposits.

[31] See Stiglitz (2002) for a long critique of the policies followed by the IMF during the emerging market crisis in the late 1990s.

Bank runs are costly to society, not only because they dilute the resources of small and uninformed depositors but also because funds are cut off from productive projects, eventually reducing the growth path of an economy. For this reason, governments across the world have established a safety net for the banking system, consisting of a set of regulations and mechanisms designed to enforce an adequate level of risk-taking on banks and to increase the confidence of investors in times of distress. The more relevant components of the safety net, which are regulation, deposit insurance and the lender of last resort policy, were surveyed in this chapter.

The safety net also has a negative aspect, though, because the commitment to insure depositors or to bailout "*too-big-to-fail*" banks induces moral hazard, which in the end increases the risk of failure and so the probability of crises. Indeed, prudential regulation that does not take account of systemic risk-shifting could increase the risk of collective failure, as banks could raise the correlation of their investments in the hope of extracting greater regulatory subsidies when they are "*too-many-to-fail*".

No analysis of the banking system would be complete without considering the effects that contagion has on financial stability. The empirical literature has shed light on the evidence of contagion and two main channels have been identified: the informational channel and the credit channel. The informational channel refers to situations where signals of bad performance of one or more banks are interpreted by depositors as valid information on the solvency of banks with similar characteristics, or for the banking industry as a whole. Contagion through this channel might be controlled by regulatory measures to restore depositors' confidence, such as those included in the safety net. The credit channel, on the other hand, explains how the failure of a single bank can spread throughout the web of linkages developed in the interbank funding market, the payment system or derivatives markets. In this case, an analysis of the possible problems in the design of financial systems is required in order to prevent contagion.

Contagion in the interbank funding market for deposits and loans can arise due to the existence of asymmetric information problems, curtailing the access to credit of solvent banks, exposing creditor banks to the failure of borrower banks, or exposing borrower banks to changes in the welfare of a common creditor. For many models of contagion in the interbank markets for loans, asymmetric information arises from close bank-firm relationships. Solutions include multi-bank financing schemes and other mechanisms of information disclosure.

Contagion through the payment system is explained in models comparing two settlement systems – net and gross – and using a cost-benefit analysis. The cost of net systems can be modelled by the risk of contagion and the opportunity cost of holding reserves. While diversified lending could reduce the risk of contagion in net systems, gross systems should be preferred when that risk is high because it eliminates the source of exposure.

Contagion through derivatives markets remains a largely unexplored area, mainly because of the opaqueness and complexity of these instruments. While these contracts are attractive because of their off-balance sheet nature, they open up new possibilities for contagion, which have only recently begun to be addressed by regulators and tested by markets.

In the same way that the propagation of a crisis among banks in a local economy serves as a justification for central bank intervention as a lender of last resort, fears of contagion across economies have given support to the existence of an international lender of last resort. An international lender of last resort, however, cannot commit itself to unlimited liquidity assistance, as a central bank in theory could do. The idea that partial official assistance will promote a complementary involvement by the private sector has become known as the "catalytic effect" of international funding. Its window of effectiveness, though, may be limited, both by the willingness of borrowing governments to embrace policy reform and the rollover decisions of private creditors. The desirability of the reforms imposed by an international lender of last resort is a point of conflict in itself. Unpopular economic measures seriously undermine the commitment of governments to implement them and it is not even clear that international creditors always support them.

Much research is still needed in the area of banking, regulation and contagion. Sunspot equilibrium models, which are popular because of their tractability, have limited capability for use in the analysis of policy design, as a relevant set of players (depositors) is neglected. For example, models of deposit insurance have concentrated on the effect that non-risk adjusted premiums have on moral hazard, more than on the direct implications that the introduction of these types of contracts have on the equilibrium behaviour of depositors and banks. The following chapter will return to this point, introducing deposit insurance in a model of information-based bank runs (Goldstein and Pauzner, 2000). I will show that pure panic based runs could persist when deposits are insured and that the effect of insurance on moral hazard is proportional to the size of the guarantee. Considerations of a lender of last resort and closure policies to limit the probability of panic runs and regulation designed to control moral hazard will also be addressed.

Chapter 2

Deposit Insurance and the Risk of Runs

2.1 Introduction

Bank runs cause real economic problems. General activity typically declines substantially during panic runs, as the payment system is suspended and productive investment is halted by the unwillingness to make new loans.

Panics runs were fairly common in the United States during the National Banking Era (1864-1913) (Miron, 1986). Many authors and policymakers regard the introduction of deposit insurance and regulation during the 1930s as one of the main causes of the decrease in the rate of bank failures in that country (Williamson, 1995).[1] Although this measures were not imitated outside the USA until the 1960s (initially in India and later in Europe), escalating banking crises and concerns over financial stability and consumer protection led in the 1980s to the widespread establishment of explicit limited deposit guarantees.[2]

Nevertheless, experience has demonstrated that limited insurance is not sufficient to protect banks from runs in a weakened financial system. In the past two decades, at least twelve countries (Ecuador, Finland, Honduras, Indonesia, Japan, Korea, Malaysia, Mexico, Nicaragua, Norway, Thailand and Turkey) have temporarily extended explicit

[1] The Federal Deposit Insurance Corporation (FDIC) was created in 1933 to guarantee checking and saving accounts in member institutions. Before its introduction, suspension of convertibility was a common practice which generated a strong pressure for monetary and banking reform (Friedman and Schwartz, 1963). Diamond and Dybvig (1983) argue that the most important reform that followed was government deposit insurance.

[2] In a survey conducted by the IMF and the World Bank among 85 different systems of deposit insurance, 67 countries were offering an explicit and limited deposit guarantee in normal times, with varying types of funding (ex-ante, ex-post), membership (compulsory in all but seven countries) and mandate across economies (see Demirguç-Kunt and Sobaci, 2000).

full coverage during times of serious financial distress. Some of them did not even have an explicit system of deposit insurance before their crises.[3]

The success of blanket guarantees in stopping bank runs has been mixed. Funding constraints and macroeconomic stability appear to be key limitations to their effectiveness. In Norway, for example, runs on the banking system started in March 1988, after a small commercial bank issued an earnings report warning that it had lost all its equity. At that time, the government had no program for shoring up the capital of troubled banks, nor did it sponsor any form of deposit insurance. But the banking industry managed its own deposit insurance system – the Commercial Bank Guarantee Fund (CBGF) – that injected capital into troubled banks to cover depositors' claims, under the guidance of the public Banking, Insurance, and Securities Commission (BISC). By the spring of 1990, capital injections from the CBGF and consolidations proposed by the BISC appeared to suppress the outbreak of banking insolvency. However, after the failure of the three largest banks nearly depleted the capital of the private insurance fund, runs were out of control and the government was forced to establish a public insurance fund, which controlled about 85 percent of commercial bank assets by the end of 1991 (Ongena et al, 2000).

A blanket guarantee was also introduced in Ecuador in December 1999, after runs in the largest bank of the country started to spread to other banks. Deposits were withdrawn from all banks (those perceived as weak suffering more than banks perceived as strong) to be put into US dollars, which were held mainly in cash or transferred offshore, despite the guarantee. After a long and painful process of deposit freezing , and an international audit process which ended with the State taking over three large banks and closing down several others, the blanket guarantee still in place worked for a time, until the authorities decided to default on the external debt and unilaterally reschedule the domestic debt. It then became clear that it would be very difficult to pay the debt accumulated to guaranteed depositors from closed banks, plus that used to save the banks taken over by the State. This prompted a three-pronged debt/currency/banking crisis, which no deposit guarantee could have stopped.[4]

Although deposit insurance is a popular tool among policymakers, even partial pro-

[3] Indonesia, Malaysia and Thailand.

[4] Several other examples could be mentioned. In Dominican Republic, although the authorities had an explicit limited deposit insurance, they decided to guarantee all deposits of one large bank upon its intervention in April 2003. However, the bank was not closed and lost over a half of its deposits in two weeks. Another example can be found in Uruguay in 2002. After a few banks had their licences suspended, depositors in those banks did not receive any protection, but the remaining domestic banks got an unlimited guarantee placed on their sight and savings accounts (but not on time deposits). While this stopped runs initially, it did not halt them completely and there were subsequent runs. The author thanks Steven Seelig, of the IMF's Systemic Issues Division, for these useful examples.

tection is a controversial issue among economists. Many authors agree that deposit insurance is a source of moral hazard, that by reducing the incentives of depositors to monitor their banks it damages financial stability by encouraging risk-taking. Deposit insurance can indeed be very costly, that cost being typically born by taxpayers. For example, the USA Savings and Loan Crises (1986-1995), the most intense series of institutional failures in the USA since the 1930s, involved a loss of US$153 billion, of which US$124 billion were borne by US taxpayers. Deposit insurance is said to have been critical in this case. In the 1980s, the insurance limit was raised from US$40,000 to US$100,000 (the current limit), encouraging depositors to continue funding an already risky industry (which was reflected on raising interest rates). According to Jameson (2003), however, deposit insurance alone was not responsible for this collapse, as loose regulation on Savings and Loans activity meant that institutions were able to use these new funds to gamble their way into profit.[5]

A successful guarantee must then be accompanied by efficient regulation in order to prevent the negative effect of moral hazard on financial stability. In an empirical work Demirguç-Kunt and Detragiache (1999) find that explicit deposit insurance increases the vulnerability of the banking system, particularly when the coverage is more extensive, as its presence tends to make economies more vulnerable to rises in real interest rates, exchange rate depreciation and to runs triggered by currency crises.[6] However, good institutions (used as an estimate of a good regulatory environment) perform an important role in curbing this negative effect.[7]

Diamond and Dybvig (1983) were the first to propose deposit insurance as a mechanism to stop inefficient runs on solvent banks, in a model with perfect information and deterministic returns. Runs in this multiple equilibrium model are a consequence of coordination failure among depositors and turn out to be a self-fulfilling prophecy. With the introduction of insurance the "bad" sunspot equilibrium (runs) is always eliminated, and the policy becomes costless as it is never used in equilibrium. Since then, and despite the limited predictive capability of this model, authors have taken for granted the power of deposit insurance as an instrument for eliminating inefficient runs, and

[5]Another well known example is the rescue of Credit Lyonnais in France. In November 1992, when the bank had become seriously insolvent, the Minister of Finance promised the savings of the bank's 8 million depositors would be safe. By 1997, this promise had cost taxpayers about US$17 billion – unofficial estimates refer to losses in the rank of US$20 to US$30 billion. But in this case again, losses mounted well before full protection was in place, due to severe mismanagement and poor regulation by the authorities.

[6]Deposit insurance usually guarantees only the domestic value of deposits.

[7]The authors use a series of indexes, measuring different aspects of the institutional environment of a country, that may be positively correlated with the quality of regulation. These indexes include: the degree to which the rule of law prevails, the quality of contract enforcement, the quality of the bureaucracy, the extent of bureaucratic delay and the degree of corruption.

have concentrated instead on studying the problem of pricing of insurance and its effect on banks' moral hazard (for a survey on this subject see section 1.4.2 in the previous chapter).

This chapter will come back to the study of the effects that deposit insurance has on the equilibrium behaviour of depositors and banks, while abstracting from the problem of insurance pricing. In particular, I want to consider whether the empirical findings described above can be supported by this model. I will consider Goldstein and Pauzner's (2000) model of information-based bank runs, where private information allows for a unique equilibrium. I will show that while consumers achieve better risk-sharing in a competitive banking system than in autarky, more solvent projects are liquidated as uninsured depositors fail to coordinate in a subset of fundamentals, and run on banks they know to be solvent. When introducing deposit insurance, I show that its effectiveness in eliminating panic runs varies with the size of coverage and the degree of supervisory involvement of the agency in charge of insurance. Under a narrow mandate (when the agency is not involved in the supervision of banks), a deposit insurance contract preserving the monitoring role of depositors involves offering less than full protection. The trade-off is that panic runs cannot be completely eliminated with a partial guarantee, although it does reduce the region of fundamentals for which that occurs. Under a broad mandate (with a high degree of supervisory involvement), I show that panic runs tend to disappear for any level of insurance as the regulator's signal becomes more precise, given that liquidity assistance is committed to solvent but illiquid banks. Moreover, it is cost efficient never to provide liquidity to insolvent banks. However, only extremely insolvent banks are closed, and those with enough funds to cover the payment of the final period guarantee are allowed to continue in operation. Therefore, the smaller the protection offered to depositors, the higher is forbearance. All these results hold, irrespective of the specific values of the guarantee, which in particular might imply the social cost of deposit insurance to be lower under a broad mandate. Finally, I show that deposit insurance increases the equilibrium value of the demand deposit contract in the interim period and so the probability of runs, at least for high levels of the guarantee, but this effect seems also to be smaller under a broad mandate. Limited insurance can contain this externality to some extent, justifying the observed conduct of governments across the world in normal times.

Given the combination of these results and the empirical evidence provided by other authors, this chapter concludes that a preferable scheme would be one in which the agency in charge of insurance has more supervisory involvement (broad mandate) or a high degree of coordination with the supervisory authority.

The chapter is organised as follows. Section 2.2 introduces the benchmark model of information-based deposit runs, as developed by Goldstein and Pauzner (2000). Deposit insurance is justified because of the inefficient liquidation of solvent banks in equilibrium. Section 2.3 introduces deposit insurance under two possible mandates for the insurer. The equilibrium under a narrow mandate is discussed in section 2.4, and that under a broad mandate, in section 2.5. Section 2.6 compares the optimal demand deposit contracts offered under the two mandates. Policy implication and possible extensions are discussed in section 2.7. Finally, conclusions are given in section 2.8.

2.2 A Model of Information-based Bank Runs

One of the simpler and better known models explaining the inherent fragility associated with the banking system belongs to Diamond and Dybvig (1983). The maturity mismatch between long-term loans financed with short-term deposits exposes banks to the risk of runs. Crucially, public information on the quality of a bank's investment portfolio leads to multiple equilibria, one of which involves coordination failure among depositors, who run on a solvent bank solely because they fear other depositors will do the same. As a result, the probability of runs is undetermined in this model, seriously limiting its usefulness as an instrument to evaluate policies designed to reduce banking fragility.

Goldstein and Pauzner (2000) modify this model, using global games' techniques.[8] By replacing common knowledge on the bank's fundamentals by noisy private signals received by depositors, a unique equilibrium emerges in which fundamentals act as a mechanism to coordinate agents' beliefs towards a more efficient outcome.

Consider an economy with three periods $t \in \{0, 1, 2\}$ and a perfectly competitive banking industry, where all banks have access to the same two investment technologies at the planning period ($t = 0$). Banks are risk neutral and decide whether to invest in a liquid technology, returning 1 unit of consumption at $t + 1$ per unit invested at t; or in a stochastic, long-term, partially illiquid technology, returning 1 if liquidated in the interim period ($t = 1$), and R if liquidated at $t = 2$. The long-term return function, $R = R(\theta)$, is a continuous and increasing function of a random variable θ, uniformly distributed in the interval $[0, 1]$, that represents underlying fundamentals of the projects

[8] Global games, first studied by Carlsson and van Damme (1993), are games of incomplete information where players observe noisy signals of an uncertain underlying economic state or fundamental, which determines the payoffs of the game. For a review of the theory see Morris and Shin (2002).

financed by a bank. After receiving deposits, banks decide which project to invest in. Assuming that $E[R(\theta)] > 1$, investment in the risky project is superior to storage and, therefore, all resources are pulled on it. Because the zero profit condition implies all banks will offer exactly the same contract, it is possible to restrict the analysis to one representative bank.[9]

A continuum of mass one consumers receive 1 unit of endowment – let us say money – at $t = 0$, that they invest in the representative bank which offers a demand deposit contract (c_1, c_2).[10] Depositors are risk averse, with preferences represented by a concave and increasing utility function, $u(c_1, c_2)$, with coefficient of relative risk aversion higher than 1. Depositors are uncertain about their time of consumption. With probability $1 - \pi$ a depositor is patient, meaning she enjoys consumption only at $t = 2$. With complementary probability, π, she is impatient and consumes only in the interim period.[11] At $t = 1$, types are privately realised and all impatient depositors withdraw to consume, whereas patient depositors evaluate the expected payoff at the final period, conditional on their belief in the response of their counterparts, and decide whether to withdraw or to remain.

Let n be the total number of withdrawals at $t = 1$, so $n - \pi$ is the number of patient depositors withdrawing in the interim period ($n \in [\pi, 1]$). If $n > 1/c_1$, the bank does not have enough resources to pay the promised value of deposits, c_1, and it is liquidated. Thus, each consumer demanding early withdrawal receives $1/n$ and those waiting until the second period receive 0. On the other hand, if $n \leq 1/c_1$, the bank survives to the final period and all depositors demanding early withdrawal receive c_1. For simplicity, assume that the bank is always liquidated at $t = 2$ and, because there was no equity in the initial period, remaining customers equally share the value of final assets, i.e. $c_2(\theta, n) = \dfrac{(1 - nc_1)R(\theta)}{(1 - n)}$.[12]

Naturally, $1 \leq c_1 \leq 1/\pi$, otherwise depositors would prefer not to invest in the bank (first inequality), or runs would be triggered by the demand of impatient depositors alone (second inequality).

[9] By assuming that depositors cannot invest in more than one bank, the contract that one bank offers does not affect the payoffs of depositors on a different bank.

[10] Every bank in this economy is ex-ante identical, therefore it is possible to normalise the size of the representative bank to 1 (there is no equity at $t = 0$). Limiting the analysis to demand deposit contracts is a standard assumption in the literature, not restrictive, because these contracts are effectively observed in banks. This assumption, however, implies that this model does not solve for the optimal contractual form. For a justification of the use of demand deposit contracts see *"A Remark about Demand Deposit Contracts"* in chapter 1.

[11] With no discounting, the utility of impatient agents is $u(c_1, c_2) = u(c_1)$, while that of patient depositors is simply $u(c_1, c_2) = u(c_2)$.

[12] This simplification becomes natural under the assumption of perfect competition in an economy with a finite planning horizon. In practical terms, it means that in the end the bank actually behaves as a mutual fund.

Patient depositors' payoffs are summarised in the following table:

	$n \leq 1/c_1$	$n > 1/c_1$
$t = 1$	c_1	$1/n$
$t = 2$	$c_2(\theta, n)$	0

TABLE 2.1: Patient depositors' payoffs in the game without insurance.

At the beginning of period 1, each depositor receives a private, non-verifiable signal on the true value of the fundamentals, $\theta_i = \theta + \varepsilon_i$; where ε_i are i.i.d. random variables, uniformly distributed in the interval $[-\varepsilon, \varepsilon]$. This distributional assumption implies that signals are equally informative among depositors.[13]

If a consumer were sure that $\theta < \theta_L$ – where θ_L is defined as the solution to $c_2(\theta_L, \pi) = c_1$ –[14] withdrawal would be a strictly dominant strategy, irrespective of the value of n (see table 2.1). In the present model, a consumer knows that $\theta < \theta_L$ if her signal satisfies $\theta_i < \theta_L - \varepsilon$, and every consumer receive signals below this level if $\theta < \theta_L - 2\varepsilon$. Hence, the interval $[0, \theta_L - 2\varepsilon]$ is the *lower dominance region,* where all patient depositors withdraw independent of the actions of other players. On the other hand, if a consumer were sure that $\theta = 1$, she would know the bank's return to be at its highest possible level and, therefore, she should prefer to remain. By continuity of the payoff function, there exists θ_U such that if $\theta_i > \theta_U + \varepsilon$ a patient depositor remains, and every consumer receive signals above this level if $\theta > \theta_U + 2\varepsilon$. The interval $[\theta_U + 2\varepsilon, 1]$ is the *upper dominance region*, where patient depositors always remain.[15]

Goldstein and Pauzner (2000) concentrate on "*equilibrium on switching strategies*", that is, an equilibrium in monotone strategies with threshold θ^*, such that if $\theta_i < \theta^*$

[13] The uniform distributional assumption is consistent with the Laplacian "*principle of insufficient reason*" –that one should apply a uniform prior to unknown events–, because it implies that around the switching point the number of agents remaining or withdrawing are uniformly distributed. Away from the switching point the density of n is not uniform (n has two atoms of probability in $\{\pi, 1\}$), but the strategy motivated by this belief coincides with the equilibrium action. The *Laplacian action* turns out to be an approximate optimal action in many binary action games. Indeed, as long as the payoff of a dominant action is increasing in the true value of the fundamentals (action monotonicity), Morris and Shin (2002) show that this action coincides with the equilibrium action. Action monotonicity is satisfied by the payoffs of the present model.

[14] An incentive compatible constraint for patient depositors to wait to the second period is that, at least for $n = \pi$, $E\left[u\left(c_2(\theta, \pi)\right)\right] \geq u\left(c_1\right)$ which, given that the utility function is concave and increasing, implies $E\left[c_2(\theta, \pi)\right] \geq c_1$. As for very low realizations of θ, $c_2(\theta, \pi) < c_1$, there must exist $\theta = \theta_L$ such that $c_2(\theta_L, \pi) = c_1$.

[15] The existence of the upper dominance region is not directly implied by the payoff structure of the game, as it was in the case of the lower dominance region. If the signal is very high, patient depositors remain, provided that other depositors wait as well and enough of them for the long-term technology not to be completely liquidated in the interim period. Alternative explanations could justify this behaviour. For example, Dasgupta (2002) proposes that when very high returns are guaranteed, a bank becomes an attractive target for potential purchase by a larger, more liquid bank, which would make it optimal for patient depositors to wait. Alternatively, for very high signal banks the supervisory authority could be willing to act as a Lender of Last Resort (LoLR, an explanation that will become natural later on, when studying the case with insurance under a broad mandate), rescuing a solvent bank when facing a liquidity shock. Anticipating that, patient depositors should remain.

patient depositors withdraw, and remain if $\theta_i > \theta^*$. Indeed, this type of solution turns out to be the only equilibrium of the game.[16]

After receiving a signal θ_i, depositor i knows that the true value of θ lies in the interval $[\max\{\theta_i - \varepsilon, 0\}, \min\{\theta_i + \varepsilon, 1\}]$. With incomplete information, depositors must condition their beliefs upon their private signals, which are positively correlated with the private signals of others. With every patient depositor following the same equilibrium strategy, a consumer rationally anticipates that if $\forall i\ \theta_i < \theta^*$, everybody will withdraw. This will be the case for values of $\theta < \theta^* - \varepsilon$. In the same way, a consumer knows that all patient depositors will wait if $\forall i\ \theta_i > \theta^*$, which is always the case if $\theta > \theta^* + \varepsilon$. Finally, in the intermediate region, because of the uniform distributional assumption on ε_i, a rational player will assign a uniform distribution to her beliefs on the number of patient depositors withdrawing early. Hence, when the equilibrium threshold is θ^*, the number of early withdrawals, n, will be given by the following non-increasing function of θ (figure 2-1):[17]

$$n(\theta,\theta^*) = \begin{cases} 1 & if \quad \theta < \theta^* - \varepsilon \\ \frac{1+\pi}{2} - \frac{(1-\pi)(\theta-\theta^*)}{2\varepsilon} & if \quad \theta^* - \varepsilon \le \theta \le \theta^* + \varepsilon \\ \pi & if \quad \theta > \theta^* + \varepsilon \end{cases} \tag{2.1}$$

In deciding whether to withdraw or to remain, a depositor must evaluate the (conditional) expected utility of these two actions. Denoting by $\delta(\theta,\theta^*)$ the difference of utilities between waiting until $t = 2$ and withdrawing at $t = 1$:

$$\delta(\theta,\theta^*) \doteq \delta(\theta, n(\theta,\theta^*)) = \begin{cases} u\left(c_2(\theta,n)\right) - u(c_1) & \text{if } n(\theta,\theta^*) \le 1/c_1 \\ u(0) - u(1/n) & \text{if } n(\theta,\theta^*) > 1/c_1 \end{cases},$$

each consumer evaluates:

[16] The usual argument to prove the uniqueness of equilibrium builds on the property of *strategic complementarity*: that the payoff of an agent choosing a determined action is non-decreasing in the number of other agents choosing the same action. For games satisfying this property, Carlsson and van Damme (1993) and Morris and Shin (2002) demonstrate the existence of a unique strategy profile surviving the iterated deletion of dominated strategies (*strategy dominance solvability*). Nevertheless, this property is not satisfied for the present game, as in the region where banks fail ($n > 1/c_1$), the payoff of a depositor who runs is decreasing in the number of other depositors running. In fact, in this region, actions are strategic substitutes. Making strong use of the uniform distributional assumption on the noise, Goldstein and Pauzner (2000) show that for any feasible belief $n(\theta)$, the regions where $\Delta(\theta_i, n(\theta)) \le 0$ and $\Delta(\theta_i, n(\theta)) > 0$ are complementary connected intervals, and therefore any equilibrium of the game must be monotone. Dasgupta (2002) extends this result to general distributional assumptions on $n(\theta)$, and shows that, for this game, there are no non-monotone equilibria in the set of all feasible beliefs over the actions of other agents.

[17] For $n(\theta,\theta^*)$ to be well defined, it has to be consistent with the beliefs implied by the existence of the dominance regions. If $\theta < \theta_L - 2\varepsilon$, everybody will receive signals in the lower dominance region and n should be equal to 1. Hence, consistency will require that $\theta_L - 2\varepsilon < \theta^* - \varepsilon$ or simply that $\theta_L - \varepsilon < \theta^*$. Similarly, if $\theta > \theta_U + 2\varepsilon$ everybody will receive signals in the upper dominance region and n should be equal to π. Therefore consistency will require $\theta^* < \theta_U + \varepsilon$.

$$\Delta(\theta_i,\theta^*) \doteq E_{\theta_i}[\delta(\theta,\theta^*)] = \int\limits_{\theta_i-\varepsilon}^{\theta_i+\varepsilon} \tfrac{1}{2\varepsilon}\delta(\theta,\theta^*)d\theta.$$

This means that upon receiving a signal θ_i, if a patient depositor's conditional expected utility of remaining is higher than the utility of withdrawing, that is if $\Delta(\theta_i,\theta^*) > 0$, she will wait to withdraw in the final period. Otherwise, if $\Delta(\theta_i,\theta^*) < 0$, she will quit the bank at $t = 1$. Finally, if $\Delta(\theta_i,\theta^*) = 0$ a depositor will be indifferent between the two actions.

Theorem 1 (Goldstein and Pauzner, 2000) *There exists a unique equilibrium threshold θ^* in the interval $]\theta_L - \varepsilon, \theta_U + \varepsilon[$ satisfying $\Delta(\theta^*,\theta^*) = 0$, $\Delta(\theta_i,\theta^*) < 0$ for all $\theta_i < \theta^*$, and $\Delta(\theta_i,\theta^*) > 0$ for all $\theta_i > \theta^*$. Moreover, $\theta^*(c_1)$ is increasing on c_1.*

The equilibrium threshold can be computed as the solution to[18]

$$\int\limits_{\pi}^{1/c_1} u\left(\frac{1-nc_1}{1-n}R\left(\theta^* + \frac{\varepsilon}{1-\pi}(1+\pi-2n)\right)\right)dn = \int\limits_{\pi}^{1/c_1} u(c_1)dn + \int\limits_{1/c_1}^{1}\{u(1/n)-u(0)\}\,dn \tag{2.2}$$

The existence of the equilibrium comes from the continuity of $\Delta(\theta_i,\theta^*)$ in both arguments, the dominance regions, and the monotonicity of $\Delta(\theta_i,\theta_i)$ as a function of one variable (see Appendix). That $\theta^*(c_1)$ is increasing on c_1 can be intuitively justified. If the payment in the interim period increases, more of the risky project has to be liquidated to pay early withdrawals. Therefore, the incentive for patient depositors to wait should decrease, both because the expected payoff at $t = 2$ is lower, and because each agent assigns a higher probability to the event of a run.

Proposition 2 *The probability of a bank run is equal to $\theta^* + \dfrac{\varepsilon}{1-\pi}\left(1+\pi-\dfrac{2}{c_1}\right)$, and it is increasing on the level of risk-sharing offered by the demand deposit contract (c_1).*

Proof. For any feasible value of c_1, there exists a unique equilibrium threshold $\theta^*(c_1)$ and, therefore, the function determining the number of early withdrawals, $n(\theta,\theta^*(c_1))$,

---

[18] Using $\theta(n) = \theta^* + \dfrac{\varepsilon}{1-\pi}(1+\pi-2n)$, the inverse function of $n(\theta,\theta^*)$ in the region $[\theta^*-\varepsilon,\theta^*+\varepsilon]$ to change variables, and rearranging terms:

$$\Delta(\theta^*,\theta^*) = \int\limits_{\theta^*-\varepsilon}^{\tilde{\theta}} \{u(0) - u(1/n)\}\,d\theta + \int\limits_{\tilde{\theta}}^{\theta^*+\varepsilon} \{u(c_2(\theta,n)) - u(c_1)\}\,d\theta = 0$$

$$\Leftrightarrow \int\limits_{1/c_1}^{1} \{u(1/n) - u(0)\}\,dn = \int\limits_{\pi}^{1/c_1} \{u(c_2(\theta(n))) - u(c_1)\}\,dn.$$

is also uniquely defined. A bank goes bankrupt if and only if depositors run on the bank in the interim period, that is, if and only if $n > 1/c_1$. Define by $\widetilde{\theta}(c_1)$ the value of θ such that $n(\widetilde{\theta}(c_1), \theta^*(c_1)) = 1/c_1$.

As n is strictly decreasing on θ in the region $[\theta^*(c_1) - \varepsilon, \theta^*(c_1) + \varepsilon]$ (or equivalently, for values of n in between π and 1), $n(\theta, \theta^*(c^{\widetilde{\theta}(c_1)}\big\}1/c_1$ if and only if $\theta < \widetilde{\theta}(c_1)$ (see figure 2-1). Therefore, $prob\,\{n > 1/c_1\} = prob\left\{\theta < \widetilde{\theta}(c_1)\right\} = \widetilde{\theta}(c_1)$, given that θ is uniformly distributed in $[0, 1]$. Using the inverse function of $n(\theta, \theta^*(c_1))$,

$\widetilde{\theta}(c_1) = \theta^* + \dfrac{\varepsilon}{1-\pi}\left(1 + \pi - \dfrac{2}{c_1}\right).$

Notice that $\left|1 + \pi - \dfrac{2}{c_1}\right| \leq 1 - \pi$ (see figure 2-1), therefore $0 < \theta^* - \varepsilon \leq \widetilde{\theta} \leq \theta^* + \varepsilon < 1$, and the probability is well defined and non-degenerated.[19] Finally, it increases in c_1 because $\dfrac{\partial\widetilde{\theta}}{\partial c_1} = \dfrac{\partial\theta^*}{\partial c_1} + \dfrac{2\varepsilon}{(1-\pi)(c_1)^2} > 0$. ■

A higher c_1 represents a gain in risk-sharing, as more resources from the final period are passed to early consumers. However, it also implies an increase in the probability of runs. Goldstein and Pauzner (2000) prove that $c_1 > 1$ provided that the probability of $R(\theta) < 1$ is small enough.

2.2.1 Inefficient Liquidation

Private information allows for the coordination of depositors' actions, in such a way that bank runs are avoided for sufficiently high values of the fundamentals (when $\theta > \theta^* + \varepsilon$). A natural question is then whether the equilibrium behaviour of depositors is desirable in terms of financial stability. Is it possible that panic runs persist for a region of the fundamentals, such that solvent banks can be liquidated? I will show in this section that depositors will still fail to coordinate in a subset of the fundamentals, and run on banks they know to be solvent.

Assume for a moment that both the safe and risky technologies are available to depositors for direct investment at $t = 0$. Also assume that $E[u(R(\theta))] > u(1)$, so that depositors invest all their resources in the long-term risky project. At $t = 1$ all impatient depositors withdraw. Suppose that patient depositors still receive private signals. Risk aversion implies that if $E_{\theta_i}[u(R(\theta))] \leq u(1)$ they should liquidate the project at $t = 1$, while if $E_{\theta_i}[u(R(\theta))] > u(1)$ they should hold it to the final period. Thus, in *autarky*, if a consumer evaluates the project to be solvent she should wait and withdraw in the final period.

[19] Because $\theta_L - \varepsilon < \theta^*$ and $2\varepsilon < \theta_L \Rightarrow \varepsilon < \theta^*$. On the other hand, $\theta^* < \theta_U + \varepsilon$ and $\theta_U < 1 - 2\varepsilon \Rightarrow \theta^* < 1 - \varepsilon$.

In an intermediated system, both the definition of solvency and the behaviour of depositors will clearly depend on the promised value of deposits at $t = 1$.

Definition 3 *When a bank offers a demand deposit contract paying $c_1 \geq 1$ in the interim period, the bank is said to be* ***fundamentally solvent*** *if $c_2(\theta, \pi) \geq c_1$, that is, if $\theta \geq \theta_L$.*[20]

A bank is fundamentally solvent if when only impatient consumers withdraw, the payoff at $t = 2$ is at least as good as the maximum certain payoff at $t = 1$. According to this definition, solvency is a property that cannot be verified in the interim period. No player in this game (not even the bank itself) is able to observe the true value of θ until $t = 2$. However, depositors observing signals $\theta_i > \theta_L + \varepsilon$ can be sure that the bank is solvent. Hence, if $\theta > \theta_L + 2\varepsilon$, everybody receive signals above $\theta_L + \varepsilon$ and all patient depositors know the bank is solvent. Is it then possible for solvent banks to go bankrupt in this model? Or put differently, is it possible that in equilibrium $\theta^* - \varepsilon > \theta_L + 2\varepsilon$, so that for certain values of θ all depositors run on a solvent bank?

Notice that if $c_1 = 1$, the solvency criteria would be the same as in autarky, and if $\varepsilon \to 0$ there would be no pure panic runs in equilibrium. Taking limit when $\varepsilon \to 0$ in equation 2.2, we obtain $u(R(\theta^*)) = u(1)$, which implies that $\theta^*(1) = \theta_L(1)$. That is, when the noise is negligible and the contract offers the value of liquidation of the project at $t = 1$, pure panic runs are eliminated (indeed, even if the noise were not negligible, partial runs do not occur because $\forall \varepsilon > 0$, $\widetilde{\theta}(1) = \theta^*(1) - \varepsilon$). However, as no risk-sharing is offered, this contract does not improve on the autarkic solution.

For the case $c_1 > 1$, take limit as ε goes to zero in equation 2.2:

$$\int_{\pi}^{1/c_1} u\left(\frac{1 - nc_1}{1 - n} R(\theta^*)\right) dn = \int_{\pi}^{1/c_1} u(c_1) dn + \int_{1/c_1}^{1} \underbrace{\{u(1/n) - u(0)\}}_{>0} dn$$

If θ^* were equal to θ_L then $c_2(\theta^*, \pi) = c_1$ but, as $c_2(\theta, n)$ is decreasing on $n \geq \pi$, $\theta^* > \theta_L$ would be needed only to compensate the first term in the RHS. As the second term is strictly positive, θ^* needs to increase further, which implies that for $c_1 > 1$ and $\varepsilon = 0$, $\theta^* > \theta_L$ and pure panic runs occur in this region.

For the case of a strictly positive amount of noise, consider the following numerical example: $u(c) = \ln(\frac{1}{2} + c) + 1$, which is increasing, concave, and has an index of relative

[20] *The definition of solvency introduces another constraint for the feasible values of c_1. While at the planning period the value of θ is not verifiable, a minimum condition for the bank to be allowed to operate should be to be solvent at least for the highest realisation of θ, this is*
$c_2(1, \pi) = \frac{1 - \pi c_1}{1 - \pi} R(1) \geq c_1 \Leftrightarrow c_1 \leq \frac{R(1)}{1 + \pi (R(1) - 1)}$.
Notice $ER(\theta) > 1$ and $R(.)$ increasing, imply that $R(1) > 1$.

risk aversion higher than 1; and $R(\theta) = (\theta + \rho)^2$, which is continuous and increasing on θ, and satisfies $Eu(R(\theta)) \geq u(1)$. Also consider the following values for the parameters: $\pi = 1/3$, $\rho = 0.75$. Figure 2-2 in the Appendix shows the solvency and equilibrium threshold levels (θ_L and θ^*, respectively) for different values of c_1 and ε. The dotted lines represent the levels $\theta_L + 2\varepsilon$ and $\theta^* - \varepsilon$, between which pure panic based runs occur. For example, if $c_1 = 1.1$ and $\varepsilon = 0.01$, $\theta_L = 0.3261$ and $\theta^* = 0.5948 \gg \theta_L$, and for all values of the fundamentals in the interval $[\theta_L + 2\varepsilon, \theta^* - \varepsilon[= [0.3461, 0.5848[\neq \varnothing$, all patient depositor run on the bank even though they know it is solvent. Observe that as c_1 increases, the region of pure panic runs becomes larger. Moreover, the smaller the noise the smaller the value of c_1 for which panic runs occur.

In conclusion, while a competitive banking system offers better risk-sharing for consumers; more solvent projects are liquidated than in autarky. This is costly to society, as output is reduced and jobs are destroyed.

2.3 Deposit Insurance in a Model of Information-Based Bank Runs

For the case of deterministic returns ($R(\theta) = R > 1$ constant), Diamond and Dybvig (1983) show that the introduction of a guarantee on deposits paying the outside option or autarkic solution, $(1, R)$, acts as a mechanism that ex-ante eliminates the inefficient equilibrium (runs on solvent banks).[21] Such a guarantee could be credibly financed by a tax on early withdrawals, in an amount depending on n,[22] and it is always effective in deterring runs (even when offering only partial coverage), and so it is not used in equilibrium. In a model with stochastic returns, however, the effectiveness of the insurance policy in eliminating panic runs will vary with the size of the guarantee and the degree of supervisory involvement of the agency in charge of insurance.

Consider a Deposit Insurance Corporation (DIC) offering an insurance contract (g_1, g_2) in case the bank fails. The DIC can operate under a *narrow mandate* – common in Europe –, acting basically as a pay box to compensate insured depositors of failed

[21] Diamond and Dybvig (1983) show this result holds for a more general case, where π is a binomial random variable.

[22] Considering a sequential servicing constraint, each consumer withdrawing early pays $\tau = c_1 - 1$ in taxes, that are immediately deposited back in the bank by the government, to make these resources available to pay other depositors. If $n = \pi$ the money is returned to depositors for consumption in the interim period. However, if $n > \pi$ each depositor withdrawing early consumes only 1, and patient depositors waiting to the second period receive R. Chari (1989) criticises this solution, arguing that depositors could consume their money before paying the tax, therefore making the scheme impossible to implement. However, if the government arranged for the tax to be directly paid by banks (as it usually happens in economies with well developed tax collection systems), this problem would be solved.

banks when instructed by the appropriate authority; or under a *broad mandate* – common in Asia and the Americas – where it also monitors the condition of the banking industry and takes responsibility for the resolution of failed insured institutions.

In order to stress the differences between the two systems, I will abstract from the presence of a regulatory authority in the case of a narrow mandate DIC. In this case, the agency will simply pay the guarantee every time the bank does not have enough resources to repay depositors claims, regardless of the bank being insolvent or just illiquid. Under a broad mandate, however, the DIC will also have the ability to monitor the bank's activities – which is modelled by a private signal on θ – providing lender of last resort (LoLR) assistance to solvent but illiquid banks with positive probability, and resolving inefficient banks according to a *least-cost* criteria.

Under a narrow mandate, the timing of the game with deposit insurance is as follows:

- At $t = 0$ the DIC announces the level of insurance it is going to offer (g_1, g_2). Depositors receive 1 unit of endowment (money) that they invest in the representative bank, which offers a demand deposit contract $(c_1, c_2(\theta, n))$. After receiving deposits, the bank invests in the risky asset.

- At $t = 1$ all impatient depositors withdraw. Patient depositors observe private signals on θ ($\theta_i = \theta + \varepsilon_i$) and decide whether to withdraw or to remain. The DIC observes the realisation of n. If $n > 1/c_1$, the bank's assets are liquidated and transferred to the DIC for the payment of the guarantee. If $n \leq 1/c_1$, the bank continues in operation until the final period.

- At $t = 2$, if the bank is open, remaining patient depositors are paid $\max\{c_2(\theta, n), g_2\}$. If $c_2(\theta, n) < g_2$ remaining assets are transferred to the DIC for the payment of the guarantee. If the bank goes bankrupt at $t = 1$, remaining depositors receive g_2.

On the other hand, when the DIC operates under a broad mandate, the timing of the game is as follows:

- At $t = 0$ the DIC announces the level of insurance it is going to offer (g_1, g_2). Depositors receive 1 unit of endowment (money) that they invest in the representative bank, which offers a demand deposit contract $(c_1, c_2(\theta, n))$. After receiving deposits, the bank invests in the risky asset.

- At $t = 1$ all impatient depositors withdraw. Patient depositors observe private signals on θ ($\theta_i = \theta + \varepsilon_i$) and decide whether to withdraw or to remain. The DIC observes the realisation of n and its own private signal ($s = \theta + \xi_s$, $\xi \leq \varepsilon$) and decides whether to leave the bank open – sometimes providing liquidity assistance – or to close it and pay the guarantee, in which case all the bank's assets are passed onto the DIC and all depositors claiming early withdrawal are paid out g_1.

- At $t = 2$, if the bank is open, remaining patient depositors are paid $\max\{c_2(\theta, n), g_2\}$; if $c_2(\theta, n) < g_2$ remaining assets are transferred to the DIC for the payment of the guarantee in the second period. If the bank goes bankrupt at $t = 1$, remaining depositors are paid g_2.

Deposit guarantees are usually expressed as a percentage of the principal or nominal value of deposits at the time of a bank failure, or as a limit up to which deposits can be recovered. Thus, a natural constraint for the value of insurance is $g_1 = g_2 = g \leq c_1$. Indeed, following the definition of solvency, if g were strictly higher than c_1 the DIC would have to pay the guarantee in the second period to depositors in a solvent bank, even if this were not subject to runs in the interim period.

The funding of a deposit insurance system varies from country to country. I consider here an ex-post funded system, getting resources through a government tax on withdrawals, as in Diamond and Dybvig (1983). However, as in Goldstein and Pauzner (2000), I will drop the "*sequential servicing constraint*" assumption and allow the bank to observe the length of the queue (n) before paying out depositors, so that all customers withdrawing at a given period receive exactly the same payoff. I assume that deposits are senior to other claims, so that when a bank fails at date t, its assets – or their liquidation value – are transferred to the DIC for the payment of the deposit guarantee. Therefore, the government can directly tax all early withdrawals at a constant rate equal to $\tau = c_1 - g$, transferring the revenue to the DIC for the payment of the guarantee in the final period.

The following analysis is divided into two phases. First, I study the equilibrium behaviour of depositors and the DIC under the two mandates. Second, I study the optimal decision problem for the bank. A complete formulation of the game should include a payoff function for the DIC, in order to compute the optimal level of insurance offered in the planning period. This chapter will not solve for that problem but I offer a discussion of the ideas that should be considered in the concluding section.

In order to be able to make comparisons later, I will denote by θ^* the equilibrium threshold in the benchmark model without deposit insurance, and by θ_g^* the one obtained in the model with insurance.

2.4 Equilibrium under a Narrow Mandate

(Interim Period Sub-Game)

Under a narrow mandate and once the DIC and the bank have announced their respective contracts (g and c_1), all impatient depositors withdraw in the interim period and patient depositors, observing private signals on θ, decide whether to withdraw or to remain. At this stage this is the only relevant decision because the action of the DIC is directly determined by the strategies played by patient depositors: the guarantee is paid out if and only if $n > 1/c_1$.[23]

Figure 2-3.a (in the Appendix) gives a representation of the extended form of this sub-game. According to this, patient depositors' payoffs are given by table 2.2.

	$n \leq 1/c_1$	$n > 1/c_1$
$t = 1$	c_1	g
$t = 2$	$\underbrace{\max\{c_2(\theta, n), g\}}_{\geq g}$	g

TABLE 2.2: Patient depositors' payoffs in the game with insurance under a narrow mandate.

Denote by $n(\theta)$ a feasible belief ($\pi \leq n(\theta) \leq 1$) about the aggregate behaviour of patient depositors consistent with the information received, and by $\delta_g(\theta, n(\theta))$ the difference of payoffs between waiting until $t = 2$ and withdrawing at $t = 1$, once the deposit guarantee g is in place:

$$\delta_g(\theta, n(\theta)) = \begin{cases} u(\max\{c_2(\theta, n), g\}) - u(c_1) & \text{if } n(\theta) \leq 1/c_1 \\ 0 & \text{if } n(\theta) > 1/c_1 \end{cases}.$$

After receiving a private signal, θ_i, each consumer evaluates :

$$\Delta_g(\theta_i, n(\theta)) = E_{\theta_i}[\delta_g(\theta, n(\theta))] = \int_{\theta_i-\varepsilon}^{\theta_i+\varepsilon} \tfrac{1}{2\varepsilon}\delta_g(\theta, n(\theta))d\theta,$$

the conditional expected premium for the action of waiting to withdraw in the second period when deposits are insured.

Given that $c_1 \geq g$, the lower and upper dominance regions can be defined as in the model without insurance. For $\theta < \theta_L$, the bottom left of table 2.2 is always less than or equal to c_1, hence to withdraw is weakly dominant. For very high levels of the

[23]Notice that the final period payoff is a result of the actions taken in the interim period, and that no relevant decision is made in the last stage of the game.

fundamentals ($\theta > \theta_U$), depositors should remain, irrespective of the actions of other players.[24]

If a switching point θ^* exists, a belief consistent with the existence of the dominance regions and the uniform distribution of the noise is again given by $n(\theta, \theta^*)$, in equation 2.1. Denote by $\delta_g(\theta, \theta^*) \doteq \delta_g(\theta, n(\theta, \theta^*))$ and $\Delta_g(\theta_i, \theta^*) \doteq \Delta_g(\theta_i, n(\theta, \theta^*))$.

Remark 4 *Notice that if $g < c_1$ the function $\delta_g(\theta, \theta^*)$ has a discontinuity at $\theta = \widetilde{\theta}$. Nevertheless, the function $\Delta_g(\theta_i, \theta^*)$ is still continuous in both arguments (see figures 2-4 and 2-5).*

Proposition 5 *For any given value of the guarantee and as a function of one variable, $\Delta_g(\theta^*, \theta^*)$ is strictly increasing.*

See proof in the Appendix.

If $c_1 = g$, $\delta_g(\theta, \theta^*) \geq 0$ for all θ, which implies that $\Delta_g(\theta_i, \theta^*) \geq 0$ for all θ_i. Hence, patient depositors do not monitor their banks ($n(\theta) = \pi$, $\forall\theta$). Such a solution means that insolvent banks are never liquidated, making the guarantee very expensive for low states of the fundamentals. Therefore, I will concentrate on the case $c_1 > g$.

Proposition 6 *If $c_1 > g$, there exists a unique equilibrium in switching strategies, θ_g^*, such that a patient consumer withdraws if $\theta_i < \theta_g^*$ and remains if $\theta_i > \theta_g^*$.*

Proof. Call $\Delta_g(\theta^*) = \Delta_g(\theta^*, \theta^*)$.

$\forall\theta \leq \widetilde{\theta} - \varepsilon,\ \Delta_g(\theta) = 0$ (see figure 2-5).

$$\forall\theta \in]\widetilde{\theta} - \varepsilon, \theta_L - \varepsilon],\ \Delta_g(\theta) = \int_{\theta-\varepsilon}^{\theta+\varepsilon<\theta_L} \tfrac{1}{2\varepsilon} \left\{u\left(\max\left\{c_2(\theta, n), g\right\}\right) - u(c_1)\right\} d\theta < 0,$$

as for any $n \geq \pi$ and $\theta < \theta_L$, $c_2(\theta, n) < c_1$.

$$\Delta_g(\theta_U + \varepsilon) = \int_{\theta_U}^{\theta_U+2\varepsilon} \tfrac{1}{2\varepsilon} \left\{u\left(\max\left\{c_2(\theta, n), g\right\}\right) - u(c_1)\right\} d\theta > 0,$$

as for $\theta > \theta_U$, $n = \pi$ and $c_2(\theta, n) > c_1$.

By continuity, there exists $\theta_U + \varepsilon > \theta_g^* > \theta_L - \varepsilon$ such that $\Delta_g(\theta_g^*) = 0$, and proposition 5 implies this solution is unique. ■

[24] When $\theta > \theta_L$ it is weakly dominant to remain. Nevertheless, as in this region the payoff of a single depositor will depend on the strategy chosen by other patient consumers, the existence of an upper dominance region will require additional assumptions as in the case without insurance. Consequently, assume it will also be given by the interval $[\theta_U + 2\varepsilon, 1]$.

Proposition 7 *The monotone equilibrium threshold, θ_g^*, defines the unique strategy surviving iterated deletion of strictly dominated strategies over the set of all feasible beliefs on the actions of players.*

Proof. Whatever the value of g, strategies in this game are complementary. In the region where $n \leq 1/c_1$, the payoff to remain is non-decreasing on the number of players waiting to withdraw in the second period. In the region $n > 1/c_1$ the payoff is constant, therefore non-decreasing on the number of players withdrawing. Hence, the result follows as a direct application of the results in Morris and Shin (2002) for binary actions games, with a continuum players, strategic complementarity, and a unique monotone equilibrium threshold. ■

Summing up, a unique equilibrium preserving the monitoring role of depositors exists in the model with deposit insurance if $c_1 > g$, and it is such that if a patient depositor receives a signal $\theta_i \leq \theta_g^*$, $\Delta_g(\theta_i, \theta_g^*) \leq 0$ and she withdraws; while if $\theta_i > \theta_g^*$, $\Delta_g(\theta_i, \theta_g^*) > 0$ and she remains.

Using that $n(\theta, \theta_g^*)$ is linear in the interval $[\theta_g^* - \varepsilon, \theta_g^* + \varepsilon]$ to change variables and rearrange terms, it is possible to see that θ_g^* solves:

$$\int_{\pi}^{1/c_1} u\left(\max\left\{\frac{1-nc_1}{1-n}R\left(\theta_g^* + \frac{\varepsilon}{1-\pi}(1+\pi-2n)\right), g\right\}\right) dn = \int_{\pi}^{1/c_1} u(c_1) dn \qquad (2.3)$$

Proposition 8 *The monotone equilibrium threshold for the game of information-based bank runs with insured deposits ($c_1 > g$), satisfies the following properties:*

1. *θ_g^* increases in c_1 iff*
$$u(g) > u(c_1) - c_1^2 \int_{\pi}^{1/c_1} \left\{u'(c_1) - u'(c_2(n))\frac{\partial c_2}{\partial c_1}\mathbb{1}_{\{n \leq \hat{n}\}}\right\} dn.$$
2. *θ_g^* is decreasing on g. That is, a higher value of insurance increases the incentives to remain.*

See proof in the Appendix.

Proposition 9 *Under a narrow mandate, a deposit insurance contract preserving the monitoring role of depositors involves $g < c_1$. Nonetheless, inefficient liquidation of solvent banks (panic runs) will persist for this type of insurance, the less so the higher the guarantee.*

Proof. Reconsider equation 2.3 when the signal's noise vanishes. Taking the limit when $\varepsilon \to 0$:

$$\int_{\pi}^{1/c_1} u\left(\max\left\{c_2(\theta_g^*, n), g\right\}\right) dn = \int_{\pi}^{\widehat{n}} u\left(c_2(\theta_g^*, n)\right) dn + \int_{\widehat{n}}^{1/c_1} u\left(g\right) dn = \int_{\pi}^{1/c_1} u(c_1) dn,$$

with $\widehat{n}$ defined by $c_2(\theta_g^*, \widehat{n}) = g$.

Because $g < c_1$, $c_2(\theta_L, \pi) = c_1$ and $c_2(\theta_g^*, n)$ is decreasing on n; θ_g^* needs to be higher than θ_L for this equality to hold, which means that in equilibrium depositors will still run on some solvent banks. Finally, it is clear from this equation that the higher g the smaller the region of panic runs. ∎

For a strictly positive amount of noise (e.g. $\varepsilon = 0.01$) consider the numerical example in figure 2-6. For every value of $c_1 > 1$ the equilibrium threshold is above θ_L, although the gap is smaller for higher levels of insurance (fewer banks are liquidated). Consider the case $g = 1$ and $c_1 = 1.1$. For these parameters, $\theta_L = 0.3261$ and the equilibrium threshold equals $\theta_g^* = 0.3928$. Hence, for every θ in the region $[\theta_L + 2\varepsilon, \theta_g^* - \varepsilon[= [0.346\,1, 0.3828[\neq \varnothing$, all patient depositor continue running on a bank they know to be solvent. This region, however, is substantially smaller than the one without insurance, on page 70.

Blanket guarantees in times of crisis are usually designed to protect the principal value of deposits, in order to enhance market confidence and secure the purchasing power of consumers. In this model, a blanket guarantee would translate into $g = 1$. However, despite the high level of protection, depositors still run on some solvent banks. Notice that this result it is not in response to a lack of confidence in the deposit insurance system or to a macroeconomic shock affecting the economy. It emerges naturally as an equilibrium in a model with asymmetric information, where depositors rationally anticipate the reaction functions of their counterparts.

2.5 Equilibrium under a Broad Mandate

Under a broad mandate, and once the DIC and the bank have announced their respective contracts (g and c_1), all impatient depositors withdraw in the interim period and patient depositors, observing private signals on θ, decide whether to withdraw or to remain. The DIC moves after depositors have played, observing the realisation of n and its own private signal s, which comes from the monitoring of the bank's activities. Based on the information revealed by these two variables, the DIC must decide whether to leave the bank open – in which case liquidity assistance may sometimes be required – or to close it and pay the guarantee.

Figure 2-3.b gives a representation of the extended form of the sub-game faced by depositors and the DIC in the interim period.

2.5.1 DIC's Sub-game

In this version of the game, bankruptcy is not determined solely by the actions of depositors; it can also be the efficient outcome of supervision and prudential regulation. On a theoretical level, Repullo (2000) justifies the allocation of supervisory activities to the DIC every time withdrawals are large enough to pose a systemic threat: *"deposit insurance...institutions have become responsible for dealing with solvency problems, leaving Central Banks with the exclusive role of handling liquidity problems"*. As the present model does not study the problem of separation of activities between the central bank and the banking supervisor, for simplicity, the DIC will be allowed to deal with both solvency and liquidity issues.

Hence, as opposed to the case with a narrow mandate, this time the DIC has access to private information which can be used to decide a closure rule and a LoLR policy for banks. At $t = 1$ the DIC receives a private, non verifiable signal $s = \theta + \xi_s$, where ξ_s is uniformly distributed on $[-\xi, \xi]$. Given the supervisory role assigned to this agency, preferential access to information will naturally imply that $\xi \leq \varepsilon$; meaning that, on average, the DIC's signal is more informative than the signals of depositors. Knowing the value of s, the DIC corrects the conditional probability distribution of θ and estimates that the true value of θ follows a uniform distribution in the interval $[s - \xi, s + \xi]$.

Once patient depositors have played, the realisation of n becomes observable to all players, in particular to the DIC. As a function of θ (which is indeed a 1-1 relationship in the region of partial runs), n also carries information about the true state of the bank. For a given value of n, the DIC has the option of closing the bank based on its estimated solvency, or leaving it open, in which case liquidity assistance could be provided under exceptional circumstances.

The IMF's code of best practice requires the resolution of a failing bank to be decided according to a "*least cost*" criteria (Hoelscher and Quintyn, 2003). Using this idea, I define the objective function of the DIC at this stage of the game as that of minimising the cost of resolution of a bank.[25]

I will first study the case of perfect information ($\xi = 0$), so that s perfectly reveals

[25] Although this assumption still allows for very interesting results, a complete welfare analysis would require a more general definition of the DIC's objectives (see section 2.7.1).

the true value of θ in the interim period. Later, I will extend these results to the case of a noisy signal ($\xi > 0$).

Closure rule

When the DIC decides to close a bank, the guarantee on deposits must be paid. I am assuming deposits are senior to other claims, therefore, when a bank fails, all of its assets are passed onto the DIC and the agency has to decide how to manage available resources. Suppose the DIC has the option of issuing debt against the future value of the bank's assets, with an expected return of $E_s[R(\theta)]$.[26] Of course, the DIC can also liquidate the assets in the interim period, obtaining a certain return of 1. Hence, the net expected cost of the decision of closing a bank in the interim period equals

$$g - \max\{E_s[R(\theta)], 1\}.$$

As I am assuming the DIC to have perfect information, the term $E_s[R(\theta)]$ simplifies to $R(\theta)$.

Proposition 10 *When the DIC has perfect information ($\xi = 0$), it is cost efficient for it to close the bank if remaining assets are insufficient to cover the value of the guarantee in the final period ($c_2(\theta, n) < g$), and otherwise to leave it open.*

Proof. Assume $n \leq 1/c_1$ and compare the costs of the two actions in the interim period:[27]

Close	*Open*	
$g - \max\{R(\theta), 1\}$	0	*, if* $c_2(\theta, n) \geq g$
	$(1-n)g - (1-nc_1)R(\theta)$	*, if not*

Closing the bank means the guarantee has to be paid at a cost equal to $g - \max\{R(\theta), 1\}$. If the bank is left open, and remaining assets are enough to pay depositors at least the value of the guarantee in the final period, this action has no cost to the DIC. However, if funds in the bank are insufficient, the guarantee must be honoured at a cost equal to $(1-n)g - (1-nc_1)R(\theta)$.

A least cost criteria implies that if $c_2(\theta, n) \geq g$ it is better to leave the bank open. Indeed, this rule is Pareto optimal, even if $g - \max\{R(\theta), 1\} \leq 0$ (in which case the DIC

[26] Debt issuance will require the DIC to provide funds to pay the guarantee in the first period. Nonetheless, as the bank's assets have been transferred to it, the counterpart risk of this loan should be minimal.

[27] If $n > 1/c_1$ all assets would be liquidated in the interim period, in which case to allow the bank to operate until the second period would not be an option, unless it receives a loan from the central bank. However, a bail out would not be efficient in this case, because the bank is insolvent (see propositions 13, 18 and 19 later in this chapter).

should be indifferent as to which of the two actions, since its objective is to minimise the cost of bank resolution and not to make a profit from this operation). In order to see that, look at the welfare of depositors:

Close	*Open*
$u(g)$	$nu(c_1) + (1-n)u(c_2(\theta, n))$

Because $c_1 \geq g$ and $c_2(\theta, n) \geq g$, depositors are better off if the bank is allowed to survive to the final period.

On the other hand, if $c_2(\theta, n) < g$ closure is the least cost solution. Comparing the costs of the two actions:

$$g - \max\{R(\theta), 1\} \leq (1-n)g - (1-nc_1)R(\theta)$$
$$\Leftrightarrow ng \leq \max\{R(\theta), 1\} - (1-nc_1)R(\theta).$$

By contradiction, assume $ng > \max\{R(\theta), 1\} - (1-nc_1)R(\theta)$.

i. If $R(\theta) \geq 1 : ng > R(\theta) - (1-nc_1)R(\theta) = nc_1 R(\theta) \Leftrightarrow 1 \geq \frac{g}{c_1} > R(\theta)$ which is a contradiction.

ii. If $R(\theta) < 1 : ng > 1 - (1-nc_1)R(\theta) \Leftrightarrow R(\theta) > \frac{1-ng}{1-nc_1} \geq 1$, which is again a contradiction.

■

LoLR policy

According to Bagehot's doctrine, a LoLR should lend only to solvent banks experiencing liquidity problems (see section 1.4.3 in the previous chapter) . In the present framework,

Definition 11 *A bank faces a liquidity shock if $n > \pi$ at $t = 1$.*

i. *For a given realisation of n, a bank is* ***liquid*** *in both periods if $c_2(\theta, n) \geq c_1$. Notice that any liquid bank must also be solvent.*

ii. *A bank is* ***fundamentally solvent but illiquid*** *if $\theta \geq \theta_L$, $n > \pi$ and $c_2(\theta, n) < c_1$.*

Solvency is a property of the bank which cannot be verified until the final period. However, and as for the moment I am assuming $\xi = 0$, the DIC can perfectly observe the true value of θ in the interim period.

Clearly, if for a given value of n a bank were liquid, no assistance would be required. If it were fundamentally solvent but illiquid, however, there would be room for a LoLR.

Proposition 12 *When the DIC has perfect information ($\xi = 0$), a cost efficient LoLR policy is to rescue fundamentally solvent but illiquid banks.*

Proof. Suppose the bank is fundamentally solvent but illiquid.

$$c_2(\theta,\pi) = \frac{1-\pi c_1}{1-\pi} R(\theta) \geq c_1 \Leftrightarrow (1-\pi c_1)\, R(\theta) - (n-\pi)\, c_1 \geq (1-n)\, c_1.$$

If the DIC provides liquidity assistance for a maximum of $(n-\pi)\, c_1$ in the interim period, the inequality above establishes that the residual return when liquidating only πc_1 units in the interim period minus the repayment of the loan – at zero interest rate –, is enough to secure remaining patient depositors to receive a least c_1 at $t = 1$. In other words, lending money to a fundamentally solvent bank has zero cost for the DIC.

Committing liquidity assistance to fundamentally solvent but illiquid banks is indeed Pareto optimal, in terms of consumers' welfare:

i. If $c_2(\theta, n) < g$ it was argued before that the bank should be closed. However, comparing the welfare of depositors it is possible to see that the bank should be bailed out (as this policy has zero cost):

Close		*LoLR + Open*
$u(g)$	$<$	$nu(c_1) + (1-n)u(c_2^L)$
		where $c_2^L = c_2(\theta,\pi) \geq c_1 \geq g$

ii. If $g \leq c_2(\theta, n) < c_1$ it was established before that the bank should be allowed to survive until the final period. Comparing the welfare of depositors when just leaving the bank open against the situation where the DIC also provides liquidity assistance:

Open		*LoLR + Open*
$nu(c_1) + (1-n)u(c_2(\theta, n))$	$<$	$nu(c_1) + (1-n)u(c_2^L)$
		where $c_2^L \geq c_1 > c_2(\theta, n)$

■

Therefore, the DIC should commit liquidity assistance to solvent but illiquid banks, and such commitment should be public information in order to deter panic runs. Should the DIC, under some circumstances, also commit liquidity to insolvent banks? The answer is no, and it will be proved in what follows.

Define by $\underline{\theta}$ the value of the fundamentals satisfying $c_2(\underline{\theta}, \pi) = g$. Clearly, as $g < c_1$, $\underline{\theta} < \theta_L$. Consider the case where the bank is fundamentally insolvent and it is also facing a liquidity shock:

$$n > \pi,\ c_2(\theta, n) < g \text{ and } \max\{\underline{\theta}, 0\} \leq \theta < \theta_L.$$

As $c_2(\theta, n) < g$, the closure rule determines that it should be closed. However, as $\theta \geq \underline{\theta}$, $c_2(\theta, \pi) \geq g$ and by lending the excess withdrawal at $t = 1$, the DIC could secure a higher return for remaining patient depositors in the final period.

Proposition 13 *The DIC should never commit liquidity assistance to fundamentally insolvent banks.*

Proof. Compare the costs of the two policies for the DIC:

Close	*LoLR + Open*
$g - \max\{R(\theta), 1\}$	$(n-\pi)\, c_1 - \{(1-\pi c_1)\, R(\theta) - (1-n)g\}$

The net cost of bailing out the bank equals the cost of the loan minus whatever assets can be recovered from the bank in the final period.

Closure is the least cost solution if and only if

$$g - \max\{R(\theta), 1\} \leq (n-\pi)\, c_1 - \{(1-\pi c_1)\, R(\theta) - (1-n)g\}$$
$$\Leftrightarrow ng \leq (n-\pi)\, c_1 + \max\{R(\theta), 1\} - (1-\pi c_1)\, R(\theta).$$

By contradiction, assume $ng > (n-\pi)\, c_1 + \max\{R(\theta), 1\} - (1-\pi c_1)\, R(\theta)$.

i. If $R(\theta) \geq 1 : ng > (n-\pi)\, c_1 + R(\theta) - (1-\pi c_1) R(\theta)$

$\Leftrightarrow 0 > \underbrace{n(c_1 - g)}_{\geq 0} + \underbrace{\pi c_1 (R(\theta) - 1)}_{\geq 0}$ which is a contradiction.

ii. If $R(\theta) < 1 : ng > (n-\pi)\, c_1 + 1 - (1-\pi c_1) R(\theta)$

$\Leftrightarrow \underbrace{(1-\pi c_1)}_{>0}\underbrace{(R(\theta) - 1)}_{<0} > \underbrace{n(c_1 - g)}_{\geq 0}$, which is again a contradiction.

■

When combining the two policies (closure rule and LoLR) it is possible to conclude that, with perfect information, solvent banks are never allowed to fail, and it is only when an insolvent bank experiences large withdrawals in the first period – and large enough for remaining assets to be insufficient to cover the payment of the guarantee in the final period – that the bank is closed by the DIC. This result indicates that, despite the supervisory role assigned to this agency, depositors retain some monitoring power but also that the DIC lacks the commitment to close all insolvent banks in the interim period (if $g < c_1$). Indeed, the smaller the guarantee, the higher the share of insolvent banks that might be allowed to survive.

Summarising the previous results:

Proposition 14 *When the DIC has perfect information ($\xi = 0$), a cost efficient policy is to leave a bank open if it is fundamentally solvent or if its remaining assets are enough to pay the guarantee on deposits in the final period. Otherwise, the bank should be closed. The LoLR should lend only to fundamentally solvent banks facing liquidity shocks (table 2.3).*

Proof. The only part of this proposition that has not been proved yet is that a bank should be closed independent of the actions of consumers if $\theta < \underline{\theta}$ (provided this value is non negative). This result is immediate because $c_2(\theta, \pi) < g$ clearly implies that $c_2(\theta, n) < g$ for all $n \geq \pi$. Following propositions 10 and 13, the bank should be closed. ■

θ	**Second period return**	**Closure rule**	**LoLR policy**	**Observations**
$\theta < \underline{\theta}$	$c_2(\theta, \pi) < g$	Close	No	Bank fundamentally insolvent
$\underline{\theta} \leq \theta < \theta_L$	$c_2(\theta, n) < g$	Close	No	
	$c_2(\theta, n) \geq g$	Open	No	
$\theta \geq \theta_L$	$c_2(\theta, \pi) \geq c_1 \geq g$	Open	Yes	Bank fundamentally solvent, lend $(n - \pi)c_1$ iff $n > \pi$ and $c_2(\theta, n) < c_1$

TABLE 2.3: Closure and LoLR policies with a broad mandate and perfect information.

The assumption that the DIC lends at a discounted rate normalised to zero does not contradict other assumptions in the model – the safe technology return was also normalised to zero –, neither is it uncommon in the literature. Allen and Gale (1998), for example, study the problem of a central bank providing emergency liquidity assistance in a model where early liquidation of assets is costly, and they also normalise the lending interest rate to zero. Other authors have argued that a LoLR should lend at a high penalty rate, in order to stop public funds from being used to finance regular investment (see Bagehot (1873) and Repullo (2000)). However, as the present model considers only one representative bank, it cannot take into account interbank lending as an alternative source of liquidity for solvent banks, as they do.

Case of imperfect information ($\xi > 0$)

Having established these results, moving to the case of imperfect information is simple. Considering a positive but small amount of noise, all the previous equations in θ can be rewritten in terms of their conditional expected value, which will allow for computing cost efficient closure and LoLR policies.

Define by s_L the value of the signal satisfying $E_{s_L}[c_2(\theta, \pi)] = c_1$, $s^* = s^*(n)$ such that $E_{(s^*,n)}[c_2(\theta, n)] = g$, and $\underline{s}$ such that $E_{\underline{s}}[c_2(\theta, \pi)] = g$ (if the solution is non-negative, and zero otherwise). Notice that for any given value of n, these parameters are uniquely determined because $c_2(\theta, n)$ is increasing on θ.

Proposition 15 *For small non-negative values of the DIC's signal noise, a cost efficient policy is to leave the bank open if $s \geq s_L$ or if $s \geq s^*(n)$, and otherwise to close it.*

Proof. Same as above, replacing all expressions by their conditional expected values.

■

The DIC can anticipate that a bank is solvent if $s > \theta_L + \xi$, and insolvent if $s < \theta_L - \xi$. Because $c_2(\theta, \pi) < c_1$ if $\theta < \theta_L$, and $c_2(\theta, \pi) > c_1$ if $\theta > \theta_L$, it follows from the definition of s_L that $\theta_L - \xi \leq s_L \leq \theta_L + \xi$. Indeed, when $s > s_L$ the expected value of $c_2(\theta, \pi)$ is computed for larger values of θ, and because $c_2(\theta, \pi)$ is increasing on θ this implies that $E_s[c_2(\theta, \pi)] > c_1$. The opposite is true when $s < s_L$. Therefore, conditional upon its private information, the DIC estimates a bank to be solvent if $s \geq s_L$, and insolvent if $s < s_L$.

Proposition 16 *For small non-negative values of ξ, the DIC should commit liquidity assistance to a bank facing runs if and only if $s \geq s_L$ and $E_s[c_2(\theta, n)] < c_1$.*

Corollary 17 $\lim_{\xi \to 0} s_L(c_1) = \theta_L(c_1)$. *That is, when the information gathered by the DIC becomes extremely precise, only insolvent banks are allowed to fail.*

These results justify the principle of *"creative ambiguity"*: depositors cannot anticipate if the LoLR will provide liquidity assistance for a subset of the fundamentals ($\theta \in]\theta_L - \xi, \theta_L + \xi[$). Nevertheless, this is not a consequence of the LoLR randomising over its set of actions (playing an equilibrium in mixed strategies, as in Freixas et al. (1999)). The DIC's strategy is perfectly determined and rational but it is not observed by consumers due to asymmetric information.

Indeed, two kind of errors are possible when $\xi > 0$. With positive probability the DIC can mistakenly allow a solvent bank to fail (by refusing liquidity assistance), or else bail out an insolvent bank. However, the information contained in $n(\theta)$ has not yet been taken into account. If a monotone equilibrium threshold for depositors θ_g^* exists, n can be expressed as an invertible function of θ in the region of partial runs (that is, where $\pi < n < 1$).

Proposition 18 *If $\theta_g^* \leq \theta_L + \varepsilon$, $n(\theta, \theta_g^*) > \pi$ reveals the true value of θ at $t = 1$. In this case, no matter what the value of ξ is, the DIC determines its closure and LoLR policies as in the case of perfect information (proposition 14). If $\theta_g^* > \theta_L - \varepsilon$, the DIC also gets perfect information in the region of no runs ($n(\theta, \theta_g^*) = \pi$).*

Proof. Assume a monotone equilibrium threshold θ_g^* existed, such that $\theta_L - \varepsilon \leq \theta_g^* \leq \theta_L + \varepsilon$. A value of $n(\theta)$ consistent with this equilibrium is given by equation 2.1. Hence,

$$n = \begin{cases} \pi & \text{if } \theta_i > \theta_g^* \text{ for all } i, \quad \text{if } \theta > \theta_g^* + \varepsilon \geq \theta_L, \text{ the bank is solvent} \\ \in]\pi, 1[& \text{if } \theta_g^* - \varepsilon < \theta < \theta_g^* + \varepsilon, \quad n \text{ fully reveals } \theta \\ 1 & \text{if } \theta_i < \theta_g^* \text{ for all } i, \quad \text{if } \theta < \theta_g^* - \varepsilon \leq \theta_L, \text{ the bank is insolvent} \end{cases}$$

■

The next step is to determine that a monotone equilibrium does indeed exists for depositors.

2.5.2 Patient Depositors' Sub-game

Once a patient depositor has received a private signal on θ, she will construct beliefs about the behaviour of other players, the value of the last period return, and the action chosen by the DIC, upon which she will derive her optimal strategy.

Dominance regions: Whatever the relationship between the first and second period guarantee, an upper dominance region does exist for this game. Formally, a depositor can be sure that the bank is solvent if $\theta_i > \theta_L + \varepsilon \geq \theta_L + \xi$. Because $s_L \leq \theta_L + \xi$, she also knows that the bank will be bailed out if facing runs. All patient depositors receive signals above this value if $\theta > \theta_L + 2\varepsilon$, and anticipating that $c_2^L = c_2(\theta, \pi) \geq c_1$ whatever the value of n, they wait until the final period.[28] Hence, the *upper dominance region* for this game is given by $[\theta_L + 2\varepsilon, 1]$.[29]

If $\underline{\theta}$ exists ($c_2(\underline{\theta}, \pi) = g$), $R(.)$ increasing implies that $\underline{\theta} - \xi \leq \underline{s} \leq \underline{\theta} + \xi$. If the DIC's signal is below $\underline{s}$ the bank will be closed. Therefore, if $\forall i \ \theta_i < \underline{\theta} - \varepsilon$, a depositor knows the bank will be closed no matter what strategies are played by other players. All depositors will receive signals below this level if $\theta < \underline{\theta} - 2\varepsilon$. Hence, if $\underline{\theta} > 2\varepsilon$ the optimal reaction of depositors in this region will be to run, in which case a *lower dominance region* exists.[30]

In the intermediate region ($[\underline{\theta} - 2\varepsilon, \theta_L + 2\varepsilon]$), only sufficiently illiquid insolvent banks will be closed, that is, those for which $s < s^*(n)$. In this region, depositors' payoffs will depend upon their actions in the following way:[31]

[28] The upper index "L" in c_2^L stands for the "liquidity" assistance by the LoLR.

[29] In the game without insurance, additional assumptions were required for depositors to unconditionally remain for higher realisations of θ. With the DIC operating under a broad mandate, these assumptions are endogenised by means of the commitment of liquidity to solvent but illiquid banks.

[30] For low values of the guarantee, $\underline{\theta}$ might well not exist. However, if coordination induces $\theta_L - \varepsilon \leq \theta_g^*$, this should not be a problem (because $\theta_L > 2\varepsilon$) and indeed, as previously established, it will provide the DIC with high quality information to assess the solvency of banks through the information contained in $n(\theta, \theta_g^*)$.

[31] $n \geq 1/c_1 \Longrightarrow s < s^*(n)$. If $s \geq s^*(n)$ the bank will remain open and the depositors' final payoff will

	$n \geq 1/c_1$	$n < 1/c_1$	
		$s < s^*(n)$	$s \geq s^*(n)$
$t = 1$	g	g	c_1
$t = 2$	g	g	$\max\{c_2(n, \theta), g\}$

TABLE 2.4: Patient depositors' payoffs in the game with insurance under a broad mandate.

Consistent with the dominance regions, $n = \pi$ if $\theta > \theta_L + 2\varepsilon$ and $n = 1$ if $\theta < \underline{\theta} - 2\varepsilon$. In the intermediate region, the action taken by the DIC will be directly determined by the behaviour of patient depositors, which payoffs are described in table 2.4. The entry in the bottom right corner satisfies $E_s[c_2(n, \theta)] \geq g$.

Once again, $n(\theta, \theta^*)$ in equation 2.1 defines a belief consistent with these dominance regions and the uniform distribution of the noise.

As in the case with a narrow mandate, it is possible to prove that:

Proposition 19 *When the DIC operates under a broad mandate and offers insurance $g < c_1$ in both periods, there exists a unique equilibrium threshold θ_g^*, such that depositors remain if $\theta_i > \theta_g^*$ and otherwise withdraw. This threshold satisfies $\underline{\theta} - \varepsilon \leq \theta_g^* \leq \theta_L + \varepsilon$.*

Proof. Similar to proposition 6.

■

Corollary 20 $\lim_{\varepsilon \to 0} \theta_g^*(c_1) \leq \theta_L(c_1)$, *which says that as the noise of depositors' signals vanishes, they run only on insolvent banks.*

As $\xi \leq \varepsilon$, when ε tends to zero ξ must go to zero as well, and the outcome of the game with a broad mandate insurance agency achieves a social optimum, in the sense that liquidity assistance is targeted exclusively to solvent banks ($\lim_{\xi \to 0} s_L = \theta_L$) but is not used in equilibrium as they do not experience runs ($\lim_{\varepsilon \to 0} \theta_g^*(c_1) \leq \theta_L(c_1)$). Extremely insolvent banks are closed ($\lim_{\xi \to 0} \underline{s} = \underline{\theta}$), and those with enough funds to cover the payment of the final period guarantee are allowed to continue operating. All these results hold irrespective of the specific values of insurance provided, which in particular might imply the insurance policy to be less expensive under a broad mandate.

be higher than g. Anticipating that, depositors should wait, so $n < 1/c_1$, which is a contradiction.

2.6 Demand Deposit Contract with Insurance

(Bank's Planning Period Sub-Game)

The previous sections solved for the equilibrium strategy of depositors and the DIC taking as given the value of c_1, and without considering the effect that deposit insurance may have on the value of the demand deposit contract. Too generous a protection could generate moral hazard if, because of limited liability, banks choose excessively high values of c_1, making them fundamentally insolvent for most realisation of θ and so increasing the probability of runs (at least for high levels of the guarantee). If insurance were sufficiently generous for depositors to be willing to accept this contract, the DIC would most likely end up paying the guarantee in the interim period.

This section will derive a condition for the optimal value of c_1 under both mandates. Because the equations become intractable for positive values of the noise, I will solve for the limit case when ε tends to zero. An analytical solution is possible when abstracting from the effect that changes in c_1 have on the equilibrium threshold. In that case, it is possible to show that the effect of deposit insurance is to reduce the level of risk-sharing, this effect being stronger under a broad mandate. However, when considering the impact of changes in c_1 on the equilibrium threshold, comparative statics show that the net effect of insurance is to increase the value of c_1. Limited insurance can contain this effect up to a certain level, justifying the observed conduct of governments across the world who offer only partial insurance on deposits in normal times in order to encourage depositors to keep monitoring their banks.

2.6.1 Narrow Mandate DIC

At the planning stage, the bank calculates the value of c_1 that maximises the ex-ante expected utility of consumers:

$$\max_{c_1} Eu(c_1) = \int_{\theta_g^*+\varepsilon}^{1} \left\{ \pi u(c_1) + (1-\pi) u\left(\max\left\{ \frac{1-\pi c_1}{1-\pi} R(\theta), g \right\} \right) \right\} d\theta$$

$$+ \int_{\tilde{\theta}_g}^{\theta_g^*+\varepsilon} \left\{ n(\theta,\theta_g^*) u(c_1) + (1-n(\theta,\theta_g^*)) u\left(\max\left\{ \frac{1-n(\theta,\theta_g^*)c_1}{1-n(\theta,\theta_g^*)} R(\theta), g \right\} \right) \right\} d\theta$$

$$+ \int_{0}^{\tilde{\theta}_g} u(g)\, d\theta$$

For $\theta > \theta_g^* + \varepsilon$ all depositors remain and receive $\max\{c_2(\theta,\pi), g\}$. When $\theta < \theta_g^* - \varepsilon$, all withdraw, the bank goes bankrupt and depositors are paid the guarantee. In the intermediate region, depositors' withdrawals are decreasing on the value of the fundamentals according to the formula of $n(\theta, \theta_g^*)$, and the guarantee is paid each time the bank's resources are insufficient to cover demanded deposits (i.e. when $\theta < \widetilde{\theta}_g$).

When $\varepsilon \to 0$ the region $[\theta_g^* - \varepsilon, \theta_g^* + \varepsilon]$ converges on $\{\theta_g^*\}$, and as payoffs in this region are discontinuous at $\widetilde{\theta}_g \longrightarrow \theta_g^*$, the problem reduces to:

$$\begin{aligned}\max_{c_1} Eu(c_1) = & \int\limits_{\theta_g^*(c_1)}^{1} \{\pi u(c_1) + (1-\pi) u\left(\max\left\{c_2\left(\theta, \pi\right), g\right\}\right)\} d\theta \\ & + \lim_{\varepsilon \to 0^+} \left\{n(\theta, \theta_g^*) u(c_1) + (1 - n(\theta, \theta_g^*)) u(\max\left\{c_2\left(\theta_g^*, n(\theta, \theta_g^*)\right), g\right\})\right\} \\ & - \lim_{\varepsilon \to 0^-} \left\{n(\theta, \theta_g^*) u(g) + (1 - n(\theta, \theta_g^*)) u(g)\right\} + \int\limits_{0}^{\theta_g^*(c_1)} u(g) d\theta\end{aligned}$$

It was established before that when $\varepsilon \to 0$, $\theta_g^*(c_1) > \theta_L(c_1)$, hence the term under the first integral $\max\{c_2(\theta, \pi), g\} = c_2(\theta, \pi) \geq c_1 \geq g$. Also, $\lim_{\varepsilon \to 0^+} n(\theta, \theta_g^*) = \pi$ and $\lim_{\varepsilon \to 0^-} n(\theta, \theta_g^*) = 1$, and the previous expression becomes:

$$\begin{aligned}\max_{c_1} Eu(c_1) = & \int\limits_{\theta_g^*(c_1)}^{1} \{\pi u(c_1) + (1-\pi) u\left(c_2\left(\theta, \pi\right)\right)\} d\theta \\ & + \pi u(c_1) + (1-\pi) u(c_2\left(\theta_g^*, \pi\right)) - u(g) + \int\limits_{0}^{\theta_g^*(c_1)} u(g) d\theta\end{aligned}$$

A sufficient condition for the value of c_1 maximising this function is given by equation 2.4:

$$\begin{aligned}& \pi \int\limits_{\theta_g^*(c_1)}^{1} \{u'(c_1) - R(\theta) u'\left(c_2(\theta, \pi)\right)\} d\theta + \pi \left\{u'(c_1) - R(\theta_g^*(c_1)) u'\left(c_2(\theta_g^*(c_1), \pi)\right)\right\} = \\ & \frac{\partial \theta_g^*(c_1)}{\partial c_1} \left[\pi u(c_1) + (1-\pi) u\left(c_2(\theta_g^*(c_1), \pi)\right) - (1 - \pi c_1) R'(\theta_g^*(c_1)) u'\left(c_2(\theta_g^*(c_1), \pi)\right) - u(g)\right]\end{aligned} \tag{2.4}$$

Hence, the marginal gain from risk-sharing due to the transfer of consumption from patient to impatient depositors (LHS) equals the positive marginal cost associated with the increase in the probability of runs (RHS).

Proposition 21 *When abstracting from the effect that changes in the demand deposit contract have on the equilibrium threshold of depositors* $\left(\frac{\partial \theta_g^*(c_1)}{\partial c_1} = \frac{\partial \theta^*(c_1)}{\partial c_1} = 0\right)$, *the introduction of deposit insurance reduces the level of risk-sharing offered by banks:* $c_1^g < c_1^{ng}$.[32]

[32] c_1^g *stands for the equilibrium value of the demand deposit contract in the interim period in the model with insurance, and* c_1^{ng} *for the one without insurance.*

See proof in the Appendix.

Given that the bank's portfolio has not changed, the reduction in the interest paid on deposits could be explained as the bank *"free riding"* on the reduction of risk on deposits that results from the DIC's guarantee.

The effect of c_1 on the equilibrium threshold, however, is not nil. For the case without insurance, theorem 1 established that $\frac{\partial \theta^*(c_1)}{\partial c_1} > 0$. For the case with insurance under a narrow mandate, proposition 8 showed that $\theta_g^*(c_1)$ is increasing on c_1 if and only if

$$u(g) > u(c_1) - c_1^2 \int_{\pi}^{1/c_1} \underbrace{\left\{ u'(c_1) - u'(c_2(n)) \frac{\partial c_2}{\partial c_1} \mathbb{1}_{\{n \leq \hat{n}\}} \right\}}_{>0} dn.$$

This inequality is clearly resolved when $g = c_1$ then, by continuity, the result extends to a small neighbourhood of values of $g < c_1$. For the parameters of my numerical example, simulations for different values of g show that indeed $\frac{\partial \theta_g^*(c_1)}{\partial c_1} > 0$ (see figure 2-7).

Define by $\Psi(c_1, g) = 0$ the equation implicitly defined by the optimal condition of this problem, equation 2.4. Notice that $Eu(c_1)$ is quasi-concave on c_1, as it is the composition of monotonic and concave functions. Therefore, given that c_1^g maximises $Eu(c_1)$, $\frac{\partial \Psi}{\partial c_1}(c_1^g) = SOC(c_1^g) \leq 0$.

Partial differentiation of equation 2.4 with respect to g gives $\frac{\partial \Psi}{\partial g} = \frac{\partial \theta_g^*(c_1)}{\partial c_1} u'(g) > 0$. Hence, by the implicit function theorem $\frac{\partial c_1^g}{\partial g} = -\frac{\frac{\partial \Psi}{\partial c_1}(c_1^g)}{\frac{\partial \Psi}{\partial g}} \geq 0$, which justifies the intuitive idea that as $g \leq c_1$, the higher the protection to depositors the less liable are banks for losses, and the higher the interest rate they offer (higher c_1).

Under a narrow mandate, as in the benchmark model, the probability of failure in the interim period is given by

$$p \doteq prob\{n > 1/c_1\} = \theta_g^* + \frac{\varepsilon}{1-\pi}\left(1 + \pi - \frac{2}{c_1}\right).$$

Hence, if g is sufficiently high for $\frac{\partial \theta_g^*}{\partial c_1} > 0$, the probability of failure in the interim period is increasing on c_1 :

$$\frac{\partial p}{\partial c_1} = \frac{\partial \theta_g^*}{\partial c_1} + \frac{2\varepsilon}{(1-\pi)\, c_1^2} > 0$$

On the other hand, for small values of g (limited protection) proposition 8 could be violated, so that $\frac{\partial \theta_g^*}{\partial c_1} < 0$ in which case $\frac{\partial c_1^g}{\partial g} < 0$ and the probability of runs could decrease with the level of insurance:

$$\frac{\partial p}{\partial g}=\frac{\partial \theta_g^*}{\partial g}+\frac{2\varepsilon}{(1-\pi)\,c_1^2}\frac{\partial c_1^g}{\partial g}<0$$

These results can be summarised in the following proposition.

Proposition 22 *High levels of deposit insurance increase the equilibrium level of risk-sharing offered by banks, $c_1^g > c_1^{ng}$, and therefore the probability of runs. Partial insurance can limit this effect.*

Figure 2-8 plot the FOC for the determination of the optimal demand deposit contract, as described by equations 2.4 and 2.5 (in the Appendix). Without insurance, the level of risk-sharing offered is very small and actually very close to 1. When offering full principal insurance ($g = 1$), the *FOC* is positive and decreasing for the values of c_1 considered in this simulation, implying the optimal level of risk-sharing will be higher.

Limited liability and a high level of insurance increase the incentives for the bank to offer higher values of c_1, rising the probability of failure. This result is consistent with empirical evidence showing that an increase in the volume of insured funds is usually accompanied by a sharp rise in interest rates (e.g. in the Savings and Loan crisis). Partial insurance can limit this effect, as seen in figure 2-9, while improving on intertemporal risk-sharing compared to the non-insurance case.

2.6.2 Broad Mandate DIC

Consider again $\varepsilon \longrightarrow 0$, which implies that ξ also vanishes.

From section 2.5, $\lim_{\xi\to 0} s_L(c_1)=\theta_L(c_1)$, $\lim_{\xi\to 0}\underline{s}(c_1)=\underline{\theta}(c_1)$ and $\underline{\theta}\le \lim_{\varepsilon\to 0}\theta_g^*(c_1)\le\theta_L(c_1)$. Thus, in the region $\underline{\theta}(c_1)\le\theta\le\theta_L(c_1)$, as $\lim_{\varepsilon\to 0^+} n(\theta,\theta_g^*)=\pi$ and $\lim_{\varepsilon\to 0^-} n(\theta,\theta_g^*)=1$, the closure of insolvent banks is strictly determined by the actions of depositors.

The problem faced by the bank in the planning period is the same as before:

$$\max_{c_1} Eu(c_1)=\int_{\theta_g^*(c_1)}^{1}\{\pi u(c_1)+(1-\pi)u\left(c_2\left(\theta,\pi\right)\right)\}\,d\theta$$
$$+\ \pi u(c_1)+(1-\pi)u(c_2\left(\theta_g^*,\pi\right))-u(g)+\int_0^{\theta_g^*(c_1)} u(g)d\theta,$$

and a sufficient condition for the optimal value of c_1 is given by equation 2.4.

As mentioned in section 2.2.1, when ε equals zero inefficient liquidation (panic runs) occurs in the model without insurance for any value of $c_1 > 1$. Therefore, $\theta^*(c_1)>\theta_L(c_1)\ge\theta_g^*(c_1)$.

Knowing this relationship, when abstracting from the effect that changes in the demand deposit contract have on the equilibrium threshold of depositors

$\left(\frac{\partial\theta_g^*(c_1)}{\partial c_1} = \frac{\partial\theta^*(c_1)}{\partial c_1} = 0\right)$, the introduction of a deposit guarantee reduces the level of risk-sharing offered by banks: $c_1^{ng} > c_1^g$. However, the shifting is stronger in this case.[33]

Hence, the effect of deposit insurance on the optimal demand deposit contract can be decomposed in two factors. When $\frac{\partial\theta_g^*}{\partial c_1} = 0$ its effect is to reduce the level of risk-sharing offered by banks with respect to the case without an explicit guarantee. When the impact of changes in c_1 on the probability of runs is included, the opposite effect is observed and c_1^g increases in an amount related to the level of the promised protection. In the case of a broad mandate, the net effect of insurance is expected to be a rise in c_1^g, albeit smaller than in the case of a narrow mandate, as the first factor (a shift to the left of c_1^{ng}) is stronger.

2.7 Robustness

I now go on to discuss some policy implication arising from the model, and how the results might change under alternative specifications and assumptions.

2.7.1 Deposit Insurance Contract (DIC's Planning Period Sub-game)

The choice of the optimal amount of coverage is a relevant issue not discussed in this model, as the value of g was considered an exogenous parameter influencing the outcome of the game. The IMF typically suggests the world average of per capita GDP as a rough rule of thumb for adequate coverage. In practice, however, coverage limits vary widely from country to country. For example, the percentage of the value of deposits covered is almost negligible in Sri Lanka and Estonia, and only about 10 percent in Brazil and Tanzania; but above 65 percent in the USA, and more than 70 percent in Norway, India and Japan (see statistical Appendix in Garcia (2000)).

A complete definition of the game (and therefore a full welfare analysis of deposit insurance) would need to specify a payoff function for the DIC, in order to determine the optimal level of insurance offered at the planning period. Such a function should include the effect that a deposit guarantee has on both the depositors' equilibrium threshold and the demand deposit contract offered by banks.

[33] Denote by $\theta_g^N(c_1)$ the equilibrium threshold under a narrow mandate and by $\theta_g^B(c_1)$, the one under a broad mandate. Because $\theta_g^N(c_1) \geq \theta_L(c_1) \geq \theta_g^B(c_1)$, from the proof of proposition 21 in the appendix it can be seen that the integration region in $LHS_{(2.4)}(c_1)$ is larger in the latter case and includes smaller values of θ, where the function $\varpi(\theta, c_1)$ is negative. Hence, an even smaller value of c_1 is needed in order to make $LHS_{(2.4)}(c_1^g) = 0$.

The objectives of deposit insurance usually include the protection of depositors and financial stability concerns. Garcia (2000) argues that because many deposit insurance schemes include all deposit-taking institutions, consumer protection is a number one concern. Financial stability, in his opinion, would be the main concern if membership were confined only to systemically important banks. Taking this author's point of view, the DIC should determine g in order to maximise the ex-ante expected utility of consumers, while considering a funding constraint (money raised through an ex-post tax, in the case of this model), and the effect that deposit insurance has on the risk of runs and on moral hazard. On the other hand, Rochet (1999) notices that prudential authorities themselves tend to insist more on the prevention of systemic risks as the main issue. Under this view, g should be chosen in order to minimise the probability of joint failure of systemic institutions.

Coverage limits are in practice determined by very complex political processes. In the USA, for example, it has been discussed whether deposit insurance should be indexed to living costs.[34] Indeed, the choice of the limit coverage could become time-inconsistent when facing systemic crises (the ex-ante chosen value of the guarantee could become ex-post inefficient), particularly so if the insurer has a narrow mandate or if its signal is too noisy under a broad mandate. In either case, some solvent banks would go bankrupt, weakening even more the financial system. The authorities could then decide to temporarily increase the guarantee in order to contain runs, however hard reducing it later to contain moral hazard (Garcia, 2000).

2.7.2 Credibility and Blanket Guarantees

Blanket guarantees are usually designed in times of crisis to protect the principal value of deposits, in order to enhance market confidence and secure the purchasing power of consumers. When $g = c_1$ runs are completely eliminated but so is market discipline. This is why deposit guarantees usually cover the principal value of deposits and not the interest accrued on them. The results in this chapter show, however, that if $g < c_1$ and the agency in charge of insurance does not get involved in the supervision of banks, depositors will still run on some solvent banks, the more so the lower the guarantee. This result is not in response to a lack of confidence in the insurance scheme, nor to a macroeconomic shock affecting the economy but it emerges naturally as an equilibrium in a model with asymmetric information, where depositors rationally anticipate the

[34] See *Alan Greenspan: Deposit Insurance.* Testimony before the Committee on Banking, Housing, and Urban Affairs, U.S. Senate, Washington, D.C., 26 February 2003. BIS Review 10/2003.

reactions of their counterparts.

For a deposit guarantee scheme to be credible and operational, initial funding is required. The IMF code of best practice establishes that the funding of a deposit insurance scheme should be adequate and perceived as sufficient to maintain public confidence and that, upon failure, legal priority over assets should be given to the DIC on behalf of depositors, as this model assumes.

Some countries run schemes with an ex-ante funding, charging participant institutions an insurance premium. Others, as in this model, run ex-post funded systems. Whatever scheme is applied, and given the scope of the losses involved when a bank fails, for insurance to be credible government backing may be needed. Indeed, in the majority of countries with explicit systems, while deposit insurance is privately funded by their member institutions, some implicit or explicit government backing always exists.

2.7.3 Macroeconomic Shocks

If, in the interim period, depositors were uncertain about the available funds for the payment of the guarantee in the subsequent period, they could precipitate a run. Consider $g < c_1$, the value of the guarantee in the interim period, and $g_2 < g$ its expected value in the final period. It is not difficult to prove that a monotonic equilibrium exists for this game, and that the equilibrium threshold is decreasing on g_2. Therefore, the smaller the expected insurance in the final period, the more banks would be suffering from runs. Such a situation could arise because of fears of asset depletion (corruption), macroeconomic shocks, or an attack on the currency.

If the economy experienced a macroeconomic shock in the interim period, reducing future returns from $R(\theta)$ to, let us say, $R(\theta) - k$, the effect would be twofold. First, the solvency threshold would move upwards, increasing the probability of runs. With more banks suddenly becoming insolvent, the DIC might not have enough resources to cover the payment of the guarantee and, if this were anticipated, more patient depositors would withdraw. Second, the promise of liquidity assistance by a lender of last resort would probably not be enough to overcome this effect, as fewer bank would be bailed out in a model with a broad mandate because expected returns are lower.

2.7.4 Twin crisis

The present model considers only one representative bank, therefore, a run on the bank could be identified with a run on the currency. If not enough funds can be readily available for the payment of deposit insurance, its expected value in the last period (g_2, as in section 2.7.3) might be reduced, raising the equilibrium threshold and therefore runs. Increasing pressure to cover guaranteed deposits could force the government into devaluation, in which case the value of the guarantee in the final period would be further reduced, generating a spiral reaction.

Devaluation could be introduced in this model as in Repullo (2000), where the government provides liquidity assistance (broad mandate) in the form of bonds that can be traded for consumption goods in the interim period, which determines their equilibrium price level. Doubts that the guarantee could be paid in the final period would precipitate a collapse in the price of these bonds, which could be interpreted as a currency crisis. Devaluation could also be modelled as in Chang and Velasco (1999), by linking the equilibrium exchange rate of the economy to the same fundamentals of the bank.

2.7.5 Moral Hazard

Many authors agree that deposit insurance induces moral hazard. In fact, a guarantee on deposits can be seen as a callable put option on the agency offering insurance (Merton, 1977; Acharya and Dreyfus, 1989), which value increases monotonically in the volatility of the investment portfolio, and then is maximised at the highest possible level of risk.

In the model introduced in this chapter is not possible to study moral hazard as commonly understood, that is, a risk shifting in the bank investment strategies. Given that by assumption there is only one dominant risky technology, banks cannot shift to riskier investment projects as a result of the introduction of insurance, simply because these projects are not available.

A different source of moral hazard is proposed by Bond and Crocker (1993), linking deposit insurance to the level of capitalisation of banks. Their model concentrates on the effect that insurance pricing has on the level of optimal reserves. In the present model banks do not hold reserves because of the assumptions about the risky technology (that its return in the interim period equals that of the safe asset). One possibility for studying this phenomenon would be to consider higher costs of early liquidation, in the sense that only a fraction $\mu < 1$ of the original investment could be recovered at $t = 1$. As in that case banks would need to keep reserves to pay impatient depositors withdrawals in the interim period, it would be possible to study how the introduction of a guarantee on deposits would shift the composition of the investment portfolio.

2.7.6 Closure Rule

Strictly speaking, closure is only one of the possible options for dealing with failed banks. Other options involve the intervention or takeover of the institution (transferring the control of the bank to the DIC), the merger or sale of the bank to a stronger institution, purchase an assumption, or a bridge bank for the administration of good assets. In the present model with a single representative bank, many of these options are not viable.

Acharya and Yorulmazer (2007a,b) show that the acquisition of failed banks by stronger institutions can be welfare improving, and even make systemic crisis less likely by reducing the correlation between banks' portfolios. The present model could be extended to study this type of policy in the following way. Consider two ex-ante identical banks that ex-post differ only on the realization of θ. If one bank became insolvent (Bad) while the other is solvent (Good), the DIC (broad mandate) could consider merging both if their combined return were enough to pay patient depositors in both banks at least c_1 in the final period:

$$E_{s_1,s_2}[(1-\pi c_1)R(\theta_G)+R(\theta_B)] \geq \pi c_1 + 2(1-\pi)c_1,$$

where s_j is the DIC's private signal on bank j.

This would require the DIC to provide liquidity assistance to the stronger institution, in order to satisfy patient depositors' redemptions in the failed bank. This would be a different form of LoLR assistance from the one proposed in Acharya and Yorulmazer (2007a,b). Given the banks have no equity; the "purchase" would effectively be a transfer of assets from the failed bank to the solvent one (if viable).

The equilibrium for the depositors' sub-game will depend on the structure of the signals they receive. If they obtained two equally informative signals (for example, if they held deposits in both banks), anticipating the merger patient depositors in the insolvent institution would prefer to wait. On the other hand, if depositors could monitor only one bank the equilibrium would depend on the order of the game, that is, which bank nature chooses to play first, as in Dasgupta (2002). Otherwise, the DIC would need to act pre-emptively, based solely on its signals and therefore missing the information contained in the number of runs on each bank.

2.7.7 Other Extensions

Deposit guarantees are designed to protect small and usually uninformed depositors. This gives a trade-off because more sophisticated depositors tend not to be covered, and could exercise monitoring power independent of the DIC. Including this type of players in the game could provide an additional explanation for the failure of blanket guarantees sometimes.

Finally, it would be interesting to compare ex-ante versus ex-post funded systems. I have chosen to discuss an ex-post tax funded system, as in Diamond and Dybvig (1983). In practice, some countries do run ex-post funded systems, charging a fee to surviving institutions participating in the scheme (e.g. in the U.K.). If an ex-ante premium were charged to the bank, the equations of the model would be modified. I expect, however, that the main results would not change.

2.8 Concluding Remarks

In this chapter I have introduced deposit insurance in a model of information-based bank runs. The model has a unique equilibrium, which allows for a proper evaluation of the effects of insurance on the behaviour of depositors, banks and the insurer. I have shown that, while consumers achieve better risk-sharing in a competitive banking system than in autarky, more solvent projects are liquidated as uninsured depositors fail to coordinate in a subset of fundamentals and run on banks they know to be solvent. While deposit insurance may prevent panic runs up to some extent, its effectiveness varies with the size of coverage and the degree of supervisory involvement of the agency in charge of insurance. I have considered two possible mandates. Under a narrow mandate, and abstracting from the presence of any other regulatory authority, the insurer's main objective is to pay the guarantee every time a bank has insufficient resources to cope with withdrawals. Under a broad mandate, the insurer also has responsibility for the resolution of insolvent and/or illiquid banks, and the ability to provide emergency liquidity assistance as a lender of last resort.

Under a narrow mandate, a deposit insurance contract preserving the monitoring role of depositors involves offering less than full protection. The trade-off is that panic runs cannot be completely eliminated with a partial guarantee, although it does reduce the region of fundamentals for which that occurs. Under a broad mandate, I showed that panic runs tend to disappear for any level of insurance as the regulator's signal becomes more precise. Given that liquidity assistance is committed to solvent but illiquid institutions, depositors do not run on solvent banks. Moreover, it is cost efficient for the authority never to provide liquidity to insolvent banks. However, only extremely insolvent banks are closed, and those with enough funds to cover the payment of the final period guarantee are allowed to continue in operation. Therefore, the smaller the protection offered to depositors, the higher is forbearance. All these results hold irrespective of the specific values of the guarantee, which in particular might imply that the social cost of deposit insurance is lower under a broad mandate.

Finally, I showed that deposit insurance increases the equilibrium value of c_1 and so the probability of runs, at least for high levels of the guarantee, but this effect seems also to be smaller under a broad mandate. Limited insurance could contain this externality up to some level, justifying the observed conduct of governments across the world in normal times.

Under a narrow mandate, pure panic runs persist even when depositors' signals become very precise, the more so the lower the guarantee. Under a broad mandate, on the other hand, panic runs can be eliminated even with partial insurance but forbearance increases. Which one should be preferred?

Both mandates are equally popular among economies. In their survey, Demirguç-Kunt and Detragiache (1999) report that 34 out of 67 deposit insurance systems have a narrow constitution but they also show that the negative externalities imposed by deposit insurance on financial stability can be curbed by effective regulation, a result in line with the main conclusions of this chapter. Indeed, during recent years some countries (e.g. France) have started to move from narrow to broad mandate schemes (Garcia, 2000). Moreover, the costs of forbearance need to be compared to the benefits of a more stable financial system when panic runs threaten it.

Therefore, a scheme where the DIC has more supervisory involvement (broad mandate), or else a high degree of coordination with the authority in charge of supervision, should be preferred.

2.9 Appendix

2.9.1 Proof of the existence of a unique monotone equilibrium in the case without insurance

Proof.

This follows an easier version of the proof, proposed by Dasgupta (2002). Notice that $\Delta(\theta_i, \theta^*)$, being the composition of continuos functions in θ_i and θ^*, is continuous in both arguments.

Take $\theta_i = \theta_L - \varepsilon$:

$$\Delta(\theta_L - \varepsilon, \theta_L - \varepsilon) = \int_{\theta_L - 2\varepsilon}^{\theta_L} \tfrac{1}{2\varepsilon}\delta(\theta, \theta_i) d\theta.$$

Notice that this integral is well defined as $\theta_L - 2\varepsilon > 0$.

$$\Delta(\theta_L - \varepsilon, \theta_L - \varepsilon) = \int_{\theta_L - 2\varepsilon}^{\widetilde{\theta}_L} \tfrac{1}{2\varepsilon} \{u(0) - u(1/n)\} d\theta + \int_{\widetilde{\theta}_L}^{\theta_L} \tfrac{1}{2\varepsilon} \{u(c_2(\theta, n)) - u(c_1)\} d\theta,$$

where $\widetilde{\theta}_L$ is such that $n(\widetilde{\theta}_L, \theta_L - \varepsilon) = 1/c_1$. The first integral is clearly negative. The second one is also negative, simply notice that $\widetilde{\theta}_L < \theta < \theta_L$ implies that $n > \pi$ and then $c_2(\theta, n) < c_1$. Hence, $\Delta(\theta_L - \varepsilon, \theta_L - \varepsilon) < 0$.

Now take $\theta_i = \theta_U + \varepsilon$:

$$\Delta(\theta_U + \varepsilon, \theta_U + \varepsilon) = \int_{\theta_U}^{\widetilde{\theta}_U} \tfrac{1}{2\varepsilon} \{u(0) - u(1/n)\} d\theta + \int_{\widetilde{\theta}_U}^{\theta_U + 2\varepsilon} \tfrac{1}{2\varepsilon} \{u(c_2(\theta, n)) - u(c_1)\} d\theta,$$

well defined given that $\theta_U + 2\varepsilon < 1$.

The first integral is again negative but I have assumed that for $\theta > \theta_U$ the return of the risky technology is extremely high, therefore the second integral should overcome the first one ($n < 1/c_1$ when $\theta > \widetilde{\theta}_U$) to make $\Delta(\theta_U + \varepsilon, \theta_U + \varepsilon) > 0$.

Hence, by continuity there exists $\theta^* \in]\theta_L - \varepsilon, \theta_U + \varepsilon[$ such that $\Delta(\theta^*, \theta^*) = 0$.

This equality establishes that the positive and negative parts of the integral exactly offset each other. Therefore, raising θ_i above θ^* increases the positive part, so that $\Delta(\theta_i, \theta^*) > 0$ if $\theta_i > \theta^*$; while lowering θ_i below θ^* increases the negative part, and then $\Delta(\theta_i, \theta^*) < 0$ if $\theta_i < \theta^*$.

That $\Delta(\theta_i, \theta_i)$ is monotonic increasing can be proved by total differentiation with respect to θ_i, and this property determines the uniqueness of the equilibrium. ■

2.9.2 Proof of proposition 5

Proof.

Consider θ^1, $\theta^2 \in [0,1]$ such that $\theta^1 < \theta^2$. I want to prove that $\Delta_g(\theta^1,\theta^1) < \Delta_g(\theta^2,\theta^2)$.

$$\Delta_g(\theta^j,\theta^j) = \int_{\widetilde{\theta}^j}^{\widehat{\theta}^j} \{u(g) - u(c_1)\}\, d\theta + \int_{\widehat{\theta}^j}^{\theta^j+\varepsilon} \left\{ u\left(\frac{1-n(\theta,\theta^j)c_1}{1-n(\theta,\theta^j)} R(\theta) \right) - u(c_1) \right\} d\theta, \ j = 1,2,$$

where $\widetilde{\theta}^j$ is defined by $n(\widetilde{\theta}^j,\theta^j) = 1/c_1$ and satisfies $\widetilde{\theta}^j = \theta^j + \frac{\varepsilon}{1-\pi}\left(1+\pi-\frac{2}{c_1}\right)$; and $\widehat{\theta}^j$ is such that $c_2(\widehat{\theta}^j, n(\widehat{\theta}^j,\theta^j)) = g$, $\widetilde{\theta}^j \leq \widehat{\theta}^j \leq \theta^j + \varepsilon$.

$\widehat{\theta}^j$ is uniquely defined, as in the interval $[\theta^j - \varepsilon, \theta^j + \varepsilon]$, $c_2(\theta, n(\theta,\theta^j))$ is strictly increasing on θ. For the same reason it is true that $\widehat{\theta}^1 - \widetilde{\theta}^1 > \widehat{\theta}^2 - \widetilde{\theta}^2$, as when θ^j is higher a relatively smaller value of $\widehat{\theta}^j$ is required for $c_2(\widehat{\theta}^j, n(\widehat{\theta}^j,\theta^j)) = g$.

Finally, notice that from the definition of $n(\theta,\theta^j)$, if $\theta \in [\theta^1 - \varepsilon, \theta^1 + \varepsilon]$, $\theta' \in [\theta^2 - \varepsilon, \theta^2 + \varepsilon]$, and $\theta - \theta^1 = \theta' - \theta^2$ then $n(\theta,\theta^1) = n(\theta',\theta^2)$. This basically establishes that over the intervals $[\widetilde{\theta}^j, \theta^j + \varepsilon]$ $j = 1,2$, the functions $n(\theta,\theta^j)$ take exactly the same values.

This information is sufficient to conclude that the function $\Delta_g(\theta^*,\theta^*)$ is increasing. Although the first integral is higher for $j = 1$ (because the function is constant and the integration region is larger), the argument under the second integral is increasing on θ – remember that n takes the same values in both regions – and the integration region is larger and covers higher values of the fundamentals for $j = 2$. Hence, the loss in the former is compensated by the gain in the latter when $j = 2$ ($c_2(\theta, n(\theta,\theta^2)) > c_2(\theta, n(\theta,\theta^1)) > g$), implying that $\Delta_g(\theta^2,\theta^2) > \Delta_g(\theta^1,\theta^1)$. ∎

2.9.3 Proof of proposition 8

Proof.

Define

$$\Phi(\theta_g^*, c_1, g) \doteq \int_{\pi}^{\widehat{n}} \{u\left(c_2(n)\right) - u(c_1)\}\, dn + \int_{\widehat{n}}^{1/c_1} \{u\left(g\right) - u(c_1)\}\, dn = 0,$$

where $c_2(n) = \dfrac{1 - nc_1}{1 - n} R\left(\theta(n)\right)$ and $c_2\left(\widehat{n}\right) = g$.

1. $$\frac{\partial \Phi}{\partial \theta_g^*} = \int_{\pi}^{\widehat{n}} u'(c_2(n)) \left(\frac{1 - nc_1}{1 - n}\right) R'(\theta(n)) dn + \frac{\partial \widehat{n}}{\partial \theta_g^*} \{u\left(g\right) - u(c_1)\} - \frac{\partial \widehat{n}}{\partial \theta_g^*} \{u\left(g\right) - u(c_1)\} > 0,$$

because by assumption $u(.)$ and $R(.)$ are increasing.

$$\frac{\partial \Phi}{\partial c_1} = \int_{\pi}^{\widehat{n}} \left\{u'(c_2(n)) \frac{\partial c_2}{\partial c_1} - u'(c_1)\right\} dn + \frac{\partial \widehat{n}}{\partial c_1} \{u\left(g\right) - u(c_1)\} - \int_{\widehat{n}}^{1/c_1} u'(c_1) dn + \frac{\partial \left(1/c_1\right)}{\partial c_1} \{u\left(g\right) - u(c_1)\} - \frac{\partial \widehat{n}}{\partial c_1} \{u\left(g\right) - u(c_1)\},$$

where $\dfrac{\partial c_2}{\partial c_1} = \dfrac{-n}{1 - n} R\left(\theta(n)\right) < 0$.

Simplifying,

$$\frac{\partial \Phi}{\partial c_1} = \int_{\pi}^{\widehat{n}} u'(c_2(n)) \frac{\partial c_2}{\partial c_1} dn - \int_{\pi}^{1/c_1} u'(c_1) dn + \left(\frac{1}{c_1}\right)^2 \{u(c_1) - u(g)\}.$$

By the Implicit Function theorem,

$$\frac{\partial \theta_g^*}{\partial c_1} = -\frac{\frac{\partial \Phi}{\partial c_1}}{\frac{\partial \Phi}{\partial \theta_g^*}} > 0 \Leftrightarrow \frac{\partial \Phi}{\partial c_1} < 0$$

$$\Leftrightarrow u(g) > u(c_1) - c_1^2 \int_{\pi}^{1/c_1} \underbrace{\left\{u'(c_1) - u'(c_2(n)) \frac{\partial c_2}{\partial c_1} \mathbb{1}_{\{n \leq \widehat{n}\}}\right\}}_{>0} dn,$$

where $\mathbb{1}_{\{n \leq \widehat{n}\}} = \begin{cases} 1 & if\ n \leq \widehat{n} \\ 0 & \sim \end{cases}$.

2. $$\frac{\partial \Phi}{\partial g} = \frac{\partial \widehat{n}}{\partial g} \{u\left(g\right) - u(c_1)\} + \int_{\widehat{n}}^{1/c_1} u'\left(g\right) dn - \frac{\partial \widehat{n}}{\partial g} \{u\left(g\right) - u(c_1)\}$$

$$\Rightarrow \frac{\partial \Phi}{\partial g} = \int_{\widehat{n}}^{1/c_1} u'\left(g\right) dn > 0 \Rightarrow \frac{\partial \theta_g^*}{\partial g} = -\frac{\frac{\partial \Phi}{\partial g}}{\frac{\partial \Phi}{\partial \theta_g^*}} < 0 \quad \blacksquare$$

2.9.4 Proof of proposition 21

Proof.

I want to compare the equilibrium condition for c_1 under a narrow mandate against the one obtained by Goldstein and Pauzner (2000) in the model without insurance:

$$\pi \int_{\theta^*(c_1)}^{1} \left\{u'(c_1) - R(\theta)u'\left(\frac{1-\pi c_1}{1-\pi}R(\theta)\right)\right\} d\theta = \frac{\partial\theta^*(c_1)}{\partial c_1}\left[\pi u(c_1) + (1-\pi)u\left(\frac{1-\pi c_1}{1-\pi}R(\theta^*(c_1))\right) - u(1)\right] \tag{2.5}$$

Take $\frac{\partial\theta_g^*(c_1)}{\partial c_1} = \frac{\partial\theta^*(c_1)}{\partial c_1} = 0$. I will prove that $c_1^g < c_1^{ng}$ (where the c_1^{ng} stands for the equilibrium without insurance, and c_1^g for the one with insurance).

Define $\varpi(\theta, c_1) \doteq u'(c_1) - R(\theta)u'\left(\frac{1-\pi c_1}{1-\pi}R(\theta)\right)$. The following result is required.

Result 1. $\varpi(\theta, c_1)$ is increasing on θ and decreasing on c_1.

Proof. In order to see that $\varpi(\theta, .)$ is increasing on θ, notice that $\varpi(\theta, .)$ is differentiable and

$$\begin{aligned}
\frac{\partial\varpi}{\partial\theta} &= -R'(\theta)u'\left(c_2(\theta,\pi)\right) - R(\theta)u''\left(c_2(\theta,\pi)\right)\left(\frac{1-\pi c_1}{1-\pi}\right)R'(\theta) \\
&= -R'(\theta)\{u'\left(c_2(\theta,\pi)\right) + u''\left(c_2(\theta,\pi)\right)c_2(\theta,\pi)\} \\
&= -\underbrace{\frac{R'(\theta)}{u'(c_2)}}\underbrace{\left\{1 + \underbrace{\frac{c_2u''(c_2)}{u'(c_2)}}_{<-1}\right\}}_{<0} > 0,
\end{aligned}$$

because by assumption $u(.)$ index of relative risk aversion is higher than 1.

In the same way, $\varpi(., c_1)$ is differentiable and

$$\frac{\partial\varpi}{\partial c_1} = u''(c_1) + \frac{\pi}{1-\pi}\left[R(\theta)\right]^2 u''\left(c_2(\theta,\pi)\right) < 0,$$

because $u(.)$ is concave.■

c_1^{ng} is the solution to

$$\pi \int_{\theta^*(c_1^{ng})}^{1} \left\{u'(c_1^{ng}) - R(\theta)u'\left(\frac{1-\pi c_1^{ng}}{1-\pi}R(\theta)\right)\right\} d\theta = 0,$$

the LHS of equation 2.5 (FOC without insurance).

Evaluating the LHS of equation 2.4 (FOC with insurance) in c_1^{ng}:

$$LHS_{(2.4)}(c_1^{ng}) = \underbrace{\pi \int_{\theta^*(c_1^{ng})}^{1} \left\{ u'(c_1^{ng}) - R(\theta) u' \left(\frac{1-\pi c_1^{ng}}{1-\pi} R(\theta) \right) \right\} d\theta}_{=0}$$

$$+ \int_{\theta_g^*(c_1^{ng})}^{\theta^*(c_1^{ng})} \left\{ u'(c_1^{ng}) - R(\theta) u' \left(\frac{1-\pi c_1^{ng}}{1-\pi} R(\theta) \right) \right\} d\theta$$

$$+\pi \left\{ u'(c_1^{ng}) - R(\theta_g^*(c_1^{ng})) u' \left(\frac{1-\pi c_1^{ng}}{1-\pi} R(\theta_g^*(c_1^{ng})) \right) \right\}$$

Result 2. For c_1 given, $\theta_g^*(c_1) < \theta^*(c_1)$.

Proof. Proposition 8 can be generalised for different values of the insurance offered in each period. It is easy to prove (by implicit differentiation) that θ_g^* is decreasing on the gap $g_1 - g_2$. As the case without insurance can be described as a particular case where $g_1 = 1/n$ and $g_2 = 0$, $\theta_g^*(c_1) < \theta^*(c_1)$ $\forall g_1 = g_2$. ■

Result 3. $\varpi(\theta, c_1)$ increasing on θ implies $\varpi(\theta_g^*(c_1^{ng}), c_1^{ng}) < \varpi(\theta^*(c_1^{ng}), c_1^{ng}) < 0$.

Proof. $\pi \int_{\theta^*(c_1^{ng})}^{1} \varpi(\theta, c_1^{ng}) d\theta = 0$ and $\frac{\partial \varpi}{\partial \theta} > 0$, implies that the function $\varpi(., c_1^{ng})$ is not constant and must change sign in the interval $[\theta^*(c_1^{ng}), 1]$. Being increasing, this means it has to be positive for high values of θ, and negative for small values of θ. In particular, $\varpi(\theta^*(c_1^{ng}), c_1^{ng}) < 0$. ■

From result 3, it follows that $LHS_{(2.4)}(c_1^{ng}) < 0$. Hence, as $\varpi(\theta, c_1)$ is decreasing on c_1, and $\varpi(\theta, c_1^{ng})$ is negative over the region $[\theta_g^*(c_1^{ng}), \theta^*(c_1^{ng})]$, $LHS_{(2.4)}(c_1^g) = 0$ if and only if $c_1^g < c_1^{ng}$.

Remember I have assumed $\frac{\partial \theta_g^*(c_1)}{\partial c_1} = \frac{\partial \theta^*(c_1)}{\partial c_1} = 0$, and then a change in c_1 does not change the limits of integration in the equations. ■

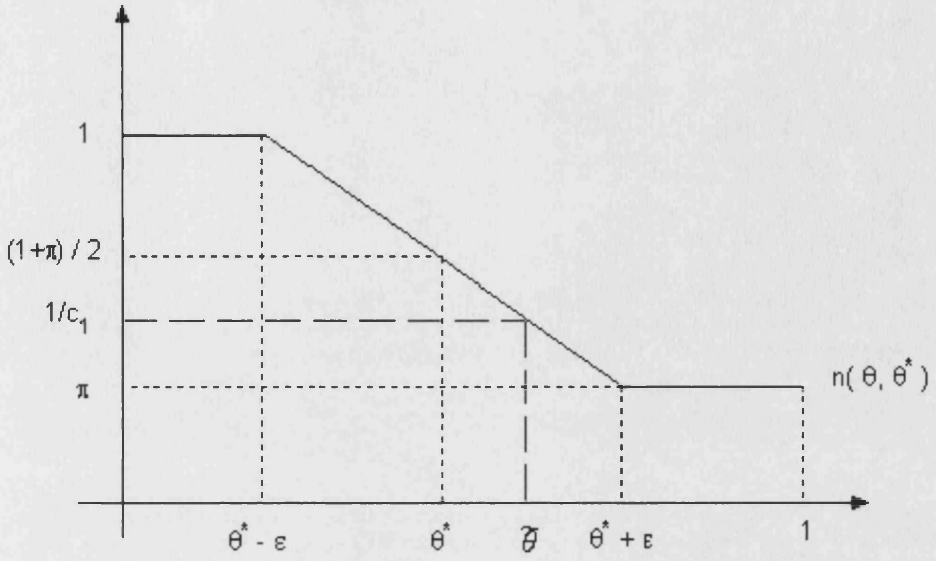

Figure 2-1: Number of total early withdrawals.

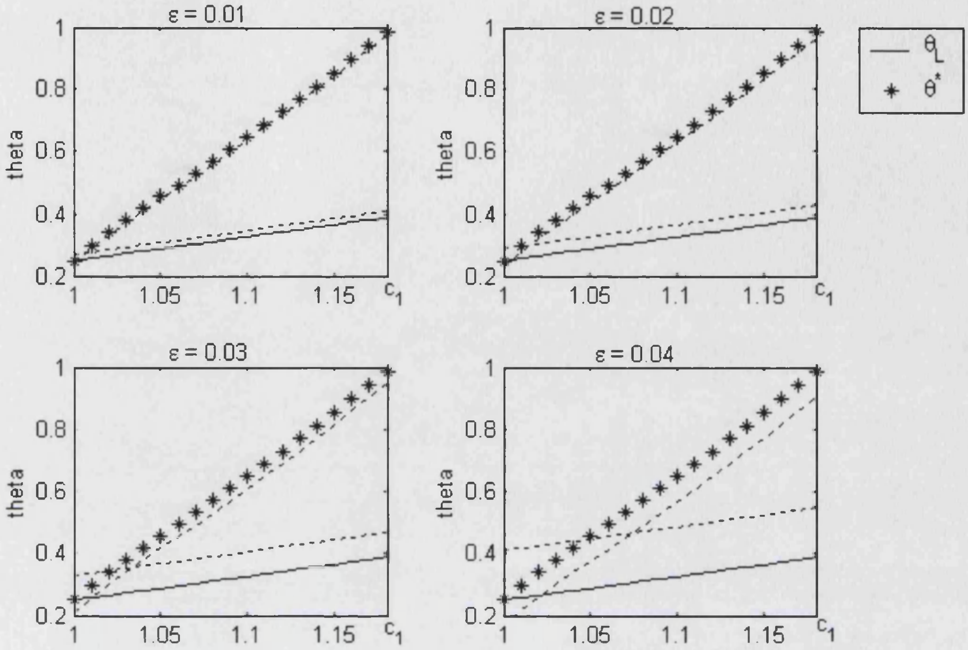

Figure 2-2: Equilibrium and solvency thresholds without insurance.

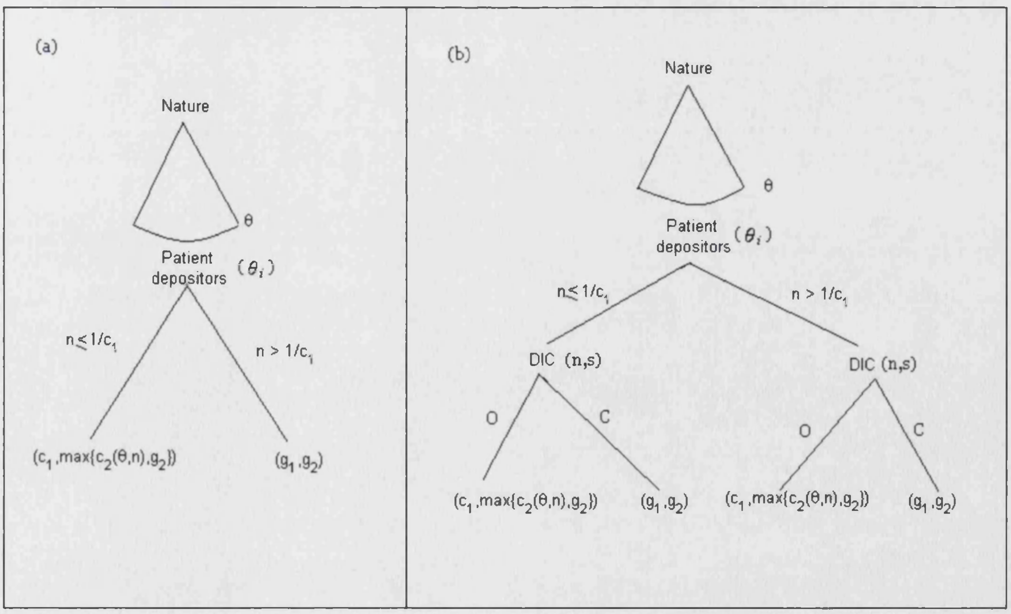

Figure 2-3: Extended form of the game in the interim period, when the DIC operates under (a) a narrow mandate and (b) a broad mandate.

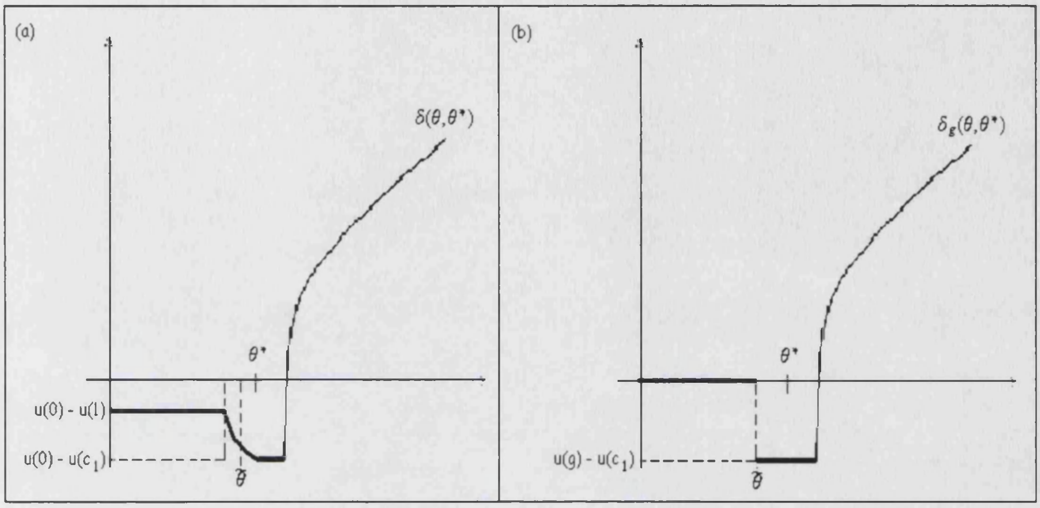

Figure 2-4: Premium to wait to withdraw in the final period (a) without deposit insurance and (b) with deposit insurance.

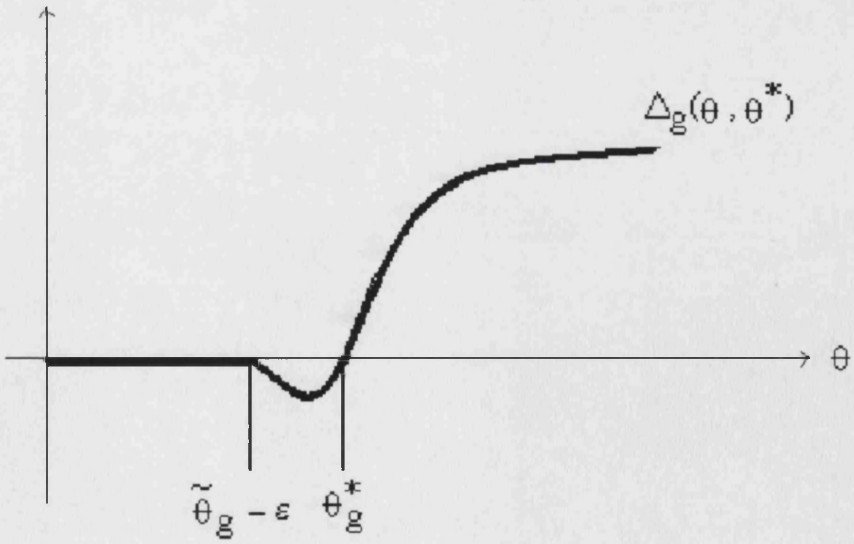

Figure 2-5: Expected premium to wait to withdraw in the final period with deposit insurance.

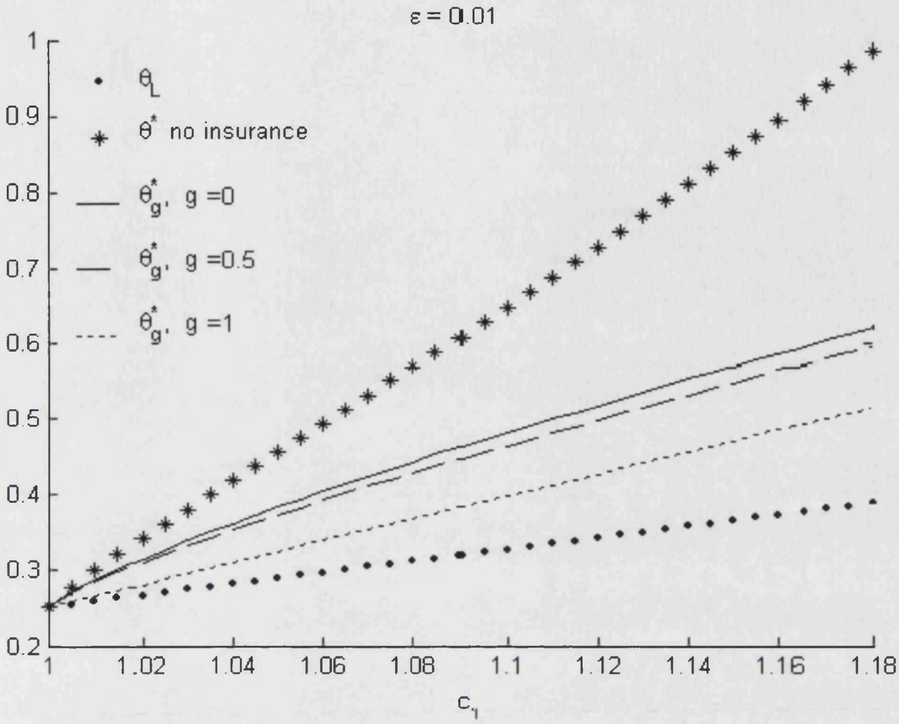

Figure 2-6: Equilibrium and solvency thresholds with and without insurance.

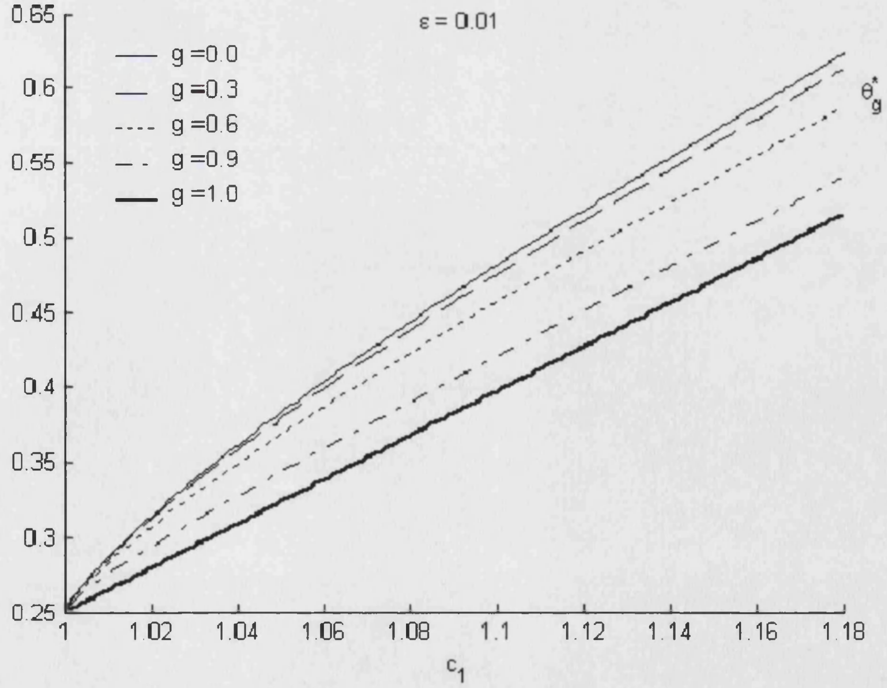

Figure 2-7: Equilibrium threshold with insurance, as a function of c_1 and g.

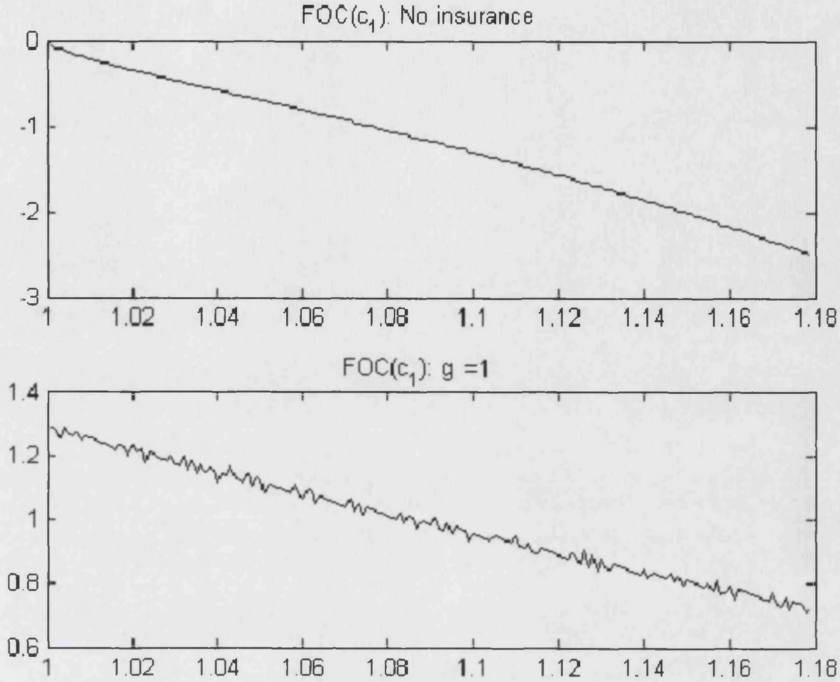

Figure 2-8: Optimal value of c_1 without insurance and with $g = 1$.

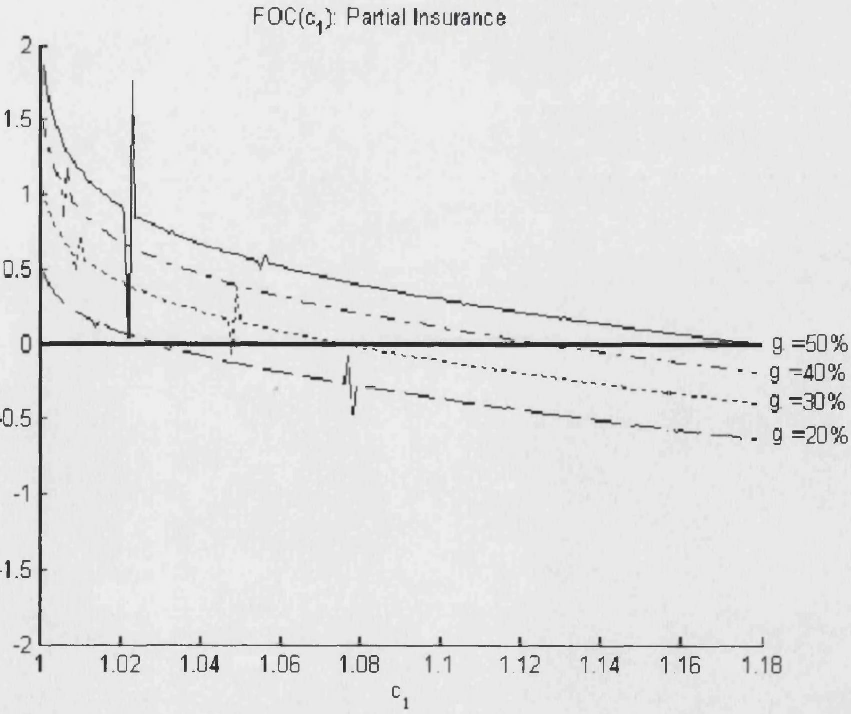

Figure 2-9: Optimal value of c_1 with partial insurance.

Chapter 3

Market Discipline and the Safety Net

3.1 Introduction

Basel II is based on three pillars, one of which is market discipline. Market discipline is broadly interpreted on the basis that private sector agents (stockholders, security holders) face costs that are positively correlated with risk and react in accordance with these costs when pricing banks' instruments (Levy-Yeyati et al, 2004).

The literature identifies two types of market discipline. *Indirect market discipline* refers to the ability of investors and other market participants to evaluate the risk of the banks through pricing information of traded securities, both in the primary and secondary markets. *Direct market discipline* refers to the responsiveness of bank managers to that evaluation because increasing risk leads to increasing costs of funding, thus constraining a bank's risk-taking (Bliss and Flannery, 2000).

One of the purposes of including market discipline in Basel II was to complement the supervisory review process (Basel Committee on Banking Supervision, 2001). There is more than one instrument, though, that could be used to implement market discipline. For example, Basel II proposes a set of disclosure recommendations in three broad areas, capital, risk exposure and capital adequacy, which are intended to reveal the financial strength of banks. On the other hand, several academic authors and central bankers have suggested the introduction of a subordinated debt requirement, on a regular basis, in order to have a widespread, comparable instrument to evaluate the risk level of different institutions.[1]

[1] See Kwast et al. (1999) for a summary of subdebt proposals and the Gramm-Leach-Bliley Financial Modernization Act of 1999 for a policymakers' view on the topic.

Subordinated debt requirements require banks to issue bonds that are senior to equity but junior to all insured claims and thus uncovered by any kind of explicit or implicit guarantee.[2] Behind this proposal is the idea that subordinated debt could allow regulators to require banks to have a bigger share of private funds at risk without harming their competitiveness, increasing the cushion for absorbing losses and saving tax-payer' money in the event of failure. Because of limited liability, equity holders tend to demand more risk than is socially efficient (see Chapter 4). Being junior claimants, subordinated debt holders should have the appropriate incentives to monitor the risk taken by the bank in order to reduce the probability of default, bringing it closer to socially efficient levels.[3]

If correctly priced, subordinated debt yields, with longer maturity and a junior status to other liabilities, should be informative about the risk preferences of banks because rational uninsured investors would demand higher risk premiums from riskier banks. Indeed, many empirical papers have found a positive and significant relationship between the interest rates paid on bank's liabilities and various forms of measuring risk (see Flannery and Nikolova (2003) for evidence in the U.S., Sironi (2003) for evidence in Europe, Demirguç-Kant and Huizinga (2004) for evidence in a sample of developing countries and Evanoff and Wall (2000) for a survey). Hence, the price of debt should provide, both to the market and to supervisors, a low cost signal of the bank riskiness (Berger, Davies and Flannery, 1999).

Without an explicit requirement for issuing subordinated debt, there may be a selection bias problem, where banks with riskier assets may simply refrain from issuing bonds. Moreover, as the price of this type of securities may reflect liquidity risk, as well as default risk, Covitz et al. (2004) argue that a policy of mandating the regular issuance of subdebt would reduce the endogeneity of liquidity premiums, improving the information content of primary and secondary markets' debt spreads.

In principle, the more frequently the bank approaches the market, the stronger will be the discipline. This comes not only through the continuous pricing of debt on a market basis but because the failure to rollover could also be a signal of trouble. However, while this mechanism may operate correctly in normal times, it may generate new sources of instability in situations of financial distress. The accumulation of short-term debt is

[2]Subordinated debt was already present in Basel I, not as a requirement but as an inferior form of equity, limited to 50% of tier 1 capital and counted as an element of tier 2.

[3]This line of argumentation has been criticized by Bliss (2001), because assuming no separation between equity holders and bank managers neglects a basic aspect of corporate governance, agency problems, downplaying the role of equity holders in enhancing market discipline. He argues that bond yield spreads provide a poor noisy predictor of risk and therefore that information should be complemented with equity prices as well as any accounting information that is publicly available.

a concern for regulators because it puts at risk the stability of the financial system. In times of financial distress prices usually depart from the underlying fundamentals and the ratings of private agencies may be delayed or may neglect relevant information, even in the absence of conflicts of interest. Coordination of private lenders might fail to provide liquidity even to solvent banks, ultimately triggering contagion (Rochet and Tirole, 1996).

Some theoretical research has borrowed tools from the corporate finance literature in order to combine market discipline with the other pillars of Basel II. In these models, capital requirements are seen as intervention thresholds for regulators, so they focus on how market discipline modifies these thresholds (Decamps et al, 2004; Bhattacharya et al, 2002). Other authors have studied how market discipline affects the agency problem between equity holders and managers, through changes in the level of monitoring effort put in by the latter, which indirectly induces a given level of risk (Hortala-Vallve, 2002). This chapter will use a variation on the three-period, imperfect information model introduced in Chapter 2, where uninformed insured depositors, informed uninsured sophisticated investors (equity holders and bond holders) and the regulator interact in the market, in order to study and compare the effects that disclosure requirements alone, and in combination with a subordinated debt requirement, have on the equilibrium probability of insolvency of the bank and on the regulator's closure policy. I also study how the maturity of the debt modifies these results.

With a disclosure requirement, I prove that the regulator uses a fully financed closure rule that avoids inefficient survival but involves an inefficient liquidation of assets in a region of the bank's fundamentals. Although the probability of insolvency is higher than the first best, it converges on it as the manager's signal noise goes to zero. When, on top of the disclosure requirement, the bank is asked to issue a zero coupon bond with long maturity, the probability of insolvency is further reduced. The regulator's closure policy in this case forces fewer banks into liquidation, given the subsidy from bond holders to deposit holders when the bank becomes fundamentally unprofitable. Finally, when debt can be rolled over in the interim period and so bond holders do monitor the bank, this period sub-game equilibrium provides very useful information that, in some cases, can completely eliminate inefficient liquidation of assets because the regulator can perfectly observe the bank's fundamentals. However, this will come at the cost of a higher probability of insolvency than in the non-rollover case.

The results in this chapter allow us to conclude that a subordinated debt requirement, when issued at a long maturity, is able to reduce the probability of insolvency for any size of noise. Indeed, intuition suggests that when the manager's signal noise is expected to be high, an appropriately high subordinated debt requirement could restore the first best. On the other hand, a subordinated bond issued with a short maturity can substantially improve on the quality of information (by reducing the noise). Therefore, a subordinated debt requirement can be used to complement disclosure requirements, providing a new set of information which is useful to the regulator.

The chapter is organised as follows. Section 3.2 sets up the model and provides a solution for the first best. Section 3.3 discusses the regulator's optimal closure policy and the equilibrium interest rate when the manager is required to disclose her information. Section 3.4 develops the new equilibrium when, in addition to disclosure, the bank issues a subordinated bond with different maturities. A discussion of the main results and possible extensions is presented in section 3.5 and the conclusions are given in section 3.6.

3.2 The Model

Consider a competitive economy that lasts for three periods $t \in \{0, 1, 2\}$, where all banks have access to the same two investment technologies at the planning period ($t = 0$). For simplicity, I will focus on a representative bank which is liquidated at $t = 2$. The bank manager must decide whether to invest in a liquid technology (storage), returning 1 unit of consumption at $t + 1$ per unit invested at t; or in a stochastic, long-term, partially illiquid technology, returning 1 if liquidated in the interim period ($t = 1$) and R if liquidated at $t = 2$. The long-term return function, $R = R(\theta)$, is continuous and increasing on θ, a random variable uniformly distributed in the interval $[0, 1]$, which represents the underlying fundamentals of the projects financed by the bank. θ is not verifiable until $t = 2$.

A continuum of mass 1 consumers receive 1 unit of endowment at $t = 0$. There exist three types of consumers: depositors (D_0), investors (B_0) and equity holders (W_0); where $0 < W_0 < B_0 < D_0$ and $D_0 + B_0 + W_0 = 1$. Equity holders are risk neutral, consume at $t = 2$ and deposit their money in the bank in exchange for future profits. Investors are also risk neutral and consume at $t = 2$ but can decide whether to invest their money in deposits or in bonds issued by the bank. Depositors are uncertain about their time of consumption. With probability $1 - \pi$ a depositor is patient, meaning that she enjoys consumption only at $t = 2$. With complementary probability, π, she is impatient and consumes only in the interim period. At $t = 1$, types are privately realised and all impatient depositors withdraw to consume. Depositors are risk averse, with preferences represented by a concave and increasing utility function $u(c_t)$, where c_t is the effective payoff in period t.

The bank offers a constant interest rate r on deposits in each period. Agents holding deposits until $t = 2$ receive $(1+r)^2$ if the bank is profitable and a residual payoff $c_2(\theta, r)$ if the bank fails. Hence, they evaluate their expected payoff in the final period, conditional on their beliefs about the response of their counterparts, and decide whether to withdraw or to wait.

I will assume that all agents have access to the storage technology but only the bank can invest in the risky project. I will also assume that the bank manager is one of the risk neutral equity holders, so she acts in the perfect interest of this group (no agency problems). After raising funds (equity, deposits and bonds), the bank manager decides which project to invest in. Assuming that $E_\theta[R(\theta)] > 1$, the risky project is superior to storage and, therefore, all resources are invested on it. Therefore, the bank's only decision is the interest rate offered on deposits.

3.2.1 Benchmark Model: Perfect Information

Assume first that the bank does not issue debt, so that investors can only make deposits. In this case, there are $D_0 + B_0$ units of deposits in the bank and the residual payoff function when the bank is unprofitable is given by $c_2(\theta, r) = \frac{1-\pi D_0(1+r)}{(1-\pi)D_0+B_0} R(\theta)$.

Notice that because $D_0 + B_0 < 1$, if the bank fails at $t = 1$ each deposit holder could receive at most $\frac{1}{D_0+B_0} > 1$, at the expense of the loss of equity holders. Assume $\pi D_0 (1+r) < 1 \left(\Leftrightarrow r < \frac{1-\pi D_0}{\pi D_0}\right)$, otherwise runs would be triggered by the demand of impatient depositors alone. If $(D_0 + B_0)(1+r) < 1$, because of the assumptions about the long-term technology, the bank will always be liquid in the interim period, even if all deposit holders run on it. This would be more likely to happen when the bank is well capitalised (W_0 high). In such a case, when the bank is fundamentally insolvent (see below for a definition), it would be a Pareto superior outcome for patient deposit holders to run because their payoffs would be higher. Indeed, equity holders would even be able to obtain some small pay back, giving private incentives for the manager to disclose her information when the bank's fundamentals are low, without the need for the intervention of a regulator. Therefore, I will work under the assumption that $(D_0 + B_0)(1+r) \geq 1 \left(\Leftrightarrow r \geq \frac{1-D_0-B_0}{D_0+B_0}\right)$, which is likely to hold when the bank is highly leveraged ($D_0 + B_0 \gg W_0$), so that a run does indeed cause the bank to fail.[4]

[4]Because depositors are risk averse, they will require a strictly positive compensation for deposits ($r > 0$). Hence, the inequality is trivially satisfied for $W_0 = 0$ (or $D_0 + B_0 = 1$). By continuity, this should also be true in a neighbourhood of $W_0 = 0$. $L = \frac{D_0+B_0}{W_0}$ is the leverage of the bank, and as

Let $\theta_i(r)$ be the values of the fundamental satisfying $c_2(\theta_i, r) = (1+r)^i$, for $i = 1, 2$.

Definition 23 *A bank is said to be fundamentally solvent if $\theta \geq \theta_1(r)$.*

The spirit of this definition is the same as in the previous chapter. This is, if patient deposit holders are able to monitor the bank, knowing $\theta \geq \theta_1(r)$, they should optimally decide to wait until $t = 2$ because $c_2(\theta, r) \geq (1+r)$.

Notice that the distributional assumption on θ implies that $prob[\theta < \theta_x] = \theta_x$. Hence, ex-ante, a bank offering a demand deposit contract with interest rate r is solvent with probability $1 - \theta_1(r)$.

Implicit differentiation of θ_1 with respect to r gives

$$\frac{d\theta_1}{dr} = \frac{(1-\pi)D_0 + B_0}{(1-\pi D_0(1+r))^2 R'(\theta_1)} > 0.$$

Proposition 24 *The higher the interest rate paid on deposits, the higher the probability of the bank being insolvent.*

Definition 25 *A bank is said to be fundamentally profitable if $\theta \geq \theta_2(r)$.*

Whenever $\theta \geq \theta_2(r)$, the bank makes profits that are shared among equity holders, which happens with probability $1 - \theta_2(r)$, deceasing in r. This comes from the implicit differentiation of θ_2 with respect to r:

$$\frac{d\theta_2}{dr} = \frac{R(\theta_1)(2-\pi D_0(1+r))}{(1-\pi D_0(1+r))R'(\theta_2)} > 0.$$

Assume for a moment that θ is perfectly observable for all agents at $t = 1$. If $\theta < \theta_1(r)$, each patient depositor would anticipate a payoff of $1+r$ if she runs but everybody else waits and $c_2(\theta, r) < 1+r$ if she waits and everybody else also waits. So she runs. Now, the optimal response when all the other patient depositors decide to run is also to run because that action gives a payoff of $\frac{1}{D_0+B_0}$, while waiting gives 0. Therefore, she optimally decides to run (table 3.1).

Conversely, if $\theta \geq \theta_1(r)$, each patient deposit holder would know the bank to be solvent and would anticipate a payoff of $1+r$ if she runs but everybody else waits and of $c_2(\theta, r) \geq 1+r$ if she waits and everybody else also waits, so she would prefer to

$D_0 + B_0 + W_0 = 1$, $D_0 + B_0 \gg W_0$ is equivalent to very high values of L. Another way of achieving this outcome would be for the liquidation of the long term technology to be costly. If the liquidation value of assets in the interim period were low enough, it would be possible for the run of impatient depositors to cause the bank to fail. While the qualitative results should be the same, the algebra would be more complicated with this assumption.

wait. However, the optimal response when all other patient deposit holders decide to run is also to run because that action gives a payoff of $\frac{1}{D_0+B_0}$ while waiting gives 0. In this case two equilibriums exist, one in which all patient deposit holders wait and the bank survives to the final period; and one where all withdraw and the bank goes bankrupt in the interim period, despite the fact that the bank is solvent. This is known as "*coordination failure*", a classical result in games with common knowledge.

	wait	*run*
$t=1$	$1+r$	$\frac{1}{D_0+B_0}$
$t=2$	$c_2(\theta,r)$	0

TABLE 3.1: Patient deposit holders' payoffs in the game with common knowledge and no insurance.

Deposit holders' participation constraint: Ex-ante, in the absence of coordination failure, depositors optimally decide to invest in the bank if $E\left[u(r)\right] \doteq E\left[u(c_1,c_2)\right] \geq \overline{u}$, where $\overline{u}$ is the reservation utility coming from the industry zero profit constraint and c_t is the effective payoff in period t, which depends on r and the strategy played by other deposit holders. Similarly, the risk neutral investors decide to take deposits in the bank if $E\left[v(r)\right] = E\left[v(c_1,c_2)\right] \geq \overline{v}$, where $v(.)$ is the identity function.

Equity holders' payoff: In the absence of coordination failure, if the bank were profitable ($\theta \geq \theta_2\left(r\right)$), the final period profit would be given by

$$W_2(\theta,r) = \left(1-\pi D_0\left(1+r\right)\right)R(\theta) - \left(\left(1-\pi\right)D_0 + B_0\right)\left(1+r\right)^2.$$

Otherwise, because of limited liability: $W_2(\theta,r) = 0$.

Hence, the manager maximises

$$\begin{aligned} E\left[W_2(\theta,r)\right] &= \int_{\theta_2(r)}^{1} \left\{\left(1-\pi D_0\left(1+r\right)\right)R(\theta) - \left(\left(1-\pi\right)D_0+B_0\right)\left(1+r\right)^2\right\} d\theta \\ &= \left(1-\pi D_0\left(1+r\right)\right) \int_{\theta_2(r)}^{1} \left\{R(\theta)-R(\theta_2)\right\} d\theta \geq 0. \end{aligned}$$

Notice that the objective function is decreasing on r. In fact, differentiating with respect to r :

$$\begin{aligned} &\tfrac{d}{dr}E\left[W_2(\theta,r)\right] = -\pi D_0 \int_{\theta_2(r)}^{1} \left\{R(\theta)-R(\theta_2)\right\} d\theta \\ &\quad -\left(1-\pi D_0\left(1+r\right)\right)\left\{\int_{\theta_2(r)}^{1} R'(\theta_2)\tfrac{\partial\theta_2(r)}{\partial r}d\theta + \left[R(\theta_2)-R(\theta_2)\right]\tfrac{\partial\theta_2(r)}{\partial r}\right\} \\ &\quad = -\pi D_0 \int_{\theta_2(r)}^{1} R(\theta)d\theta - 2\left(1-\theta_2(r)\right)\left[\left(1-\pi\right)D_0+B_0\right]\left(1+r\right) < 0. \end{aligned}$$

Benchmark Case Interest Rate: First Best

In order to compute the first best, assume that there are no inefficient runs, that is, patient deposit holders' only withdraw their money from insolvent banks. Therefore, the bank manager solves the following problem:

(P_0) min r

st.

(3.1) $E[u(r)] \geq \overline{u}$

(3.2) $E[v(r)] \geq \overline{v}$

(3.3) $r \geq \frac{1-D_0-B_0}{D_0+B_0}$

(3.4) $r \leq \frac{1-\pi D_0}{\pi D_0}$

Denote by r^0 the optimal solution to this problem.

3.2.2 Imperfect Information

The usual argument in favour of the safety net is the protection of dispersed uninformed depositors. In fact, in general, θ will be imperfectly observable only to the manager. Let me assume that depositors are unable (because of their lack of expertise, or because it is too costly for them) to monitor the bank, so that the game is one of "common absence of knowledge". Hence, two equilibriums persist: either all run or all wait, irrespective of the value of the fundamentals.

Let me introduce one additional player in this game, a regulator, that offers a deposit insurance contract $g_1 = 1$ if the bank is closed in the interim period and $g_2 = 1 + r$ if the bank goes bankrupt in the last period.[5]

	wait	*run*
$t = 1$	$1 + r$	$\frac{1}{D_0+B_0}$
$t = 2$	$\begin{cases} (1+r)^2, & \text{if } \theta \geq \theta_2 \\ \max\{c_2(\theta, r), 1+r\}, & \text{if } \theta < \theta_2 \end{cases}$	$\max\{c_2(\theta, r), 1+r\}$

TABLE 3.2: Patient depositors' payoffs in the game with common absence of knowledge and deposit insurance.

It is clear that, with these payoffs, waiting is the only dominant strategy. If all patient depositors wait:

[5] g_2 is set equal to $1 + r$ in order to eliminate runs (see Chapter 2) but this chapter does not discuss whether the chosen guarantee is optimal. Moreover, remember that we are assuming $(D_0+B_0)(1+r) \geq 1$. If $(D_0+B_0)(1+r) < 1$, insurance in the second period would need to be $\frac{1}{D_0+B_0}$ in order to prevent inefficient runs. Therefore, the compensation for patient depositors would be higher than the capital invested in the interim period, making the insurance system more expensive. In this case, however, the manager would have market driven incentives to declare her information on the quality of the bank's assets in the interim period.

$$E_{wait}^{patient}\left[u(r,g)\right] = \int_{0}^{\theta_2} u(\max\{c_2\left(\theta,r\right),1+r\})d\theta + \int_{\theta_2}^{1} u((1+r)^2)d\theta$$
$$\geq \theta_2 u(1+r) + (1-\theta_2)\, u((1+r)^2)$$

If all run:

$$E_{run}^{patient}\left[u(r,g)\right] = u(\tfrac{1}{D_0+B_0})$$

$$E_{wait}^{patient}\left[u(r,g)\right] - E_{run}^{patient}\left[u(r,g)\right] \geq (1-\theta_2)\left(u((1+r)^2) - u(\tfrac{1}{D_0+B_0})\right) > 0.$$

Investors are risk neutral, so waiting is clearly an optimal strategy for them when the risk averse patient depositors have chosen to do so. Then, with common absence of knowledge and the guarantee on deposits, it is never optimal for deposit holders to withdraw early, even if the bank is insolvent.

Participation constraints: If the regulator does not intervene in the case of insolvent banks, given that patient depositors will always wait until the final period, ex-ante they will optimally decide to invest in the bank if

$$E\left[u(r,g)\right] = \pi u\left(1+r\right) + \left(1-\pi\right)\theta_1 u\left(1+r\right)$$
$$+\left(1-\pi\right)\int_{\theta_1}^{1} u\left(\min\left\{c_2\left(\theta,r\right),(1+r)^2\right\}\right) d\theta \geq \overline{u}.$$

In this case, investors are also covered by the guarantee on deposits, which requires that $E\left[v(r,g)\right] \geq \overline{v}$. Equity holders' expected payoff is the same as in the benchmark case.

Without monitoring, and if the regulator did not intervene in the case of insolvent banks, the bank manager would free ride on deposit insurance, offering a suboptimal interest rate (see Appendix). This is intuitive, given the probability of runs is zero and deposit holders' payoffs are higher $\forall\theta < \theta_1$.

Therefore, in the game where the bank issues only insured deposits, market discipline can only be exerted by the regulator, which needs to get information about the bank. This can be done through disclosure requirements, available market information, or by the combination of these and other instruments. Assume that the regulator has a broad mandate, which means that she has the obligation to pay deposit holders the guarantee but can also take over the bank's assets if it becomes insolvent and manage them in order to pay investors according to their seniority. This implies paying first the guarantee to deposit holders, then liquidating the assets of the bank to repay the insurance contract (plus any fees due) and finally, provided any funds remain, allocating them to creditors according to their seniority, in the following order: deposit holders, bond holders and equity holders.

3.3 Disclosure

Disclosure requires the manager to reveal her signal, $\widetilde{\theta}$, to the regulator. Assume $\widetilde{\theta} = \theta + \varepsilon$, and $\varepsilon \sim U[-\varepsilon_m, \varepsilon_m]$. Hence, given a signal $\widetilde{\theta}$, she knows the true value of θ to be in the interval $\left[\max\left\{\widetilde{\theta} - \varepsilon_m, 0\right\}, \min\left\{\widetilde{\theta} + \varepsilon_m, 1\right\}\right]$.

- $\widetilde{\theta} < \theta_1 - \varepsilon_m$ implies $\theta < \theta_1$, so the bank is insolvent and should be liquidated at $t = 1$.

- $\widetilde{\theta} > \theta_1 + \varepsilon_m$ implies $\theta > \theta_1$, so the bank is solvent and should be allowed to survive until the final period.

- When $\theta_1 - \varepsilon_m \leq \widetilde{\theta} \leq \theta_1 + \varepsilon_m$ it is not possible to distinguish between a solvent or an insolvent bank, though it is possible to assign probabilities to these events.[6]

3.3.1 Closure Policy

If the bank were closed and liquidated at $t = 1$, the regulator would have to pay $D_0 + B_0$ (one unit of consumption per depositor) with a cost equal to

$$\max\{D_0 + B_0 - 1, 0\} = 0.$$

If the regulator assumed control of the bank at $t = 1$ but closed it at $t = 2$, she would have to pay $((1-\pi) D_0 + B_0)(1+r)$ with a cost equal to

$$\max\{((1-\pi) D_0 + B_0)(1+r) - (1 - \pi D_0 (1+r)) R(\theta), 0\}.$$

Hence, when the regulator observes a signal $\widetilde{\theta}$, she closes the bank in the interim period if

$$C\left(\widetilde{\theta}\right) = \int_{\widetilde{\theta}-\varepsilon_m}^{\widetilde{\theta}+\varepsilon_m} \max\{((1-\pi) D_0 + B_0)(1+r) - (1 - \pi D_0 (1+r)) R(\theta), 0\} \, d\theta > 0.$$

Notice that because $R(\theta)$ is increasing on $\theta, C\left(\widetilde{\theta}\right)$ is decreasing on $\widetilde{\theta}$. In fact, for high values of the signal $\left(\widetilde{\theta} \geq \theta_1 + \varepsilon_m\right)$ $C\left(\widetilde{\theta}\right) = 0$, while for low values $\left(\widetilde{\theta} < \theta_1 - \varepsilon_m\right)$ $C\left(\widetilde{\theta}\right) > 0$.

The IMF and BIS codes of best practices requires the resolution of a failing bank to be decided according to a "*least cost*" criterion (Hoelscher and Quintyn, 2003 and Basel

[6] In the absence of deposit insurance, disclosure of the manager's information should trigger depositors to run on insolvent banks with signals in the region $[0, \theta_1 - \varepsilon_m]$, but multiple equilibriums will persist in the region $[\theta_1 - \varepsilon_m, 1]$.

Committee on Banking Supervision, 2002). Hence, define by $\widetilde{\theta}_C = \min\left\{\widetilde{\theta} : C\left(\widetilde{\theta}\right) = 0\right\}$. $\widetilde{\theta}_C$ determines the closure rule of the bank, in the following way: $\forall\ \widetilde{\theta} < \widetilde{\theta}_C$ the bank is closed and $\forall\ \widetilde{\theta} \geq \widetilde{\theta}_C$ the bank is left open.

Proposition 26 *When the manager discloses her signal to the regulator, it is cost efficient for the latter to close the bank in the interim period when $\widetilde{\theta} < \widetilde{\theta}_C = \theta_1 + \varepsilon_m$ and leave it open otherwise.*

Proof. It only remains to prove that $\widetilde{\theta}_C = \theta_1 + \varepsilon_m$. By definition, for all signals $\widetilde{\theta} \geq \theta_1 + \varepsilon_m$, $C\left(\widetilde{\theta}\right) = 0$.

Consider $\widetilde{\theta} = \theta_1 + \varepsilon_m - \delta$, with $\delta > 0$. [7]

$$C\left(\widetilde{\theta}\right) = \int_{\theta_1-\delta}^{\theta_1+2\varepsilon_m-\delta} \max\left\{\left(\left(1-\pi\right)D_0 + B_0\right)\left(1+r\right) - \left(1-\pi D_0\left(1+r\right)\right)R\left(\theta\right), 0\right\} d\theta$$
$$= \int_{\theta_1-\delta}^{\theta_1} \left\{\left(\left(1-\pi\right)D_0 + B_0\right)\left(1+r\right) - \left(1-\pi D_0\left(1+r\right)\right)R\left(\theta\right)\right\} d\theta > 0,\ \forall\delta > 0.$$

Therefore, $\widetilde{\theta}_C = \theta_1 + \varepsilon_m$. ■

The proposed guarantee is then fully financed, without the need to resort to taxes because, for signals below $\widetilde{\theta}_C$ (i.e., for $\theta < \theta_1$), banks are closed in the interim period and depositors receive $\frac{1}{D_0+B_0}$, which is higher than the promised guarantee in that period ($g_1 = 1$).

Proposition 27 *With the previous closure rule there is no inefficient survival, although inefficient liquidation of assets might persist for a region of the fundamentals.*

Proof. Notice that if $\theta < \theta_1$, the signal would necessarily be smaller than $\widetilde{\theta}_C$, hence all insolvent banks are liquidated. When $\theta > \theta_1 + 2\varepsilon_m$ the signal is above $\widetilde{\theta}_C$ and solvent banks in that region survive. However, in the interval $[\theta_1, \theta_1 + 2\varepsilon_m]$ inefficient liquidation of assets might persist. ■

Interest rate with insurance and disclosure

Given the closure rule, the manager solves

(P_D) min r

st.

(3.9) $E_{\widetilde{\theta}_C}\left[u(r,g)\right] \geq \overline{u}$

(3.10) $E_{\widetilde{\theta}_C}\left[v(r,g)\right] \geq \overline{v}$

(3.11) $r \geq \frac{1-D_0-B_0}{D_0+B_0}$

(3.12) $r \leq \frac{1-\pi D_0}{\pi D_0}$

[7] The proof assumes $\delta \leq 2\varepsilon_m$, but the result follows directly if $\delta > 2\varepsilon_m$.

Proposition 28 *The probability of insolvency is higher in the case with deposit insurance and disclosure than in the first best. Nevertheless, it converges on the benchmark solution as the noise goes to zero.*

Proof. See Appendix. ■

The proof is intuitive. Patient deposit holders' payoffs are the same under the benchmark case as with disclosure and intervention, except in the shaded area of figure 3-1. In this region, their payoff in the benchmark case is higher:

$$c_2(\theta, r) \geq prob(\widetilde{\theta} < \widetilde{\theta_C})\frac{1}{D_0 + B_0} + prob(\widetilde{\theta} \geq \widetilde{\theta_C})c_2(\theta, r).$$

Therefore, in order to meet the participation constraints, the equilibrium interest rate must be higher.

So far, I have implicitly assumed that the manager voluntarily tells the truth to the regulator and reveals her signal but a more likely outcome is that the regulator obtains this information by direct on-site monitoring, which is costly. Being this a sunk cost, however, it will not modify the regulator's policy in any relevant way. For simplicity, throughout the rest of this chapter I will assume monitoring costs to be nil for all agents, although an extension of the results for positive values of these costs will be discussed in section 3.5.

In any case, the regulator might decide not to rely exclusively on the signal declared by the manager but also on other available market information and instruments, such as subordinated debt, which is the topic of the next section.

3.4 Subordinated Debt

Assume now that the bank issues a subordinated bond in the initial period, which implies that bond holders will be junior to depositors but senior to equity holders, in case the bank defaults on its obligations.

I am going to study two cases: one with a zero coupon bond with maturity at $t = 2$ and one where the bond can be rolled over at $t = 1$.

3.4.1 Zero Coupon Bond with Maturity at t=2

Consider the case of a zero coupon bond returning $b_2(r, s_0)$ units of consumption at maturity per unit invested at $t = 0$. That is, the return demanded by investors depends both on the number of subordinated bonds issued and the interest rate paid on deposits.

The bank is required to issue s_0 units of this bond at $t = 0$. For each value of r and θ, the profit of the bank at $t = 2$ is given by

$$\begin{aligned} W_2(\theta, r) &= (1 - \pi D_0 (1+r)) R(\theta) - [(1-\pi) D_0 + B_0 - s_0] (1+r)^2 - s_0 b_2 (r, s_0) \\ &= (1 - \pi D_0 (1+r)) R(\theta) - [(1-\pi) D_0 + B_0] (1+r)^2 - s_0 \left(b_2 (r, s_0) - (1+r)^2 \right), \end{aligned}$$

where $s_0 \leq B_0$. Naturally, $b_2 (r, s_0)$ must be higher than $(1+r)^2$ in order to attract investors away from deposits.

This time, the bank is profitable if $\theta \geq \theta_S$, where $W_2(\theta_S, r) = 0$ or equivalently $c_2 (\theta_S, r) = (1+r)^2 + s_0 \frac{b_2(r,s_0) - (1+r)^2}{(1-\pi)D_0 + B_0}$.

Notice that because $c_2 (\theta, r)$ is increasing on θ, $\theta_S(r) > \theta_2(r)$ and $\theta_S(r)$ is increasing on r.

Define by $\overline{\theta_i}$ the value of the fundamentals satisfying

$$\begin{aligned} (1 - \pi D \Leftrightarrow c_2 \left(\overline{\theta_i}, r\right) &= (1+r)^i - \tfrac{s_0(1+r)^i}{(1-\pi)D_0 + B_0} < (1+r)^i \, .^i = 0, \; i = 1, 2 \\ \Leftrightarrow c_2 \left(\overline{\theta_i}, r\right) &= (1+r)^i - \tfrac{s_0(1+r)^i}{(1-\pi)D_0 + B_0} < (1+r)^i . \end{aligned}$$

Clearly, $\overline{\theta_i} (r) < \theta_i (r)$ for $i = 1, 2$.

Players' payoffs:

When the bank's assets are liquidated in the interim period, given the form of the guarantee, each deposit holder receives $\min \left\{ \frac{1}{D_0 + B_0 - s_0}, 1 + r \right\}$ units of consumption, bondholders receive $\max \left\{ \frac{1 - (1+r)[D_0 + B_0 - s_0]}{s_0}, 0 \right\}$ and, because of limited liability, equity holders receive nothing.

When the bank is open in the final period and $\theta \geq \theta_S$, deposit holders receive $(1+r)^2$, bondholders $b_2 (r, s_0)$ and equity holders share profits and obtain $\frac{W_2(\theta, r)}{W_0}$ each.

Otherwise, if $\overline{\theta_2} \leq \theta < \theta_S$: deposit holders receive $(1+r)^2$, bond holders obtain a residual payoff and, because of limited liability, equity holders receive nothing.[8]

When $\theta < \overline{\theta_2}$ the regulator takes over on the bank's assets and players other than deposit holders receive nothing. If $\overline{\theta_1} \leq \theta$ depositors receive $c_2(\theta, r) \geq 1 + r$ and if $\theta < \overline{\theta_1}$ they receive $1 + r$, the guaranteed value of deposits in that period.

Ex-ante, bond holders optimally decide to invest in the bank's bond instead of deposits if $E_{\widetilde{\theta_C^b}} [v(b_t)] \geq E_{\widetilde{\theta_C^b}} [v(r, g, s_0)]$. Given that they cannot liquidate their assets in the interim period (there is no secondary market in this model), the only relevant information for them is the ex-ante information. In particular, they need to anticipate

[8] It is important that bond holders are senior to equity holders, otherwise the latter will get a subsidy in default.

the regulator's closure rule, which is incorporated in the required compensation for the bond, $b_2(r, s_0)$, increasing on r (see the Appendix).[9] Hence, once again, the only information that the regulator can recover in the interim period is that disclosed by the bank manager.

Closure policy with disclosure and a long maturity subordinated bond

If the bank is closed and liquidated at $t = 1$, the regulator has to pay D_0 with a cost equal to $\max\{D_0 + B_0 - s_0 - 1, 0\} = 0$.

If the bank is intervened at $t = 2$, the regulator has to pay $[(1-\pi) D_0 + B_0 - s_0](1+r)$ with a cost equal to

$$\max\{[(1-\pi) D_0 + B_0 - s_0](1+r) - (1 - \pi D_0 (1+r)) R(\theta), 0\}.$$

Hence if the regulator observes a signal $\widetilde{\theta}$, she closes the bank in the interim period if

$$C\left(\widetilde{\theta}\right) = \int_{\widetilde{\theta}-\varepsilon_m}^{\widetilde{\theta}+\varepsilon_m} \max\{[(1-\pi) D_0 + B_0 - s_0](1+r) - (1 - \pi D_0 (1+r)) R(\theta), 0\} d\theta > 0$$

Notice the function under the integral is decreasing on θ. For high values of the signal $\left(\widetilde{\theta} \geq \overline{\theta_1} + \varepsilon_m\right) C\left(\widetilde{\theta}\right) = 0$, while for low values $\left(\widetilde{\theta} < \overline{\theta_1} - \varepsilon_m\right) C\left(\widetilde{\theta}\right) > 0$.

Define by $\widetilde{\theta}_C^b = \min\left\{\widetilde{\theta} : C\left(\widetilde{\theta}\right) = 0\right\}$. $\widetilde{\theta}_C^b$ determines the closure rule of the regulator when the bank issues a long term subordinated bond and the manager discloses her information, in the following way: $\forall\, \widetilde{\theta} < \widetilde{\theta}_C^b$ the bank is closed and $\forall\, \widetilde{\theta} \geq \widetilde{\theta}_C^b$ the bank is left open until $t = 2$.

Proposition 29 *When the bank issues a long-term zero coupon subordinated bond and the manager discloses her signal to the regulator, it is cost efficient for the latter to close the bank in the interim period when* $\widetilde{\theta} < \widetilde{\theta}_C^b = \overline{\theta_1} + \varepsilon_m$ *and leave it open otherwise.*

Proof. As in proposition 26 ∎

Notice that if $\theta < \overline{\theta_1}$, then $\widetilde{\theta} < \widetilde{\theta}_C^b$ and the bank is closed. On the other hand, for all $\theta > \overline{\theta_1} + 2\varepsilon_m$ it is necessary true that $\widetilde{\theta} > \widetilde{\theta}_C^b$ and the bank is allowed to survive until the final period. Once again, this rule promotes no inefficient survival, although inefficient liquidation of assets might persist in a region of the fundamentals $[\overline{\theta_1}, \overline{\theta_1} + 2\varepsilon_m]$.

[9] Being both the market for deposits and the market for subdebt competitive, in equilibrium investors demand a return $b_2(r, s_0)$ so that they break even, i.e., so that the participation constraint is satisfied with equality. This assures that any amount of subdebt $s_0 \leq B_0$ can be issued and will meet the demand of risk neutral investors.

Optimal interest rate

Given the payoffs described above, the bank manager maximises

$$E_\theta\left[W_2(\theta,r)\right] = \left(1-\pi D_0\left(1+r\right)\right) \int_{\theta_S(r)}^{1} \left\{R(\theta)-R(\theta_S\left(r\right))\right\} d\theta,$$

which is again a decreasing function of r.

Hence, she solves:

(P_B) min r

st.

(3.13) $E_{\overline{\theta}_C^b}\left[u(r,g,s_0)\right] \geq \overline{u}$

(3.14) $E_{\overline{\theta}_C^b}\left[v(r,g,s_0)\right] \geq \overline{v}$

(3.15) $E_{\overline{\theta}_C^b}\left[v(b_t)\right] \geq E_{\overline{\theta}_C^b}\left[v(r,g,s_0)\right]$

(3.16) $r \geq \frac{1-D_0-B_0}{D_0+B_0}$

(3.17) $r \leq \frac{1-\pi D_0}{\pi D_0}$

Proposition 30 *Under deposit insurance and disclosure, both the interest rate offered by the bank and the probability of insolvency are lower in the presence of subordinated debt.*

Proof. See Appendix, section 3.7.5. ■

An intuitive proof of this result can be seen in figure 3-2. Patient depositors' payoffs are the same only when $\theta < \overline{\theta_1}$. For all other values of θ, the payoffs with subdebt and disclosure are higher than or equal to those in the case without subdebt, which means that the compensation required to make the depositors' participation constraint binding must be lower.

Defining the requirement

The subordinated debt requirement was introduced in this model as an obligation to issue a minimum amount of a subordinated bond (Calomiris, 1997). Different possibilities for the implementation of this requirement have been discussed in the literature (see Kwast et al, 1999), which can be broadly summarised as:

(i) a minimum issuance of a fraction α of equity: $s_0 \geq \alpha W_0$, $0 \leq \alpha \leq 1$.

(ii) a minimum issuance of a fraction α of deposits or total liabilities: $s_0 \geq \alpha\left(D_0 + B_0\right)$, $0 \leq \alpha \leq 1$.

(iii) a minimum issuance of a fraction α of risk weighted assets, I_0: $s_0 \geq \alpha I_0$, $0 \leq \alpha \leq 1$. In the model as stated, $I_0 = W_0 + D_0 + B_0$, therefore this type of requirement would be a linear combination of cases (i) and (ii).

The demand for bonds is constrained above by B_0, the number of investor type of consumers, all endowed with one unit of consumption. Hence, under the model's assumptions, any value of $0 < \alpha \leq 1$ will generate a risk reduction in case (i) ($\alpha W_0 \leq s_0 \leq B_0$) but only small values will be feasible under schemes (ii) and (iii) (because by assumption $D_0 > B_0$).

Notice that for the long-term subordinated bond to reduce the bank's probability of insolvency, in this model, the choice of s_0 is not crucial provided, of course, that it is strictly positive. This may seem to contradict the results obtained in Niu (2008), where a minimum issuance is required in order to affect the banks' risk-taking. However, his model explicitly allows for the manager to choose between two risky technologies, whereas here I am focusing on an indirect indicator, which is the bank's probability of insolvency. Anyway, I will show later (proposition 32) that the incentives of investors to take the bond instead of deposits may be enhanced when the requirement is above a given threshold.

The introduction of "more sophisticated" agents, unprotected by the guarantee on deposits and senior to equity holders, has modified the behaviour of the bank manager; achieving in equilibrium a lower probability of insolvency, even when bond holders do not monitor the bank in the interim period.

Corollary 31 *In the model with deposit insurance, zero noise disclosure and subordinated debt, both the interest rate chosen by the bank and the probability of insolvency are lower than in the benchmark case.*

Proof. The result in proposition 30 holds for all values of ε_m, in particular for $\varepsilon_m = 0$. ■

Whether this outcome is efficient depends on where the social optimum lies with respect to the first best. If the social optimum implies a lower interest rate than the private optimum (as seen in Chapter 4), then subordinated debt could help in achieving that goal.

Another corollary of proposition 30 is that the closure threshold in this case is smaller: $\widetilde{\theta}_C^b \left(r^B\right) < \widetilde{\theta}_C \left(r^D\right)$, $\forall s_0 > 0$. Therefore, fewer banks are liquidated in the interim period because there is a subsidy from bond holders to deposit holders in a region of (low) fundamentals.

From the proof of proposition 28 in the Appendix it can be seen that for small values of ε_m,

$$\Phi(r^D, \varepsilon_m) \approx 2\varepsilon_m \theta_1 \left[u\left(1 + r^D\right) - u\left(\tfrac{1}{D_0 + B_0}\right)\right].$$

Hence, at least in a small neighbourhood of zero, the higher ε_m the wider the gap between r^D and r^0, i.e., the interest rate increases with the noise. On the other hand, when $s_0 \to 0$, the equilibrium interest rate with subdebt converges on r^D. Therefore, at least in a small neighbourhood of zero, r^B is decreasing on s_0. If these results were sustained for all values of ε_m and s_0, it would follow that the higher the noise, the higher the subdebt requirement necessary to restore the benchmark equilibrium.

Intuitively, the more numerous the participants in the market for subordinated bonds, the higher their final payoff, given that the pool of money subsidising depositors is higher in case the bank's fundamental falls below θ_S. However, because the partial derivative of $\overline{\theta_2}$ with respect to s_0 is negative, this should be consistent with a lower value for this threshold. Indeed, from figure 3-3 it is clear that the lower $\overline{\theta_2}$ with respect to θ_1, the higher the residual payoff of the bond in the default region. The following proposition defines the set of values of s_0 for which this is the case.

Proposition 32 *In the region where the bank is fundamentally unprofitable, bond holders' payoffs are higher if* $s_0 \geq [(1-\pi) D_0 + B_0] \dfrac{r}{1+r}$.

Proof. See Appendix, section 3.7.6. ■

3.4.2 Roll-over at t=1

This time the bond has the following structure: it can be redeemed at $t = 1$ at a value $K = 1$, or held until $t = 2$ in exchange for $b_2(r, s_0)$ units of consumption. If bond holders decide not to redeem their bonds in the interim period, I interpret this as a rollover decision.

Define by θ_K the value of θ such that bond holders' payoff in the final period is equal to $K = 1$. Clearly, $\theta_K < \theta_S$ (figure 3-3). Therefore, and because for all $\theta < \theta_S$ bond holders receive only a residual payoff, θ_K is the solution to

$$(1 - \pi D_0 (1+r)) R(\theta_K) - [(1-\pi) D_0 + B_0 - s_0] (1+r)^2 = s_0 .$$

Bond holders are now sophisticated, informed investors, which through monitoring receive a private signal

$$\theta_i = \theta + \varepsilon_i, \quad \varepsilon_i \sim U[-\varepsilon_b, \varepsilon_b].$$

They use this information to evaluate whether to rollover or to redeem the bond in the interim period. As in Chapter 2, and following Goldstein and Pauzner (2000), I will concentrate here on equilibriums on *"switching strategies"*, that is, an equilibrium in monotone strategies with threshold θ^*, such that if $\theta_i < \theta^*$ bond holders redeem the bond at $t = 1$, and roll it over if $\theta_i > \theta^*$.

After receiving a non verifiable signal θ_i at $t = 1$, bond holder i knows that the true value of θ lies in the interval $[\max\{\theta_i - \varepsilon_b, 0\}, \min\{\theta_i + \varepsilon_b, 1\}]$. With incomplete information, they must condition their beliefs upon their private signals, which are known to be positively correlated with the signals of others. Define by $0 \leq n \leq s_0$ the number of bond holders redeeming their bonds at $t = 1$. Naturally, if $\theta < \overline{\theta_2}(r) - 2\varepsilon_b$, all will receive a signal $\theta_i < \overline{\theta_2}(r) - \varepsilon_b$, where the payoff is zero and therefore each one will redeem the bond irrespective of what the other players do. This is known as the *lower dominance region*, where n equals s_0. Similarly, if $\theta > \theta_S(r) + 2\varepsilon_b$ all debt holders receive signals $\theta_i > \theta_S(r) + \varepsilon_b$, and so will find themselves in the *upper dominance region,* where the payoff is b_2 and everyone rolls over, regardless of the actions of other players.

With all bond holders following the same equilibrium strategy, each one will rationally anticipate that if $\forall i \ \theta_i < \theta^*$, everybody will liquidate the bond at $t = 1$. This will be the case for values of $\theta < \theta^* - \varepsilon_b$. In the same way, everybody will rollover if $\forall i \ \theta_i > \theta^*$, which is always the case if $\theta > \theta^* + \varepsilon_b$. Finally, and because of the uniform distributional assumption on ε_i, in the intermediate region a rational player will assign a uniform distribution to her belief about the number of bond holders liquidating the bond at $t = 1$, $n(\theta, \theta^*)$, that will be given by the following non-increasing function of θ (figure 3-4):

$$n(\theta, \theta^*) = \begin{cases} s_0 & if \quad \theta < \theta^* - \varepsilon_b \\ -\dfrac{(\theta - \theta^* - \varepsilon_b)s_0}{2\varepsilon_b} & if \quad \theta^* - \varepsilon_b \leq \theta \leq \theta^* + \varepsilon_b \\ 0 & if \quad \theta > \theta^* + \varepsilon_b \end{cases}$$

Notice that for $n(\theta, \theta^*)$ to be well defined, it has to be consistent with the beliefs implied by the existence of the dominance regions. Hence, $\overline{\theta_2}(r) - 2\varepsilon_b < \theta^* - \varepsilon_b$ and $\theta^* + \varepsilon_b < \theta_S(r) + 2\varepsilon_b$, which implies $\overline{\theta_2}(r) - \varepsilon_b < \theta^* < \theta_S(r) + \varepsilon_b$.

In deciding whether to redeem the bond at $t = 1$ or to roll it over, bond holders must compare the conditional expected utility of these two actions. Denoting by $\delta(\theta, \theta^*)$ the difference of utilities between waiting until $t = 2$ and redeeming the bond at $t = 1$:

$$\delta(\theta, \theta^*) \doteq \delta(\theta, n(\theta, \theta^*))$$
$$= \max\left\{\min\left\{\frac{[1-n(\theta,\theta^*)-\pi D_0(1+r)]R(\theta)-[(1-\pi)D_0+B_0-s_0](1+r)^2}{s_0-n(\theta,\theta^*)}, b_2\left(r, s_0\right)\right\}, 0\right\} - b_1,$$

where $b_1 \leq 1$ is the effective payoff at $t = 1$ when redeeming the bond in that period.[10]

Upon receiving a signal θ_i each bond holder evaluates:

$$\Delta(\theta_i, \theta^*) \doteq E_{\theta_i}[\delta(\theta, \theta^*)] = \int_{\theta_i-\varepsilon_b}^{\theta_i+\varepsilon_b} \tfrac{1}{2\varepsilon_b}\delta(\theta, \theta^*)d\theta.$$

If a bond holder's conditional expected utility of rolling over is higher than the utility of redeeming the bond in the interim period, that is if $\Delta(\theta_i, \theta^*) > 0$, she will rollover. Otherwise, if $\Delta(\theta_i, \theta^*) < 0$, she will quit at $t = 1$. Finally, if $\Delta(\theta_i, \theta^*) = 0$ she will be indifferent between the two actions, so I assume that she rolls over.

Using the continuity of $\Delta(\theta_i, \theta^*)$ in both arguments[11], the existence of the dominance regions, and the monotonicity[12] of $\Delta(\theta_i, \theta_i)$, it is easy to prove that there exists a unique equilibrium threshold, satisfying $\theta^* \in]\overline{\theta_2}(r) - \varepsilon_b, \theta_S(r) + \varepsilon_b[$, and such that:[13]

- $\Delta(\theta^*, \theta^*) = 0$,
- $\Delta(\theta_i, \theta^*) < 0$ for all $\theta_i < \theta^*$ and
- $\Delta(\theta_i, \theta^*) > 0$ for all $\theta_i > \theta^*$.

Making a strong use of the uniform distributional assumption on the noise, Goldstein and Pauzner (2000) show that for any feasible belief $n(\theta)$, the regions where $\Delta(\theta_i, n(\theta)) \leq 0$ and $\Delta(\theta_i, n(\theta)) > 0$ are complementary connected intervals and, therefore, any equilibrium of the game must be monotone. That is to say, the strategy previously defined is the unique equilibrium of this game.

Finally, from the definition of $\Delta(\theta_i, \theta^*)$ (figure 3-5) it is possible to see that

[10] If the bank survives until $t = 2$, bond holders receive $b_1 = 1$. If the bank is closed by the regulator in the interim period, the effective payoff of debt holders is $b_1 = \max\left\{\frac{1-(1+r)[D_0+B_0-s_0]}{s_0}, 0\right\} < 1$.

[11] $\delta(\theta, \theta^*)$ is discontinuous only at one point (where b_1 jumps to 1) but this discontinuity disappears when integrating over θ (figure 3-5).

[12] Notice $R(.)$ and $-n(., \theta_i)$ are monotonically increasing in $\theta \in [\theta_i - \varepsilon_b, \theta_i + \varepsilon_b]$ for $\theta^* = \theta_i$. Although the property may fail for values of θ_i in the neighbourhood of the point where b_1 jumps (well below $\overline{\theta_2}(r) - 2\varepsilon_b$ if ε_b is small enough), at that point $\Delta(\theta_i, \theta_i)$ is non-positive and monotonically increasing afterwards (see figure 3-5).

[13] The proof is standard and can be found in the Appendix of Chapter 2. In case $b_1 = \max\left\{\frac{1-(1+r)[D_0+B_0-s_0]}{s_0}, 0\right\} = 0$, the second condition could be generalised to $\Delta(\theta_i, \theta^*) \leq 0$ for all $\theta_i < \theta^*$. The equilibrium will still be unique, because other values of θ_i satisfying $\Delta(\theta_i, \theta_i) = 0$ will be outside the interval $]\overline{\theta_2}(r) - \varepsilon_b, \theta_2(r) + \varepsilon_b[$.

$$\delta(\theta,\theta^*) < 0\ \forall\theta < \theta_K \Rightarrow \int_{\theta_K-2\varepsilon_b}^{\theta_K} \delta(\theta,\theta^*)\,d\theta \doteq \Delta(\theta_K-\varepsilon_b,\theta^*) < 0,$$
$$\delta(\theta,\theta^*) > 0\ \forall\theta > \theta_K \Rightarrow \int_{\theta_K}^{\theta_K+2\varepsilon_b} \delta(\theta,\theta^*)\,d\theta \doteq \Delta(\theta_K+\varepsilon_b,\theta^*) > 0.$$

Proposition 33 $\theta^* \in]\theta_K-\varepsilon_b,\theta_K+\varepsilon_b[$. *Therefore, the bond holders' equilibrium in the game with rollover is close to the point where they break even.*

Closure policy

The regulator does not observe the bond holders' signals but, from the managers' disclosure and direct monitoring, can observe $\widetilde{\theta}$ and n, and can also compute the value of θ^*. Define by $\overline{\theta_i}(n)$ the value of θ satisfying

$$(1-n-\pi D_0(1+r))\,R\left(\overline{\theta_i}(n)\right) = [(1-\pi)D_0+B_0-s_0](1+r)^i.$$

Notice $\overline{\theta_i}(n) \underset{n\to 0}{\rightarrow} \overline{\theta_i}$ and $\overline{\theta_i} < \overline{\theta_i}(n)$ for $i=1,2$.

Given all the available information, the regulator decides to close the bank in the interim period if the expected cost of leaving it open is strictly positive, which gives the following rule:

- If $0 < n < s_0$, using the inverse function of $n(\theta,\theta^*)$ in the region $[\theta^*-\varepsilon_b,\theta^*+\varepsilon_b]$, the regulator can infer the true value of $\theta(n) = \theta^* + \frac{\varepsilon_b}{s_0}(s_0-2n)$ and can evaluate, without any noise, whether to leave the bank open (if $\theta(n) \geq \overline{\theta_1}(n)$) or close it in the interim period (if $\theta(n) < \overline{\theta_1}(n)$).

- If $n=0$: $\overline{\theta_2}(0) = \overline{\theta_2}$ and $\theta > \theta^*+\varepsilon_b > \overline{\theta_2}$. Therefore, it is cost-efficient for the regulator and welfare improving for consumers to leave the bank open.

- If $n=s_0$: $\theta < \theta^*-\varepsilon_b < \theta_K$ but it is not clear if θ is higher or lower than $\overline{\theta_1}(s_0)$. Therefore, the regulator will need to resort to the bank manager's signal and the closure policy will be similar to the one described in section 3.3.

Proposition 34 *The closure rule when the bank issues a subordinated bond which can be rolled over in the interim period, is such that the regulator closes the bank at* $t=1$ *if* $\widetilde{\theta} < \widetilde{\theta}_C^R = \overline{\theta_1}(n)+\varepsilon(n)$ *and leaves it open otherwise; where*
$$\varepsilon(n) = \begin{cases} 0 & 0 \leq n < s_0 \\ \varepsilon_m & n = s_0 \end{cases}.$$

Whenever $n < s_0$, the information coming from bond holders' behaviour eliminates the noise on the bank's signal. In particular, inefficient liquidation of assets is completely ruled out because the regulator can perfectly observe the value of the bank's fundamentals.[14] When $n = s_0$, inefficient liquidation of assets persists if $\varepsilon_m > 0$.

A classical result for global games applies here, that public disclosure would not improve on the equilibrium of the game. If, on the contrary, either the manager or any bond holder publicly declared her signal to the market, that information would become common knowledge, breaking bond holders' coordination and prompting multiple equilibrium in the interim period sub-game, outside of the dominance regions.

A final remark regarding the interest rate equilibrium in this case: given $\overline{\theta_i}(r) \leq \overline{\theta_i}(r,n)$ for all r, n and $i \in \{1,2\}$; the expected subsidy from bond holders to deposit holders would be lower, so the equilibrium interest rate offered by the manager, r^R, will be higher than r^B, the optimal solution under disclosure and a long-term subordinated bond.

Also, if the fraction of impatient depositors is low (so that $r\pi D_0 < W_0$), $\overline{\theta_i}(r,n) \leq \theta_i(r)$ and then $\overline{\theta_1}\left(r^B\right) \leq \overline{\theta_1}\left(r^R\right) \leq \overline{\theta_1}\left(r^R,n\right) \leq \theta_1\left(r^R\right) \leq \theta_1\left(r^D\right)$, implying that the probability of insolvency will be higher than in the non-rollover case but lower than in the disclosure only case. The following proposition summarises these results.

Proposition 35 *Rollover increases the quality of information but at the cost of also increasing interest rates and the probability of insolvency of the bank, with respect to the non-rollover case.*

3.5 Discussion

3.5.1 Diversifying the bank's technology choice

In this chapter, I have studied how the probability of insolvency of the bank and the regulator's closure policy change when using and combining different instruments of market discipline.

A bank's probability of insolvency can be regarded as an indirect means of evaluating its risk-taking. The assumption on the technology choice of the bank made here can be limiting, although it considerably simplifies the algebra and allows for important conclusions to be drawn. In fact, by constraining the bank to place all its funds in

[14] When $n = 0$, θ is not perfectly observable but it is known to be higher than $\overline{\theta_2} > \overline{\theta_1}$.

one risky asset, the model does not allow for studying assets substitution. A possible extension might be to introduce two risky technologies, defined over the same support $[0, 1]$ but such that $R_2(0) \ll R_1(0)$, $R_2(1) \gg R_1(1)$, and $E[R_2(\theta)] = E[R_1(\theta)]$ (figure 3-6).

Given the guarantee on deposits and because of limited liability, the bank manager will adopt the riskier technology $(R_2(\theta))$. The aim of the model will then be to study whether disclosure or subordinated debt requirements could modify the risk choice of the bank towards the less risky asset $(R_1(\theta))$. For that, it will be necessary to modify the first stage of the game in the spirit of Niu's (2008) work. In the initial period, the manager will have to commit to a technology before offering deposits and the subordinated bond. Bond holders will move first, demanding an appropriate compensation for the bond, and then the rest of the consumers will take deposits. It is not clear what the equilibrium of this game will be. On the one hand, the solvency and profitability thresholds (θ_1 and θ_2, respectively) could be smaller under the riskier technology (figure 3-6) but, on the other hand, this technology will generate a higher payoff demanded by bond holders. Moreover, any result will be subject to the credibility of the manager's commitment to invest in the declared technology (Blum, 2000).[15]

3.5.2 Monitoring costs

I have assumed the manager's monitoring costs to be nil. However, as in Hortala-Vallve (2002) the size of the noise could be inversely proportional to the manager's monitoring effort, so that the smaller the noise the higher the disutility of monitoring experienced by her. This could push in the direction of more noise (a higher ε_m) and higher interest rates, reducing equity holders' expected payoff and increasing the closure rule, so more solvent banks could be inefficiently liquidated in the interim period.

The monitoring role of equity holders was overlooked in this exercise by assuming that there were no agency problems between them and the manager. An obvious extension will then be to lift this assumption, by assuming the manager's monitoring costs to be non-nil and allowing equity holders to observe the manager's signal.

3.5.3 Other extensions

Capital requirements could be introduced in this framework, in order to study the combined effects of these policies on the regulator's closure rule and the probability of

[15] In Blum (2000), market discipline reduces risk taking when a bank can credibly commit to a given level of risk before setting interest rates on deposits.

insolvency of the bank.

In this model, it is very important that bond holders are senior to equity holders, otherwise the latter will get a subsidy in default (in the region $[\overline{\theta_2}, \theta_S[)$, increasing moral hazard. Also, the model assumed that any funds on top of the guarantee that could be recovered from the liquidation of the banks' assets would be allocated to deposit holders, in order to compensate them for foregone interests (the guarantee covers only capital). Another option might be to prorate these funds equally between deposit holders and bond holders. On the one hand, this might increase bond holders' moral hazard, reducing monitoring (in the rollover case). On the other hand, their future payoff might rise, making less likely their withdrawal in the interim period, and so diminishing the probability of insolvency.

I modelled the rollover decision by giving bond holders the option of redeeming the bond in the interim period at no loss ($K = 1$). If a loss were made ($K < 1$) a rollover would be more likely, not only because the interim period payoff is lower but also because the withdrawal of one bond holder will increase the others' future payoff. Another possibility would be to introduce a secondary market for bonds in the interim period. In that case a new signal, the price of the bond, would be available from the market even if bonds could not be redeemed at $t = 1$.

Systemic risk considerations (too big or too many to fail) would affect the closure rule, possibly reducing this threshold if the social cost of the failure of a systemically important institution were included in the objective function of the regulator. Too big to fail problems could be addressed here by studying, for example, how the closure thresholds and equilibrium interest rates change with the exogenous parameters W_0, D_0 and B_0.

Demand driven liquidity shocks were eliminated by the introduction of deposit insurance. However, the market for the liquidation of the bank's asset in the interim period could suffer an exogenous liquidity shock, so that instead of obtaining one unit of consumption per unit of investment liquidated, the return could be lower. In that scenario, it would be interesting to study the role of the regulator as a "lender of last resort" or as a "market maker of last resort".

3.6 Conclusions

In this chapter, I have studied how the bank's probability of insolvency and the regulator's closure policy change when using and combining different instruments of market discipline. Given that depositors are dispersed and uninformed, the regulator needs to introduce deposit insurance in order to rule out sunspot equilibriums involving inefficient liquidation of assets (runs when the bank is solvent). Moral hazard forces the regulator to supervise the bank and the market, in order to get information allowing her to decide, in a cost efficient manner, whether to permit a bank to continue in operation or to suspend its licence, liquidate its assets and pay out to the depositors the guarantee in the interim period.

With a disclosure requirement, I prove that the regulator uses a fully financed closure rule that avoids inefficient survival but involves inefficient liquidation of assets in a region of the bank's fundamentals. Although the probability of insolvency is higher than the first best, it converges on it as the manager's signal noise goes to zero.

When on top of the disclosure requirement the bank is asked to issue a zero coupon bond with long maturity, even in a very small amount, the probability of insolvency is further reduced. The regulator's closure policy in this case forces fewer banks into liquidation, given the subsidy from bond holders to deposit holders when the bank becomes fundamentally unprofitable.

Finally, when debt can be rolled over in the interim period, and so bond holders do monitor the bank, this period sub-game equilibrium provides very useful information that, in some cases, can completely eliminate inefficient liquidation of assets because the regulator can perfectly observe the bank's fundamental. However, this will come at the cost of a higher probability of insolvency than in the non-rollover case.

The results in this chapter allow us to conclude that a subordinated debt requirement, when issued with long maturity, is able to reduce the probability of insolvency for any size of noise. Indeed, intuition suggests that when the manager's signal noise is expected to be high, an appropriately high subordinated debt requirement could restore the first best. On the other hand, a subordinated bond issued with a short maturity can substantially improve on the quality of information (by reducing the noise). Therefore, a subordinated debt requirement can be used to complement disclosure requirements, providing a new set of information which is useful to the regulator.

3.7 Appendix

3.7.1 Density function of $\widetilde{\theta}$

$\widetilde{\theta} = \theta + \varepsilon$ is the sum of two independent uniformly distributed random variables, $\theta \sim U[0,1]$ and $\varepsilon \sim U[-\varepsilon_m, \varepsilon_m]$. Therefore, its density of probability is equal to the convolution of

$$f_\theta(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & \sim \end{cases} \text{ and } f_\varepsilon(x) = \begin{cases} \frac{1}{2\varepsilon_m} & x \in [-\varepsilon_m, \varepsilon_m] \\ 0 & \sim \end{cases},$$

the density of probability of θ and ε, respectively. Notice $\widetilde{\theta}$ may take values in $[-\varepsilon_m, 1+\varepsilon_m]$.

$$\begin{aligned} f_{\widetilde{\theta}}(z) &= \int_{-\varepsilon_m}^{1+\varepsilon_m} f_\theta(y) f_\varepsilon(z-y)dy \\ &= \int_{-\varepsilon_m}^{0} \underbrace{f_\theta(y)}_{0} f_\varepsilon(x-y)dy + \int_{0}^{1} \underbrace{f_\theta(y)}_{1} f_\varepsilon(z-y)dy + \int_{1}^{1+\varepsilon_m} \underbrace{f_\theta(y)}_{0} f_\varepsilon(z-y)dy \end{aligned}$$

Using the following change of variable: $u = z - y \Rightarrow du = -dy$, $u(0) = z$, $u(1) = z-1$.

$$f_{\widetilde{\theta}}(z) = \int_{z-1}^{z} f_\varepsilon(u)du = \begin{cases} \int_{-\varepsilon_m}^{\min\{z,\varepsilon_m\}} \frac{1}{2\varepsilon_m} du & if\ z \leq 1-\varepsilon_m \\ \int_{z-1}^{\min\{z,\varepsilon_m\}} \frac{1}{2\varepsilon_m} du & if\ z > 1-\varepsilon_m \end{cases}$$

$$f_{\widetilde{\theta}}(z) = \begin{cases} \frac{1}{2\varepsilon_m}\left(\min\{z,\varepsilon_m\} + \varepsilon_m\right) & if\ z \leq 1-\varepsilon_m \\ \frac{1}{2\varepsilon_m}\left(\min\{z,\varepsilon_m\} - z + 1\right) & if\ z > 1-\varepsilon_m \end{cases}$$

This function is depicted in figure 3-7. Hence, the probability that a signal is below the closure rule can be written as (taking the case of a small noise signal, such that $\theta_1 + \varepsilon_m \leq 1 - \varepsilon_m$):

$$\begin{aligned} prob\left[\widetilde{\theta} < \widetilde{\theta}_C\right] &= \int_{-\varepsilon_m}^{\theta_1+\varepsilon_m} \frac{1}{2\varepsilon_m}\left(\min\{z,\varepsilon_m\} + \varepsilon_m\right) dz \\ &= \int_{-\varepsilon_m}^{\varepsilon_m} \frac{1}{2\varepsilon_m}(z+\varepsilon_m)\, dz + \int_{\varepsilon_m}^{\theta_1+\varepsilon_m} dz \\ &= \frac{1}{2\varepsilon_m} \frac{z^2}{2}\Big|_{-\varepsilon_m}^{\varepsilon_m} + \varepsilon_m + \theta_1 \\ &= \theta_1 + \varepsilon_m \end{aligned}$$

Finally, notice that $\lim_{\varepsilon_m \to 0} prob\left[\widetilde{\theta} < \widetilde{\theta}_C\right] = \theta_1$.■

3.7.2 Interest rate with deposit insurance and no monitoring

Denote by r^0 the optimal interest rate in problem (P_0), and by r^I the solution to:

(P_I) min r

st.

(3.5) $E\left[u(r,g)\right] \geq \overline{u}$

(3.6) $E\left[v(r,g)\right] \geq \overline{v}$

(3.7) $r \geq \frac{1-D_0-B_0}{D_0+B_0}$

(3.8) $r \leq \frac{1-\pi D_0}{\pi D_0}$

With deposit insurance there are no runs, hence the bank can only go bankrupt in the second period. Notice problems (P_0) and (P_I) are identical, except for the incentive compatibility constraints of deposit holders. In the optimum, these must be satisfied with equality, so $E[u(r^I,g)] = \overline{u} = E[u(r^0)]$.

$$E[u(r^I,g)] = \theta_1 u\left(\tfrac{1}{D_0+B_0}\right) - \theta_1 u\left(\tfrac{1}{D_0+B_0}\right) + \pi\left(1-\theta_1\right)u\left(1+r\right) - \pi\left(1-\theta_1\right)u\left(1+r\right)$$
$$+\pi u\left(1+r\right) + \left(1-\pi\right)\theta_1 u\left(1+r\right) + \left(1-\pi\right)\int_{\theta_1}^{1} u\left(\min\left\{c_2\left(\theta,r\right),\left(1+r\right)^2\right\}\right)d\theta$$

$$E[u(r^I,g)] = E[u(r^I)] + \theta_1\left\{\underbrace{u\left(1+r\right) - \pi u\left(\tfrac{1}{D_0+B_0}\right)}_{>0}\right\} = E[u(r^0)]$$

$\Longrightarrow E[u(r^I)] < E[u(r^0)] \Longrightarrow r^I < r^0.$

The same result arises when looking at the incentive compatibility constraint of investors.■

3.7.3 Proof of proposition 28

Denote by r^0 the optimal interest rate in the benchmark case and by r^D the one obtained with deposit insurance and disclosure. In equilibrium, the manager will choose an interest rate such that the participation constraints of deposit holders bind. For the case of the risk averse depositor type consumers:

$$E\left[u(r^0)\right] = E_{\widetilde{\theta}_C}\left[u(r^D,g)\right] = \overline{u}$$

Consider the case $\theta_1(r) + 2\varepsilon_m \leq \theta_2(r)$ (so that taking limit when $\varepsilon_m \to 0$ has sense). Denote by $\rho\left(\theta,\varepsilon_m\right) = prob\left[\widetilde{\theta} < \widetilde{\theta}_C\right]$:

$$E_{\widetilde{\theta}_C}\left[u(r,g)\right] = \pi\left\{\int_0^{\theta_1(r)} u\left(\tfrac{1}{D_0+B_0}\right)d\theta + \int_{\theta_1(r)}^{\theta_1(r)+2\varepsilon_m} \rho\left(\theta,\varepsilon_m\right)u\left(\tfrac{1}{D_0+B_0}\right)d\theta\right.$$
$$\left. + \int_{\theta_1(r)}^{\theta_1(r)+2\varepsilon_m}\left(1-\rho\left(\theta,\varepsilon_m\right)\right)u\left(1+r\right)d\theta + \int_{\theta_1(r)+2\varepsilon_m}^{1} u\left(1+r\right)d\theta\right\}$$
$$+\left(1-\pi\right)\left\{\int_0^{\theta_1(r)} u\left(\tfrac{1}{D_0+B_0}\right)d\theta + \int_{\theta_1(r)}^{\theta_1(r)+2\varepsilon_m} \rho\left(\theta,\varepsilon_m\right)u\left(\tfrac{1}{D_0+B_0}\right)d\theta\right.$$

$$
\begin{aligned}
&+\int_{\theta_1(r)}^{\theta_1(r)+2\varepsilon_m}\left(1-\rho\left(\theta,\varepsilon_m\right)\right)u\left(c_2(\theta,r)\right)d\theta+\int_{\theta_1(r)+2\varepsilon_m}^{\theta_2(r)}u\left(c_2(\theta,r)\right)d\theta \\
&\left.+\int_{\theta_2(r)}^{1}u\left((1+r)^2\right)d\theta\right\}
\end{aligned}
$$

$$
\begin{aligned}
E_{\widetilde{\theta}_C}\left[u(r,g)\right]=&\ \theta_1(r)u\left(\tfrac{1}{D_0+B_0}\right)+\pi\left(1-\theta_1(r)\right)u\left(1+r\right) \\
&+(1-\pi)\left\{\int_{\theta_1(r)}^{\theta_2(r)}u(c_2(\theta,r))d\theta+\int_{\theta_2(r)}^{1}u\left((1+r)^2\right)d\theta\right\} \\
&-\int_{\theta_1(r)}^{\theta_1(r)+2\varepsilon_m}\left[\pi u\left(1+r\right)+(1-\pi)\,u\left(c_2(\theta,r)\right)-u\left(\tfrac{1}{D_0+B_0}\right)\right]d\theta \\
&+\int_{\theta_1(r)}^{\theta_1(r)+2\varepsilon_m}\left(1-\rho\left(\theta,\varepsilon_m\right)\right)\left[\pi u\left(1+r\right)+(1-\pi)\,u\left(c_2(\theta,r)\right)-u\left(\tfrac{1}{D_0+B_0}\right)\right]d\theta
\end{aligned}
$$

$$
\begin{aligned}
E_{\widetilde{\theta}_C}\left[u(r,g)\right]=&\ \theta_1(r)u\left(\tfrac{1}{D_0+B_0}\right)+\pi\left(1-\theta_1(r)\right)u\left(1+r\right) \\
&+(1-\pi)\left\{\int_{\theta_1(r)}^{\theta_2(r)}u(c_2(\theta,r))d\theta+\int_{\theta_2(r)}^{1}u\left((1+r)^2\right)d\theta\right\} \\
&-\int_{\theta_1(r)}^{\theta_1(r)+2\varepsilon_m}\rho\left(\theta,\varepsilon_m\right)\left[\pi u\left(1+r\right)+(1-\pi)\,u\left(c_2(\theta,r)\right)-u\left(\tfrac{1}{D_0+B_0}\right)\right]d\theta
\end{aligned}
$$

Define $\Phi(r,\varepsilon_m)\doteq\int_{\theta_1(r)}^{\theta_1(r)+2\varepsilon_m}\rho\left(\theta,\varepsilon_m\right)\left[\underbrace{\pi u\left(1+r\right)+(1-\pi)\,u\left(c_2(\theta,r)\right)-u\left(\tfrac{1}{D_0+B_0}\right)}_{>0}\right]d\theta.$

Notice $\forall r:\Phi(r,\varepsilon_m)>0$ if $\varepsilon_m>0$ and $\Phi(r^D,\varepsilon_m)=0$ if $\varepsilon_m=0$.

In the benchmark model (without deposit insurance and ruling out inefficient runs):

$$
\begin{aligned}
E\left[u(r)\right]=&\ \pi\left\{\int_0^{\theta_1(r)}u(\tfrac{1}{D_0+B_0})d\theta+\int_{\theta_1(r)}^{1}u(1+r)d\theta\right\} \\
&+(1-\pi)\left\{\int_0^{\theta_1(r)}u(\tfrac{1}{D_0+B_0})d\theta+\int_{\theta_1(r)}^{\theta_2(r)}u(c_2(\theta,r))d\theta+\int_{\theta_2(r)}^{1}u\left((1+r)^2\right)d\theta\right\} \\
=&\ \theta_1(r)u(\tfrac{1}{D_0})+\pi\left(1-\theta_1(r)\right)u\left(1+r\right) \\
&+(1-\pi)\left\{\int_{\theta_1(r)}^{\theta_2(r)}u(c_2(\theta,r))d\theta+\int_{\theta_2(r)}^{1}u\left((1+r)^2\right)d\theta\right\}
\end{aligned}
$$

Then

$$
\begin{array}{rcll}
E_{\widetilde{\theta}_C}\left[u(r^D,g)\right]= & E\left[u(r^D)\right]-\Phi(r,\varepsilon_m) & =\overline{u} \\
 & E\left[u(r^0)\right] & =\overline{u} \\
\hline
\multicolumn{3}{l}{E\left[u(r^D)\right]-\Phi(r,\varepsilon_m)-E\left[u(r^0)\right]=0} \\
\multicolumn{3}{l}{\Rightarrow E\left[u(r^D)\right]-E\left[u(r^0)\right]=\Phi(r^D,\varepsilon_m)\geq 0}
\end{array}
$$

Given that $u(.)$ is increasing, this inequality implies $r^D>r^0$ for all $\varepsilon_m>0$ and $r^D\to r^0$ as $\varepsilon_m\to 0$. A similar result can be obtained for investors holding deposits, replacing $u(.)$ by the identity function in all equations and the competitive reservation utility by $\overline{v}$. Finally, given that $\frac{d\theta_1(r)}{dr}\geq 0$, $\theta_1\left(r^D\right)\geq\theta_1\left(r^0\right)$.■

3.7.4 $b_2(r, s_0)$ increasing on r

Denote by $\rho(\theta, \varepsilon_m) = prob[\widetilde{\theta} < \widetilde{\theta}_C^b]$, $\omega = \max\left\{\frac{1-(1+r)[D_0+B_0-s_0]}{s_0}, 0\right\}$ and $\sigma = \min\left\{\frac{1}{D_0+B_0-s_0}, 1+r\right\}$. Recalling $\widetilde{\theta}_C^b = \overline{\theta}_1 + \varepsilon_m$:

$$\begin{aligned}
E_{\widetilde{\theta}_C^b}[v(b_t)] &= \int_0^{\overline{\theta_1}(r)} \omega d\theta + \int_{\overline{\theta_1}(r)}^{\overline{\theta_1}(r)+2\varepsilon_m} \rho(\theta, \varepsilon_m)\omega d\theta + \int_{\overline{\theta_1}(r)}^{\overline{\theta_1}(r)+2\varepsilon_m} (1-\rho(\theta,\varepsilon_m)) \cdot 0 d\theta \\
&+ \int_{\overline{\theta_1}(r)+2\varepsilon_m}^{\overline{\theta_2}(r)} 0 d\theta + \int_{\overline{\theta_2}(r)}^{\theta_S(r)} \left\{\frac{(1-\pi D_0(1+r))R(\theta) - [(1-\pi)D_0+B_0-s_0](1+r)^2}{s_0}\right\} d\theta \\
&+ \int_{\theta_S(r)}^{1} b_2(r, s_0)\, d\theta
\end{aligned}$$

$$\begin{aligned}
E_{\widetilde{\theta}_C^b}[v(r,g,s_0)] &= \int_0^{\overline{\theta_1}(r)} \sigma d\theta + \int_{\overline{\theta_1}(r)}^{\overline{\theta_1}(r)+2\varepsilon_m} \rho(\theta,\varepsilon_m)\sigma d\theta \\
&+ \int_{\overline{\theta_1}(r)}^{\overline{\theta_1}(r)+2\varepsilon_m} (1-\rho(\theta,\varepsilon_m))\left\{c_2(\theta,r) + \frac{s_0(1+r)}{(1-\pi)D_0+B_0}\right\} d\theta \\
&+ \int_{\overline{\theta_1}(r)+2\varepsilon_m}^{\overline{\theta_2}(r)} \left\{c_2(\theta,r) + \frac{s_0(1+r)}{(1-\pi)D_0+B_0}\right\} d\theta \\
&+ \int_{\overline{\theta_2}(r)}^{\theta_S(r)} (1+r)^2 d\theta + \int_{\theta_S(r)}^{1} (1+r)^2 d\theta
\end{aligned}$$

In order to switch from deposits to bonds, investors require $E_{\widetilde{\theta}_C^b}[v(b_t)] \geq E_{\widetilde{\theta}_C^b}[v(r,g,s_0)]$.

$$\begin{aligned}
E_{\widetilde{\theta}_C^b}[v(b_t)] - E_{\widetilde{\theta}_C^b}[v(r,g,s_0)] &= \int_0^{\overline{\theta_1}(r)} \underbrace{(\omega-\sigma)}_{<0} d\theta + \int_{\overline{\theta_1}(r)}^{\overline{\theta_1}(r)+2\varepsilon_m} \rho(\theta,\varepsilon_m)\underbrace{(\omega-\sigma)}_{<0} d\theta \\
&- \int_{\overline{\theta_1}(r)}^{\overline{\theta_1}(r)+2\varepsilon_m} \underbrace{(1-\rho(\theta,\varepsilon_m))\left\{c_2(\theta,r) + \frac{s_0(1+r)}{(1-\pi)D_0+B_0}\right\}}_{>0} d\theta \\
&- \int_{\overline{\theta_1}(r)+2\varepsilon_m}^{\overline{\theta_2}(r)} \left\{\underbrace{c_2(\theta,r) + \frac{s_0(1+r)^2}{(1-\pi)D_0+B_0}}_{>0}\right\} d\theta \\
&- \int_{\overline{\theta_1}(r)+2\varepsilon_m}^{\overline{\theta_2}(r)} \left\{\underbrace{c_2(\theta,r) + \frac{s_0(1+r)^2}{(1-\pi)D_0+B_0}}_{>0}\right\} d\theta \\
&+ \int_{\overline{\theta_2}(r)}^{\theta_S(r)} \left\{\underbrace{\frac{(1-\pi D_0(1+r))R(\theta) - [(1-\pi)D_0+B_0-s_0](1+r)^2}{s_0} - (1+r)^2}_{A}\right\} d\theta \\
&+ \underbrace{(1-\theta_S(r))\left(b_2(r,s_0) - (1+r)^2\right)}_{>0}
\end{aligned}$$

$$A = \left[\frac{(1-\pi)D_0+B_0}{s_0}\right]\left(c_2(\theta,r) - (1+r)^2\right)\begin{cases} <0 & \overline{\theta_2}(r) \leq \theta < \theta_2(r) \\ >0 & \theta_2(r) < \theta \leq \theta_S(r) \end{cases}$$

Therefore, for investors to be willing to buy the subordinated bond, $b_2(r, s_0)$ will need to be sufficiently higher than $(1+r)^2$; and as $\overline{\theta_2}(r)$ and $\theta_S(r)$ are both increasing on r, so must be $b_2(r, s_0)$. ■

3.7.5 Proof of proposition 30

Denote by r^D the optimal interest rate with deposit insurance and disclosure and by r^B the one obtained when adding a long-term subordinated bond issued in an amount $s_0 \geq 0$. In equilibrium, the additional equation in program (P_B) (equation 3.15) will be satisfied with equality, defining b_2 as a function of r. Given the previous result (section 3.7.4), the manager will have an additional incentive for reducing r because that would reduce the payoff demanded by bondholders, increasing her final profit. However, deposit holders' participation constraints (equations 3.13 and 3.14) have also changed.

We know that in equilibrium $E_{\widetilde{\theta}_C^b}\left[u(r^B, g, s_0)\right] = E_{\widetilde{\theta}_C}\left[u(r^D, g)\right] = \overline{u}$. Consider the case of a small noise, so that $\overline{\theta_1}(r) + 2\varepsilon_m \leq \overline{\theta_2}(r)$ and $\theta_1(r) + 2\varepsilon_m \leq \theta_2(r)$, denote by $\rho(\theta, \varepsilon_m) = prob[\widetilde{\theta} < \textit{closure rule}]$ and by $\sigma = \min\left\{\frac{1}{D_0+B_0-s_0}, 1+r\right\} \geq \frac{1}{D_0+B_0}$.

$$
\begin{aligned}
E_{\widetilde{\theta}_C^b}\left[u(r,g,s_0)\right] = \pi \Bigg\{ & \int_0^{\overline{\theta_1}(r)} u(\sigma)\, d\theta + \int_{\overline{\theta_1}(r)}^{\overline{\theta_1}(r)+2\varepsilon_m} \rho(\theta,\varepsilon_m)\, u(\sigma)\, d\theta \\
& + \int_{\overline{\theta_1}(r)}^{\overline{\theta_1}(r)+2\varepsilon_m} (1-\rho(\theta,\varepsilon_m))\, u(1+r)\, d\theta + \int_{\overline{\theta_1}(r)+2\varepsilon_m}^{1} u(1+r)\, d\theta \Bigg\} \\
& + (1-\pi) \Bigg\{ \int_0^{\overline{\theta_1}(r)} u(\sigma)\, d\theta + \int_{\overline{\theta_1}(r)}^{\overline{\theta_1}(r)+2\varepsilon_m} \rho(\theta,\varepsilon_m)\, u(\sigma)\, d\theta \\
& + \int_{\overline{\theta_1}(r)}^{\overline{\theta_1}(r)+2\varepsilon_m} (1-\rho(\theta,\varepsilon_m))\, u\left(c_2(\theta,r) + \tfrac{s_0(1+r)}{(1-\pi)D_0+B_0}\right) d\theta \\
& + \int_{\overline{\theta_1}(r)+2\varepsilon_m}^{\overline{\theta_2}(r)} u\left(c_2(\theta,r) + \tfrac{s_0(1+r)}{(1-\pi)D_0+B_0}\right) d\theta + \int_{\overline{\theta_2}(r)}^{1} u\left((1+r)^2\right) d\theta \Bigg\}
\end{aligned}
$$

$$
\begin{aligned}
E_{\widetilde{\theta}_C^b}\left[u(r,g,s_0)\right] = {} & \overline{\theta_1}(r)u(\sigma) + \pi\left(1-\overline{\theta_1}(r)\right)u(1+r) \\
& + (1-\pi)\left\{ \int_{\overline{\theta_1}(r)}^{\overline{\theta_2}(r)} u\left(c_2(\theta,r) + \tfrac{s_0(1+r)}{(1-\pi)D_0+B_0}\right) d\theta + \int_{\overline{\theta_2}(r)}^{1} u\left((1+r)^2\right) d\theta \right\} \\
& - \int_{\overline{\theta_1}(r)}^{\overline{\theta_1}(r)+2\varepsilon_m} \rho(\theta,\varepsilon_m)\left[\pi u(1+r) - u(\sigma)\right] d\theta \\
& - \int_{\overline{\theta_1}(r)}^{\overline{\theta_1}(r)+2\varepsilon_m} \rho(\theta,\varepsilon_m)(1-\pi)\, u\left(c_2(\theta,r) + \tfrac{s_0(1+r)}{(1-\pi)D_0+B_0}\right) d\theta
\end{aligned}
$$

Replacing $\rho(\theta, \varepsilon_m)$ by its formula in each case (section 3.7.1), and using $E_{\widetilde{\theta}_C}\left[u(r,g)\right]$ from section 3.7.3:

$$
\begin{aligned}
&E_{\widetilde{\theta}_C}\left[u(r,g)\right]-E_{\widetilde{\theta}_C^b}\left[u(r,g,s_0)\right]\\
&\quad\le\left(\theta_1(r)-\overline{\theta_1}(r)\right)u\left(\tfrac{1}{D_0+B_0}\right)-\pi u\left(1+r\right)\left(\theta_1(r)-\overline{\theta_1}(r)\right)\\
&\quad+(1-\pi)\left\{\int_{\theta_1(r)}^{\theta_2(r)}u(c_2(\theta,r))d\theta-\int_{\overline{\theta_1}(r)}^{\overline{\theta_2}(r)}u\left(c_2(\theta,r)+\tfrac{s_0(1+r)}{(1-\pi)D_0+B_0}\right)d\theta\right\}\\
&\quad+(1-\pi)\left\{\int_{\theta_2(r)}^{1}u\left((1+r)^2\right)d\theta-\int_{\overline{\theta_2}(r)}^{1}u\left((1+r)^2\right)d\theta\right\}\\
&\quad-\int_{\theta_1(r)}^{\theta_1(r)+2\varepsilon_m}\left(\theta_1\left(r\right)+\varepsilon_m\right)\left[\pi u\left(1+r\right)-u\left(\tfrac{1}{D_0+B_0}\right)\right]d\theta\\
&\quad+\int_{\overline{\theta_1}(r)}^{\overline{\theta_1}(r)+2\varepsilon_m}\left(\overline{\theta_1}\left(r\right)+\varepsilon_m\right)\left[\pi u\left(1+r\right)-u\left(\tfrac{1}{D_0+B_0}\right)\right]d\theta\\
&\quad-\int_{\theta_1(r)}^{\theta_1(r)+2\varepsilon_m}\left(\theta_1\left(r\right)+\varepsilon_m\right)\left[(1-\pi)\,u\left(c_2(\theta,r)\right)\right]d\theta\\
&\quad+\int_{\overline{\theta_1}(r)}^{\overline{\theta_1}(r)+2\varepsilon_m}\left(\overline{\theta_1}\left(r\right)+\varepsilon_m\right)\left[(1-\pi)\,u\left(c_2(\theta,r)+\tfrac{s_0(1+r)}{(1-\pi)D_0+B_0}\right)\right]d\theta
\end{aligned}
$$

By definition, $\int_{\overline{\theta_1}(r)}^{\overline{\theta_2}(r)}u\left(c_2(\theta,r)+\tfrac{s_0(1+r)}{(1-\pi)D_0+B_0}\right)d\theta=\int_{\theta_1(r)}^{\theta_2(r)}u(c_2(\theta,r))d\theta$. So

$$
\begin{aligned}
&E_{\widetilde{\theta}_C}\left[u(r,g)\right]-E_{\widetilde{\theta}_C^b}\left[u(r,g,s_0)\right]\\
&\quad\le\left(\theta_1(r)-\overline{\theta_1}(r)\right)u\left(\tfrac{1}{D_0+B_0}\right)-\pi u\left(1+r\right)\left(\theta_1(r)-\overline{\theta_1}(r)\right)\\
&\quad-(1-\pi)\left(\theta_2(r)-\overline{\theta_2}(r)\right)u\left((1+r)^2\right)\\
&\quad-2\varepsilon_m\left(\theta_1\left(r\right)-\overline{\theta_1}\left(r\right)\right)\left[\pi u\left(1+r\right)-u\left(\tfrac{1}{D_0+B_0}\right)\right]\\
&\quad-\left(\theta_1\left(r\right)-\overline{\theta_1}\left(r\right)\right)\int_{\theta_1(r)}^{\theta_1(r)+2\varepsilon_m}(1-\pi)\,u\left(c_2(\theta,r)\right)d\theta
\end{aligned}
$$

Notice

$$
c_2(\theta_2,r)-c_2(\overline{\theta_2},r)=\tfrac{s_0(1+r)^2}{(1-\pi)D_0+B_0}>\tfrac{s_0(1+r)}{(1-\pi)D_0+B_0}=c_2(\theta_1,r)-c_2(\overline{\theta_1},r)
$$

which implies $\theta_2-\overline{\theta_2}\ge\theta_1-\overline{\theta_1}$.

Hence

$$
\begin{aligned}
&E_{\widetilde{\theta}_C}\left[u(r,g)\right]-E_{\widetilde{\theta}_C^b}\left[u(r,g,s_0)\right]\\
&\quad\le\left(\theta_1(r)-\overline{\theta_1}(r)\right)\Big\{(1-2\varepsilon_m)\left[u\left(\tfrac{1}{D_0+B_0}\right)-\pi u\left(1+r\right)\right]\\
&\quad-\int_{\theta_1(r)}^{\theta_1(r)+2\varepsilon_m}(1-\pi)\,u\left(c_2(\theta,r)\right)d\theta\Big\}\\
&\quad-\left(\theta_1(r)-\overline{\theta_1}(r)\right)(1-\pi)\,u\left((1+r)^2\right)
\end{aligned}
$$

$$\leq \left(\theta_1(r) - \overline{\theta_1}(r)\right)\left\{(1-2\varepsilon_m)\left[u\left(\tfrac{1}{D_0+B_0}\right) - \pi u(1+r)\right]\right.$$
$$\left. -(1-2\varepsilon_m)(1-\pi)u(1+r)\right\}$$

$$= \left(\theta_1(r) - \overline{\theta_1}(r)\right)(1-2\varepsilon_m)\left[u\left(\tfrac{1}{D_0+B_0}\right) - u(1+r)\right] \leq 0$$

Define

$$-\overline{\Phi}(r,\varepsilon_m) = \left(\theta_1(r) - \overline{\theta_1}(r)\right)(1-2\varepsilon_m)\left[u\left(\tfrac{1}{D_0+B_0}\right) - u(1+r)\right] \leq 0$$

Therefore

$$E_{\widetilde{\theta}_C}\left[u(r^B,g)\right] + \overline{\Phi}(r^B,\varepsilon_m) \leq E_{\widetilde{\theta}_C^b}\left[u(r^B,g,s_0)\right] = \overline{u} = E_{\widetilde{\theta}_C}\left[u(r^D,g)\right]$$
$$\Rightarrow E_{\widetilde{\theta}_C}\left[u(r^B,g)\right] \leq E_{\widetilde{\theta}_C}\left[u(r^D,g)\right] \Rightarrow r^D \geq r^B \text{ for all } \varepsilon_m \geq 0.$$

A similar result can be obtained for investors holding deposits, replacing $u(.)$ by the identity function in all equations and the competitive reservation utility by $\overline{v}$.

Finally, the probability of insolvency is smaller because

$$\overline{\theta_1}\left(r^B\right) \leq \theta_1\left(r^B\right) \leq \theta_1\left(r^D\right). \blacksquare$$

3.7.6 Proof of proposition 32

Recall

$$\left(1-\pi D_0(1+r)\right)R\left(\overline{\theta_2}\right) = \left[(1-\pi)D_0 + B_0 - s_0\right](1+r)^2$$

and

$$\left(1-\pi D_0(1+r)\right)R(\theta_1) = \left[(1-\pi)D_0 + B_0\right](1+r).$$

Then

$$\frac{R(\overline{\theta_2})}{R(\theta_1)} = \frac{[(1-\pi)D_0+B_0-s_0](1+r)}{(1-\pi)D_0+B_0} \leq 1$$

$$\Leftrightarrow \left[(1-\pi)D_0 + B_0 - s_0\right](1+r) \leq (1-\pi)D_0 + B_0$$

$$\Leftrightarrow \left[(1-\pi)D_0 + B_0\right]\frac{r}{1+r} \leq s_0$$

$$\Leftrightarrow R\left(\overline{\theta_2}\right) \leq R(\theta_1) \Leftrightarrow \overline{\theta_2} \leq \theta_1. \blacksquare$$

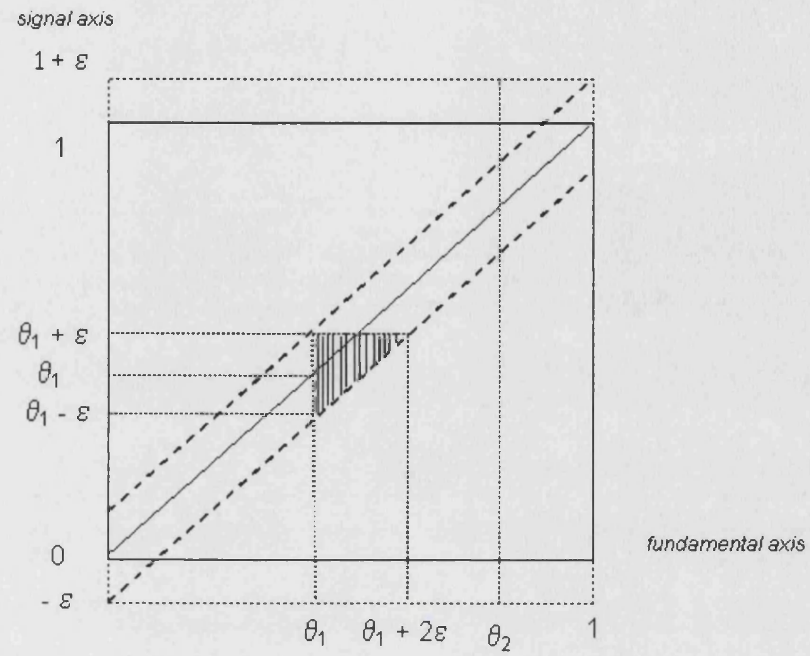

Figure 3-1: Proof of proposition 28.

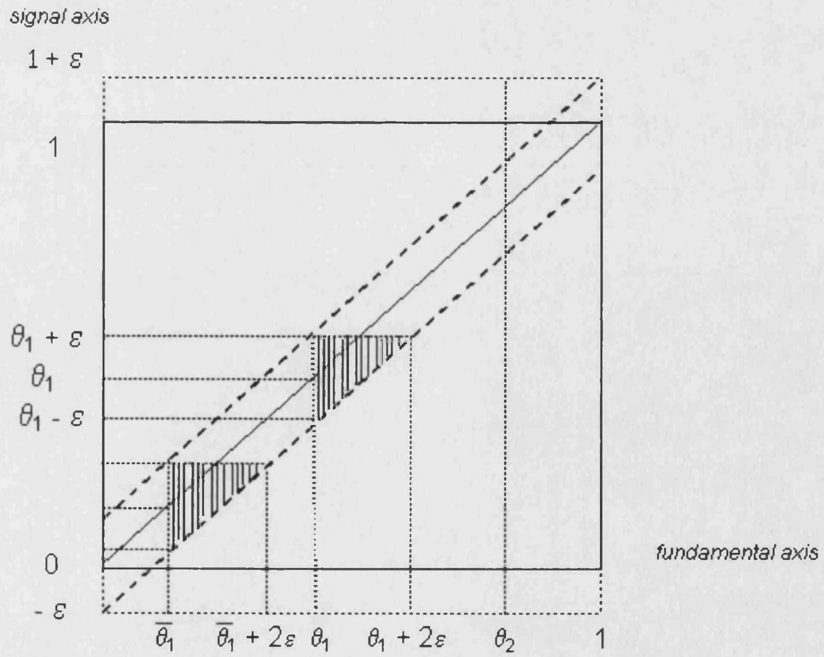

Figure 3-2: Proof of proposition 30.

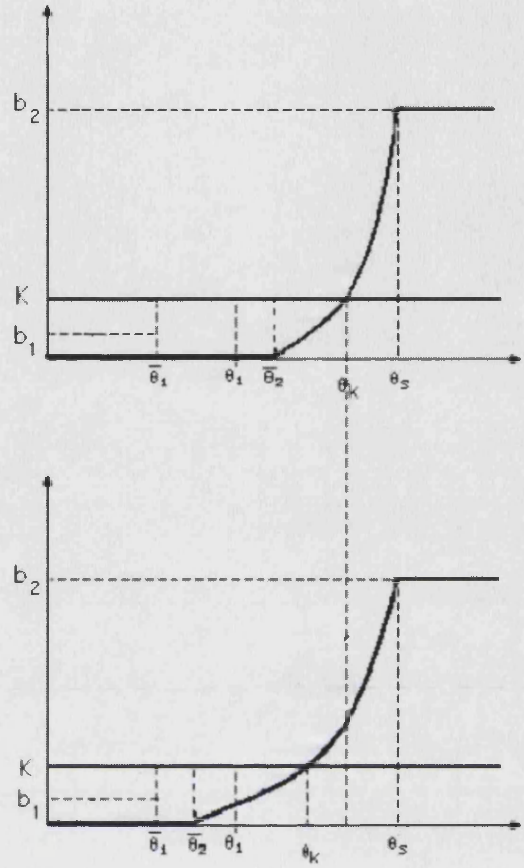

Figure 3-3: Bond holders' payoffs.

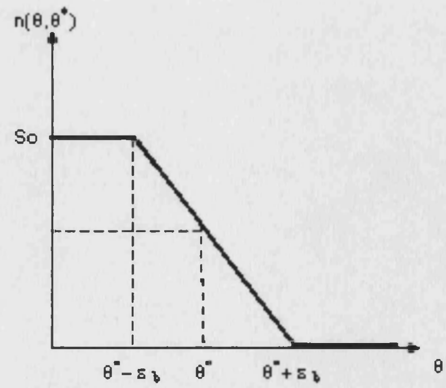

Figure 3-4: Number of bond holders redeeming the bond in the interim period.

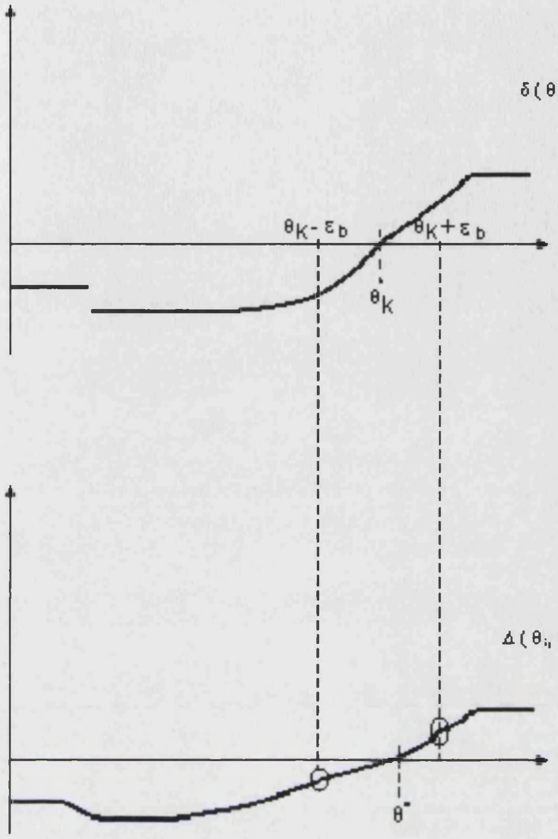

Figure 3-5: Bond holders' conditional expected utility: redemption versus rollover.

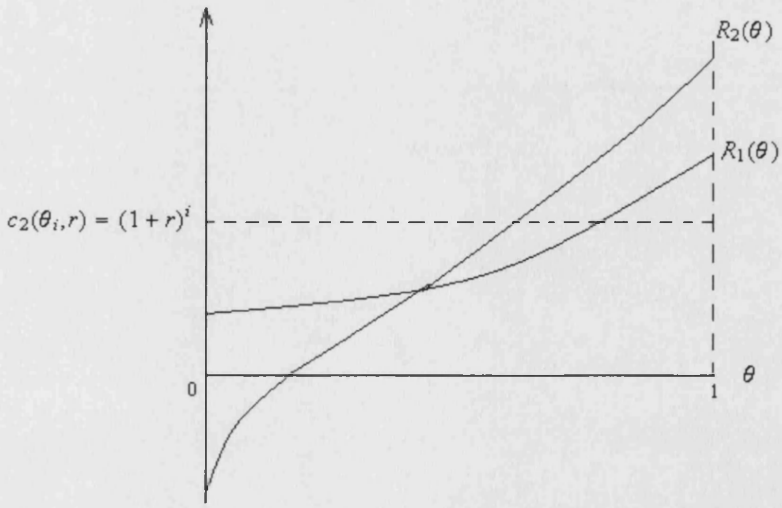

Figure 3-6: Example of two risky technologies with the same mean and different variances.

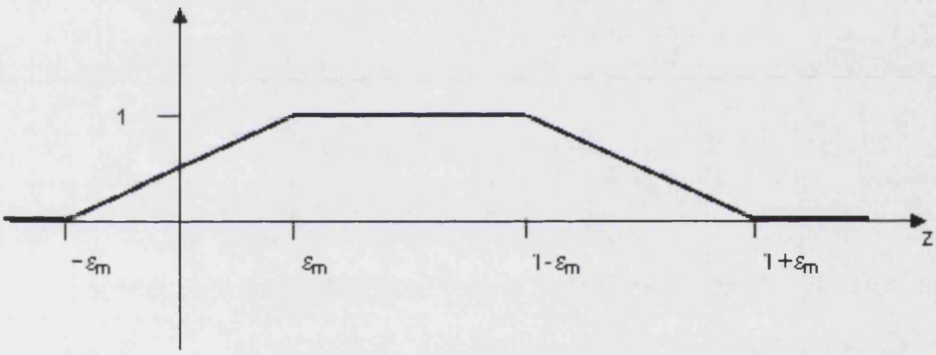

Figure 3-7: Density function of $\widetilde{\theta}$: case $\varepsilon_m < \frac{1}{2}$.

Chapter 4

Capital Regulation and Bank Risk-Taking. Completing Blum's Picture

4.1 Introduction

The Basel Committee was established at the end of 1974 by the central bank governors of the Group of Ten countries, with the aim of gathering central bankers and bank supervisors and regulators to discuss issues related to prudential banking supervision. As a result of these talks, in 1988 emerged the first version of the Basel Capital Accord, introducing a common minimum 8 percent risk-weighted capital to asset ratio for internationally active G-10 banks, which in the earliest version only considered credit risk. Although some countries had adopted minimum requirements before the agreement (the USA and the UK in 1981, for example), it was only after the agreement that capital requirements became common ground for the banking industry worldwide.[1]

The objective of this form of regulation is said to be to strengthen the soundness and stability of the international banking system, and to reduce competitive inequalities across markets. However, some theoretical results suggest that banks have found ways of overcoming the limitations that fixed capital requirements impose on their risk-taking relative to capital, either through asset substitution (Koehn and Santomero, 1980; Kim and Santomero, 1988; Flannery, 1989; Rochet, 1992), the reduction of monitoring incen-

[1] In 1993 all commercial banks in the European union were subject to a common solvency requirement. By 1999 the Basel capital accord was being implemented in about 100 countries. Indeed, since the introduction of the capital accord, risk weighted capital ratios in developed countries have increased significantly. Nonetheless, it is not clear whether this responds to regulation itself or to increased market discipline (Jackson, 1999).

tives (Besanko and Kanatas, 1993; Boot and Greenbaum, 1993) or through substantial volumes of securitisation (Jones, 2000).

The empirical evidence as to whether capital requirements reduce the probability of default or induce banks to increase risk-taking in some periods is not conclusive. In a study for 98 USA banks over the period 1975-1986, Furlong (1988) inverts the Black and Scholes (1973) pricing formula to infer the volatility of the portfolio assets of banks. He concludes that volatility was higher after the introduction of capital requirements in 1981, though it grew both for badly and well capitalised banks. In a different study over 219 G-10 banks in the period 1987-1994, Sheldon (1996) finds that while the volatility of US banks increased in the period, independent of the level of the capital requirement, that of Japanese banks fell as capital ratios rose.

On the theoretical side the picture is blurry too. The work of Kahane (1977), Kareken and Wallace (1978) and Sharpe (1978), justifies the use of capital requirements to control the solvency of banks whose asset allocation is distorted by the presence of deposit insurance, but both assume complete markets. Under an incomplete market approach, Koehn and Santomero (1980) and Kim and Santomero (1988), using a mean variance portfolio model with fixed liabilities, prove that in the absence of a solvency requirement and abstracting from the limited liability clause, the probability of bank failure is a decreasing function of its capital ratio, which is independent of the (non-negative) weights used in the computation of the ratio. However, the introduction of capital requirements changes the asset allocation of the bank, so that while the volume of the risky portfolio decreases (because the bank shifts to those assets within a lower weight category), its composition is distorted in the direction of more risk (inside the chosen weight category), increasing the probability of failure. As a way of correcting this problem, they propose the introduction of risk weights proportional to the systemic risk of assets.

Since then, the literature has given a lot of attention to market based refinements on risk weights.[2] For example, Thakor (1996) shows how a bad selection of risk weights could have a negative impact on the real sector through credit crunches, given that the asset allocation of a bank can be distorted by the difference between market and regulatory assessments of asset risks. Furfine (2001) uses a panel of large US banks between 1990 and 1997, and a structural dynamic model of bank behaviour to show that the credit crunch in the USA in the 1990s could be explained by increasing non-market based risk weighted capital requirements and excessive regulatory monitoring,

[2] Compare the great deal of attention devoted to it in the last release of Basel II (BIS, 2006). More than a half of the document goes about the way of calculating appropriate risk weighted capital ratios (Pillar I), both in a standardized and non-standardized fashion.

instead of pure demand effects.

However, when limited liability is taken into account, Rochet (1992) shows that, even with the correct weights, capital requirements are not enough to control for moral hazard and that additional requirements, in the form of minimum levels of capital independent of the size of the assets, may be needed.

Furlong and Keeley (1989) advocate capital requirements, arguing that when limited liability and the option value of (flat) deposit insurance are properly taken into account, a bank that maximises the value of its stock, and therefore diversifies its portfolio, will always reduce risk with more stringent capital requirements. The same result is obtained by Santos (1999), in a model that considers asymmetric information between the bank and the borrowing firm, and the distortions induced by the presence of deposit insurance on the optimal funding contract. More stringent capital requirements make the bank ask for a (larger) equity stake in the firm, which in turn induces the firm to lower its risk, reducing the bank's probability of default.

Nonetheless, a static framework fails to capture important intertemporal effects that capital requirements might have on the behaviour of banks. One of the first theoretical models studying the intertemporal effects of capital constraints is given by Blum (1999). In a discrete time model he studies the incentives for asset substitution coming from the reduction in expected profits imposed by the requirement. In order to raise the amount of equity in the following period, a bank may find it optimal to increase risk today, in which case strengthening the requirement would have the opposite effect for which it was designed, to curb bank risk-taking.

In this chapter I will build on Blum's model to obtain some important lessons neglected in his original work.[3] I will develop threshold values for which capital regulation becomes binding in each period, and study how the regulated equilibrium is compared to the unregulated solution and to the social optimum, the effects of regulation over financial intermediation, and its impact on the distribution of risk among banks.

This model has the advantage of allowing for the study of asset substitution in a much simpler fashion than it could have been done in the Diamond and Dybig (1983) setup used in the previous chapters: the bank manager can choose among different returns of the portfolio, which have associated a unique probability of default. Also, and unlike in the previous chapters, here I assume common knowledge, so that deposit insurance completely eliminates runs. Liquidity shocks are excluded, because it is ex-

[3] This is why the chapter is subtitled "completing Blum's picture". It is not a new model, but it brings into the scene some important pieces overlooked in the original paper.

ante known that in each period all deposits will be withdrawn for consumption and, finally, thanks to the assumption of universal risk neutrality, the risk choice of the bank becomes completely isolated from any other decision in the game.

I show that for unregulated banks risk-taking is decreasing on the level of initial equity, and converges on the social optimum when equity is sufficiently high. When introducing capital requirements, I show there exist critical threshold values in each period for which regulation starts binding. When the requirement binds in the initial period only, risk can be reduced below the unregulated solution – even to the social optimum for sufficiently tight regulation –, but fewer deposits are taken, which reduces financial intermediation. Moreover, capital requirements are not sufficient to control moral hazard because, among the binding banks, the better capitalised ones raise relatively more insured deposits and take on relatively more risk. When the requirement binds only in the interim period, bank risk-taking increases, most likely above the unregulated solution for all values of the requirement. In that case, risk would be decreasing on equity, making risk-taking even more aggressive for poorly capitalised banks. Therefore, interim period binding capital requirements will not only worsen the risk choice of banks, but make smaller banks weaker. When a constant capital requirement binds in both periods, the tighter the regulation the fewer deposits are taken from the public, though better capitalised banks raise relatively more deposits. The dynamic of risk in this case depends strongly on the relationship between the threshold values of the requirement in each period, which in turn depend on the level of initial equity of the bank. Finally, a policy recommendation discussed here is to combine a small capital requirement, in order to build a buffer against financial shocks, with a minimum equity requirement, which has the advantage of reducing risk-taking with a smaller welfare loss in terms of financial intermediation.

These results are of extreme interest in the current situation, where the crisis in the international financial markets has called for tighter regulation in the banking industry. The discussion in this chapter shows that an anticipated increase in capital requirements in the next period, combined with a shock reducing the expected return of the risky technology, increases the likelihood of a more aggressive risk-taking response by banks. In the light of these results, any amendment to the current regulatory framework should be carefully analysed.

The chapter is organised as follows. Section 4.2 sets up the basic three period's model for a regulated bank, and establishes an upper bound for the social optimum level of risk. Section 4.3 studies the equilibrium when capital requirements are slack, and compute threshold values for which they start binding. Section 4.4 studies the equilibrium when capital requirements bind in one or two periods. A discussion of the main results and policy implications are given in section 4.5, and conclusions and possible extensions are provided in section 4.6.

4.2 The Model

Consider a bank operating in an economy over three periods $t \in \{0, 1, 2\}$, with an exogenous initial equity of W_0. The bank manager is risk neutral and acts perfectly in the interest of shareholders (no agency problems), maximising the expected value of equity.

A safe asset is available in periods 0 and 1, which gross rate of return is normalised to one. That is, for each unit of consumption invested at t, this technology returns 1 unit at $t+1$. At $t = 0$ there is also a risky portfolio, which risk-return structure can be influenced by the bank, that with probability $p(R)$ returns R units at $t = 1$ per unit invested at $t = 0$, and zero otherwise. The probability function, $p(.)$, is strictly decreasing and concave for all $R \geq 1$, and satisfies $p(1) = 1$. The safe asset is (weakly) dominated by this technology if, in accordance with finance theory, there is a range of values (though eventually small) where a positive trade-off exists between risk and expected returns.[4] With the assumptions above, the expected return of the risky portfolio, $p(R)R$, is strictly concave for $R \geq 1$, and corner solutions with infinite risk are ruled-out. The unique level of risk that maximises this expected return function is given by

$$p'(R^*)R^* + p(R^*) = 0,$$

where $R^* > 1$ iff $p'(1) + p(1) > 0$.

For the sake of tractability, at $t = 1$ only one "risky" project is available, returning $\overline{R} > 1$ with probability 1. Even though it would be desirable to replicate period 0 structure of the risky asset in the intermediate period, the model would become analytically intractable. While this simplification eliminates the incentive for asset substitution in the intermediate period, it also allows concentrating on the risk choice of the bank in the initial period. Moreover, it is realistic to assume that at $t = 0$ the bank manager does not know the full spectrum of investment possibilities available to her in the following period, though she might have an idea of their average return, $\overline{R}$, which is what I assume here.

[4] Clearly, all projects with $R < 1$ are dominated by the safe asset.

Finally, assume the bank is able to raise fully insured deposits D_t in any amount at period t, at a cost $C(D_t) \geq D_t$ at $t+1$, a strictly increasing and convex function satisfying $C(0) = 0$ and $C'(0)$ bounded; which can be justified by incomplete competition arguments.[5] The assumption of full deposit insurance, in a model of complete information, makes the demand for deposits independent of the level of risk chosen by the bank. In other words, depositors are risk neutral. The assumption of universal risk neutrality is useful here to separate risk effects due to risk choice from those due to risk aversion of agents.

4.2.1 First Best

Given the assumption of universal risk neutrality, without loss of generality assume the utility function of the representative agent is $U(y) = y$. Assume consumption is postponed until $t = 2$, and in each period there is an endowment of M_t, which needs to be allocated between the risk free and risky technologies. Let me call $x_0 \leq M_0$, the amount invested in the risky asset at $t = 0$. If the project succeeds, which happens with probability $p(R)$, society will have a wealth of $(M_0 - x_0) + x_0 R$ at $t = 1$; while if the project failed, with probability $1 - p(R)$, the wealth of society will only be $M_0 - x_0$.

A new endowment of M_1 is realised at $t = 1$, which is fully invested in $\overline{R} > 1$ (because this technology is dominant in that period). Hence with probability $p(R)$, society will have a wealth of $[(M_0 - x_0) + x_0 R + M_1]\,\overline{R}$ at $t = 2$, and with probability $1 - p(R)$, a final wealth of $[M_0 - x_0 + M_1]\,\overline{R}$ for consumption.

A risk neutral social planner then maximises

$$\begin{aligned}&\max_{x_0,R} p(R)\,[(M_0 - x_0) + x_0 R + M_1]\,\overline{R} + (1 - p(R))\,[M_0 - x_0 + M_1]\,\overline{R}\\&\text{st. } x_0 \leq M_0\end{aligned}$$

or equivalently

$$\begin{aligned}&\max_{x_0,R} x_0 \overline{R}\,[p(R)R - 1] + [M_0 + M_1]\,\overline{R}\\&\text{st. } x_0 \leq M_0\end{aligned}$$

For the region where the risky technology is dominant $(p(R)R \geq 1)$ this function is increasing on x_0, then at the optimum $x_0 = M_0$. Hence, the first order condition for this

[5] Blum (1999) assumes horizontal differentiation. Another view could be that banks are ex-ante competitive, but due to high searching and switching costs they become ex-post monopolistic (see Acharya and Yorulmazer, 2006; Gondat-Larralde and Nier, 2006). The cost function $C(.)$ can be thought of as the gross interest rate paid on deposits in each period. Assuming it to be equal in both periods may be a strong simplification, given that in this model the risk borne by depositors is higher in the initial period.

problem is: $p'(R)R + p(R) = 0$, and as the objective function is concave, this condition is sufficient for a social optimum.

Of course, the previous exercise did not consider the social cost of bank failure, understood as the foregone value of intermediation, or the cost borne by the deposit insurance agency in case of failure. Therefore, absent any bankruptcy cost, the social return should equal private profits and the risk neutral social planner should choose the level of risk that maximises expected returns, i.e., R^*.

When considering bankruptcy costs, the social efficient level of risk would be lower than the private optimum.[6] However, given that R^* proposes a simpler framework for comparison, in the remainder of this chapter, when referring to the social optimum I will be talking about R^*, the zero bankruptcy cost, social efficient level of risk; keeping in mind this is in fact an upper bound for the true value.

4.2.2 Capital Requirements

Capital requirements limit the resources that can be invested in the risky technology – though any remaining funds can be invested in the safe asset without restrictions. For a regulated bank, a capital requirement c_t on its original formulation (the Cooke ratio) imposes that capital should be at least equal to an 8 percent of risk of weighted loans. Denoting by $I_t \leq W_t + D_t$, the investment in the risky portfolio in period t, capital requirements in this model translate into:[7]

$$\frac{W_t}{I_t} \geq c_t.$$

Clearly, for a given level of equity at any period, the more stringent the requirement (the higher the value of c_t) the lower the allowed investment in the risky portfolio.

The expected equity of the bank in each period is given by the return of the investment in the risky asset, plus the return of any remaining funds invested in the safe technology, minus the return on deposits; provided the bank has survived to that period (figure 4-1). Otherwise, and because of limited liability, all remaining resources are

[6] A more general way of stating this problem would be to consider bankruptcy costs as a convex and increasing function of R, $\phi(R)$. The conditional expected return of the risky portfolio would then be $p(R)R - (1-p(R))\phi(R)$.
The FOC would be given by: $p'(R)R+p(R)+p'(R)\phi(R)-(1-p(R))\phi'(R) = 0$ or $p'(R^*)R^*+p(R^*) = (1-p(R^*))\phi'(R^*) - p'(R^*)\phi(R^*) \geq 0$, which implies that because marginal returns are decreasing, by including bankruptcy costs, risk would be reduced.

[7] By convention, risk free assets have zero weight. The definition of capital requirements considered here only takes into account credit risk. It also assumes that the risky assets are weighted 100%, although Pillar I of Basel II includes weights as high as 350%. A justification for the assumption made here would be for the assets in the risky portfolio to be unrated (BIS, 2006).

transferred to the deposit insurance agency and the bank closes down (in other words, its equity equals 0).[8] Therefore, the bank's expected equity in each period is given by:

$$\begin{aligned}
&W_0 \\
&\mathbb{E}[W_1] &&= p(R)\left\{I_0 R + (W_0 + D_0 - I_0) - C(D_0)\right\} \\
& &&= p(R)\left\{I_0 (R-1) + W_0 - (C(D_0) - D_0)\right\} \\
&\mathbb{E}[W_2] &&= p(R)\left\{I_1 \overline{R} + (W_1 + D_1 - I_1) - C(D_1)\right\} \\
& &&= p(R)\left\{I_1 \left(\overline{R} - 1\right) + W_1 - (C(D_1) - D_1)\right\}
\end{aligned}$$

The regulated, risk neutral bank manager maximises the expected value of final equity, subject to capital constraints and standard feasibility conditions for investment in each period:[9]

$$(P) \quad \max_{R, D_0, I_0, W_1, D_1, I_1} \quad p(R)\left\{I_1\left(\overline{R} - 1\right) + W_1 + D_1 - C(D_1)\right\}$$

$$st.$$

$$\begin{array}{lll}
(1) & I_0 (R-1) + W_0 + D_0 - C(D_0) - W_1 \geq 0 & (\theta) \\
(2) & W_0 + D_0 - I_0 \geq 0 & (\lambda_0) \\
(3) & W_0 - c_0 I_0 \geq 0 & (\mu_0) \\
(4) & W_1 + D_1 - I_1 \geq 0 & (\lambda_1) \\
(5) & W_1 - c_1 I_1 \geq 0 & (\mu_1)
\end{array}$$

where θ is the shadow price of equity in $t = 1$, λ_t is the shadow value of the risky portfolio in period t, and μ_t is the shadow cost of the capital requirement in period t.

Sufficient conditions for an optimum are (see Appendix):

[8] If the bank fails at $t = 1$ (which happens with probability $1 - p(R)$), because of limited liability its equity is $\max\{(W_0 + D_0 - I_0) - C(D_0), 0\}$. However, it is always the case that $(W_0 - I_0) + (D_0 - C(D_0)) \leq 0$. The first term is negative because the risky technology is weakly dominant, therefore investment in the safe asset is effective only when capital requirements are binding, that is, if $I_0 = \frac{1}{c_0} W_0 > W_0$. The second term is also negative because, as I said before, by assumption $C(D_0) \geq D_0$.

[9] In order to rule out the possibility for a short sale of assets, non negativity constraints in all the variables should also be considered. Instead of explicitly including them, and in order to simplify the algebra, I will check they are satisfied at the optimum (see appendix).

$$\theta = \begin{cases} p(R^r)\overline{R} & \text{if } c_1 = 0 \\ \dfrac{p(R^r)}{c_1}\left\{\overline{R} + (c_1 - 1)\, C'(D_1^r)\right\} & \text{if } c_1 \neq 0 \end{cases} \tag{4.1}$$

$$\lambda_0 = \theta\left(C'(D_0^r) - 1\right) \tag{4.2}$$

$$\mu_0 c_0 = \theta\left(R^r - C'(D_0^r)\right) \tag{4.3}$$

$$\lambda_1 = p(R^r)\left(C'(D_1^r) - 1\right) \tag{4.4}$$

$$\mu_1 c_1 = p(R^r)\left(\overline{R} - C'(D_1^r)\right) \tag{4.5}$$

$$p'(R^r)\left\{I_1^r\left(\overline{R} - 1\right) + W_1^r + D_1^r - C(D_1^r)\right\} + \theta I_0^r = 0 \tag{4.6}$$

where "r" stands for the "*regulated*" solution. Notice from equations (4.3) and (4.5) that μ_t are well defined for $c_t = 0$, provided $W_t > 0,\ t = 1, 2$. Second order conditions are satisfied, as shown in the Appendix.

First notice that $\theta > 0$ for all R such that $p(R) \neq 0$. As I said before, θ can be interpreted as the shadow price of equity in period one, which is always valuable to the bank that has not gone bankrupt, and is equal to the earnings realised up to that period.

The non-negativity of all the multipliers implies that $R^r \geq C'(D_0^r) \geq 1$ and $\overline{R} \geq C'(D_1^r) \geq 1$. Observe that, because $R^r > 1$ and $\overline{R} > 1$, solutions involving $\lambda_0 = \mu_0 = 0$ or $\lambda_1 = \mu_1 = 0$ are not feasible. When money is invested in the safe asset in a determined period, either equation (2) or (4) in program (P) are slack, and the corresponding capital requirement (equation (3) or (5), respectively) should be binding. In such a case, investment in the safe asset takes place because the marginal cost of deposits equals the marginal return of the safe technology in that period. On the other hand, every time a capital requirement constraint is slack, the corresponding investment in the safe asset in that period is nil (because of weak dominance), as in that case the marginal cost of deposits would be strictly higher than the marginal return of the safe technology (equations (4.2) and (4.4)). Summing up, because the risky technology is weakly dominant, money is invested in the safe asset if and only if capital adequacy requirements in a determined period are binding.

4.3 Unregulated Solution

(Slack Capital Adequacy Requirements)

When capital requirements do not bind in either of the two periods, all funds are invested in the risky portfolio ($I_0^u = W_0 + D_0^u$ and $I_1^u = W_1^u + D_1^u$). By complementary slackness $\mu_0 = \mu_1 = 0$, $\lambda_1 > 0$, $\lambda_2 > 0$, and first order conditions become:

$$\theta = p(R^u)\overline{R} \tag{4.7}$$

$$\lambda_0 = \theta\,(R^u - 1) > 0 \tag{4.8}$$

$$C'(D_0^u) = R^u \tag{4.9}$$

$$\lambda_1 = p(R^u)\left(\overline{R} - 1\right) > 0 \tag{4.10}$$

$$C'(D_1^u) = \overline{R} \tag{4.11}$$

$$p'(R^u)R^u + p(R^u) = \frac{p'(R^u)}{\overline{R}(W_0 + D_0^u)}\left\{\overline{R}C(D_0^u) - \left(D_1^u\overline{R} - C(D_1^u)\right)\right\} \tag{4.12}$$

where "u" stands for the non binding case or "*unregulated*" solution. The second order condition relevant to this problem is (see Appendix):

$$p''(R^u)W_2^u + 2p'(R^u)\overline{R}\,(W_0 + D_0^u) + \frac{p(R^u)\overline{R}}{C''(D_0^u)} < 0 \tag{4.13}$$

I want to compare the risk choice of the unregulated bank with the social optimum. If $\overline{R}C(D_0^u) < D_1^u\overline{R} - C(D_1^u)$, the RHS of equation 4.12 would be positive, then $p'(R^u)R^u + p(R^u) > 0 = p'(R^*)R^* + p(R^*)$, and because marginal returns are decreasing this inequality would imply that the risk chosen by the unregulated bank would be below the efficient level ($R^u < R^*$). This is so because future income is so high that the bank would be willing to ration credit in the initial period in order to increase the probability of getting that income. In that case other policies, different from minimum capital requirements, would be needed.

Therefore, in the remainder of this chapter I will assume (as in Blum, 1999) that future income is bounded above by period 0 costs, $\overline{R}C(D_0^u) > D_1^u\overline{R} - C(D_1^u)$, which is equivalent to

$$p'(R^u)R^u + p(R^u) < 0 \text{ or } R^u > R^*.$$

In principle, one would expect the correlation between risk and equity to be negative, because the more capital the bank has, the more is at stake in the event of failure. The numerical example shown in figure 4-2 confirms this conjecture, as the optimal level of risk chosen by the unregulated bank decreases with the level of initial equity.[10] This

[10]Consider a quadratic form for the cost function, $C(x) = ax + bx^2$, and a linear probability function

result can be formally proved as follows.

Proposition 36 *The risk chosen by the unregulated bank is decreasing on initial equity, and converges on the social optimum as* $W_0 \to +\infty$.

Proof. For $\dfrac{dR^u}{dW_0} < 0$ see comparative statics in the Appendix (section 4.7.2).

Taking the limit as $W_0 \longrightarrow +\infty$ in equation 4.12, it is clear that the RHS goes to zero (R^u, D_0^u and D_1^u are bounded), and so the risk chosen by the unregulated bank converges on the social optimum. ∎

Proposition 37 *Banks with more equity depend relatively less on deposits to finance investment.*

Proof. It is proved in the Appendix that in this case $\dfrac{dD_0^u}{dW_0} < 0$, and as $C'(D_1^u) = \overline{R}$ then $\dfrac{dD_1^u}{dW_0} = 0$. ∎

Proposition 38 *There exist critical threshold values for capital requirements in each period, depending on the value of initial equity, for which regulatory constraints just start to bind. These are* $\widetilde{c_0} = \dfrac{W_0}{W_0 + D_o^u}$ *and* $\widetilde{c_1} = \dfrac{R^u (W_0 + D_0^u) - C(D_0^u)}{R^u (W_0 + D_0^u) - C(D_0^u) + D_1^u}$, *respectively.*

Proof. Period 0 requirements do not bind if

$W_0 > c_0 I_0^u = c_0 (W_0 + D_0^u) \Longleftrightarrow c_0 < \widetilde{c_0} = \dfrac{W_0}{W_0 + D_o^u}$.

Period 1 requirements do not bind if

$W_1^u > c_1 I_1^u = c_1 (W_1^u + D_1^u)$

$\Longleftrightarrow (W_0 + D_0^u) R^u - C(D_0^u) = c_1 ((W_0 + D_0^u) R^u - C(D_0^u) + D_1^u)$,

$\Longleftrightarrow c_1 < \widetilde{c_1} = \dfrac{R^u (W_0 + D_0^u) - C(D_0^u)}{R^u (W_0 + D_0^u) - C(D_0^u) + D_1^u}$. ∎

This result is consistent with the intuition that for well capitalised banks the application of small values of the capital requirement should be irrelevant; and indeed, with the empirical evidence that capital requirements are slack in the majority of banks in countries that have adopted the Basel principles. Nevertheless, sufficiently tight regulation will eventually force them into modifying their capital to asset ratios.

The impact of initial equity (W_0) over the effectiveness of capital requirements in period 0 is clear (figure 4-3).

$p(R) = \frac{U}{U-1} - \frac{1}{U-1}R$, with support $[1, U]$, satisfying the assumptions of this model.

$\int_1^U p(R)dR = 1 \Rightarrow U = 3$, $C(D_t) \geq D_t$ and $C'(D_t) \geq 1 \; \forall t \Rightarrow a = 1$, $C''(D_t) > 0 \Rightarrow b > 0$.

Proposition 39 *The initial period's threshold is increasing on the bank's equity.*

Proof. Using equations 4.9 , 4.11 and 4.12, total differentiation of $\widetilde{c_0}$ leads to

$$\frac{d\widetilde{c_0}}{dW_0} = \frac{D_0^u}{(W_0 + D_0^u)^2}\left(1 - \frac{W_0}{D_0^u C''(D_0^u)}\frac{dR^u}{dW_0}\right) > 0. \blacksquare$$

Hence, as risk decreases with initial equity $\left(\frac{dR^u}{dW_0} < 0\right)$, the threshold for which the initial period capital requirement starts to bind raises because, with a higher equity, the bank is at the same time more solvent and less risky.

On the other hand, total differentiation of $\widetilde{c_1}$ leads to

$$\frac{d\widetilde{c_1}}{dW_0} = \frac{D_1^u}{(W_1^u + D_1^u)^2}\left(R^u + (W_0 + D_0^u)\frac{dR^u}{dW_0}\right) = \frac{D_1^u R^u}{\widetilde{c_0}(W_1^u + D_1^u)^2}\left(\widetilde{c_0} + \frac{W_0}{R^u}\frac{dR^u}{dW_0}\right)$$

Hence,

$$\frac{d\widetilde{c_1}}{dW_0} = \begin{cases} > 0 & \text{if } \left|\dfrac{W_0}{R^u}\dfrac{dR^u}{dW_0}\right| < \widetilde{c_0} \\ \leq 0 & \text{if } \left|\dfrac{W_0}{R^u}\dfrac{dR^u}{dW_0}\right| \geq \widetilde{c_0} \end{cases}$$

and the evolution of this threshold depends upon a sort of "income elasticity" of the demand for risk.

At least for the chosen parameters of the numerical example shown here, $\widetilde{c_1}$ appears to depend increasingly on the level of initial equity (figure 4-3). However, this does not constitute a proof and, in principle, $\widetilde{c_1}$ could be non-monotonic in W_0.

Proposition 40 *Both thresholds converge on 1 as* $W_0 \longrightarrow +\infty$.

Proof. Direct from their definition (in proposition 38), given that R^u, D_0^u and D_1^u are bounded. $\blacksquare$

Finally, given the definitions of $\widetilde{c_0}$ and $\widetilde{c_1}$ and the assumption on R^u,

$$\widetilde{c_1} - \widetilde{c_0} > \frac{1}{I_0^u I_1^u}\left\{\underbrace{[D_0^u R^u - C(D_0^u)]I_0^u}_{\geq 0} - \underbrace{\frac{W_0}{\overline{R}}C(D_1^u)}_{\geq 0}\right\}.$$

Notice D_1^u is independent of W_0, hence, at least for small values of W_0, $\widetilde{c_1} > \widetilde{c_0}$. In particular, for $W_0 = 0$, $\widetilde{c_0} = 0$ and $\widetilde{c_1} = \dfrac{C(D_0^u)}{C(D_0^u) - D_1^u} > 0$, and though capital requirements bind for arbitrarily small values in period 0, a much tighter regulation would be needed in period 1 for it to bind. A general relationship between $\widetilde{c_0}$ and $\widetilde{c_1}$ for different values of W_0, however, cannot be established at this point.

4.4 Regulated Solution

The main results in Blum's paper establish that tightening the requirement in the initial period leads to less risk-taking, while increasing the requirement in future periods raises the level of risk above that chosen by the unregulated bank.

Blum (1999) main results: *(i) If a bank faces a binding requirement in the initial period, an increase in the requirement reduces the level of risk.*
(ii) When the capital requirement in the intermediate period first becomes binding, tightening the requirement raises the level of risk. If the requirement is further increased, risk eventually falls again but never below the level of an unregulated bank.

The first part of the result is intuitive, because by rising c_0 the return per unit invested in the risky portfolio is reduced as well (each unit invested at $t = 0$ returns $\dfrac{1}{c_0}R$ units at $t = 1$). These results, however, do not say much about when the requirement is active, how is the regulated equilibrium compared to the unregulated one or to the social optimum, the effects of regulation over financial intermediation, or its impact on the distribution of risk among banks. These questions will be addressed in the following sections.

4.4.1 Initial Period Binding Requirement

A numerical example for the case of a binding requirement in period 0 only, considering the same functions as before, is shown in figure 4-4.[11] The discontinuous line shows the social optimum. Up to $\widetilde{c}_0$ (which depends on the value of initial equity), the capital requirement is slack and the bank chooses the unregulated level of risk. After that, risk is reduced until at some point the social optimum is achieved. The following proposition formally establishes this result.

Proposition 41 *If capital requirements bind only in the initial period, tightening the requirement always reduces risk below the unregulated solution and, indeed, for each positive value of the initial equity there exists a unique value of the requirement, c_0^*, for which the bank chooses the social optimum.*

Proof. Comparative statics in the Appendix show that $\dfrac{dR}{dc_0} \leq 0$ in this case.

When the requirement just starts binding ($c_0 = \widetilde{c}_0$), $R^r = R^u > R^*$. Conversely, for the tightest possible regulation ($c_0 = 1$) equation 4.6 can be re-written:[12]

[11] Different forms of the probability distribution, which keep the assumptions of the model, give similar qualitative results.

[12] When $c_0 = 1$, if $\lambda_0 > 0 : D_0 = \left(\frac{1-c_0}{c_0}\right) W_0 = 0 \Rightarrow C(D_0) = D_0$. Else, if $\lambda_0 = 0 : C'(D_0) = 1 \Rightarrow C(D_0) = D_0$.

$(p(R)+p'(R)R)\,W_0\overline{R}=-p'(R)\left\{D_1\overline{R}-C(D_1)\right\}>0.$

Therefore $p(R)+p'(R)R>0=p(R^*)+p'(R^*)R^*$ and $R^r<R^*$. Continuity and strict monotonicity implies that there exists a unique value of c_0 for which $R^r=R^*$. ■

Remark 42 *Notice $\widetilde{c_0}=0$ for $W_0=0$ and, as regulation always binds, the bank cannot invest in the risky portfolio ($W_0=0\geq c_0I_0$, $\forall c_0>0 \Longrightarrow I_0=0$), so $R=1<R^*$. Therefore, c_0^* does not exist for $W_0=0$.*

Additionally, as the requirement binds only for values above $\widetilde{c_0}$, $c_0^*\geq\widetilde{c_0}$. Hence, by proposition 40, $c_0^*\to 1$ as $W_0\to+\infty$ (figure 4-5).

Proposition 43 *The higher the initial equity of the bank, the tighter the regulation required to make its risk converge on the social optimum.*

Proof. When the requirement binds only in the initial period, equation 4.6 can be re-written as

$(p'(R)R+p(R))\,\frac{W_0}{c_0}\overline{R}=-p'(R)\left\{\left[D_0-C(D_0)-\frac{1-c_0}{c_0}W_0+D_1\right]\overline{R}-C(D_1)\right\}.$

Therefore, $R=R^*$ if and only if $\left[D_0-C(D_0)-\frac{1-c_0^*}{c_0^*}W_0+D_1\right]\overline{R}-C(D_1)=0.$

Implicit differentiation of the equation above leads to

$\frac{dD_0}{dW_0}\left(1-C'(D_0)\right)+\frac{W_0}{c_0^2}\frac{dc_0^*}{dW_0}-\frac{1-c_0^*}{c_0^*}=0$

If $\lambda_0=0\Rightarrow C'(D_0)=1$ and $\frac{dc_0^*}{dW_0}=\frac{c_0^*(1-c_0^*)}{W_0}\geq 0.$

If $\lambda_0>0\Rightarrow I_0=W_0+D_0\Rightarrow D_0=\left(\frac{1-c_0}{c_0}\right)W_0$, therefore $\frac{dD_0}{dW_0}=-\frac{W_0}{c_0^2}\frac{dc_0}{dW_0}+\frac{1-c_0}{c_0}$, and the expression above becomes $C'(D_0)\left(\frac{W_0}{c_0^2}\frac{dc_0^*}{dW_0}-\frac{1-c_0^*}{c_0^*}\right)=0\Longleftrightarrow\frac{dc_0^*}{dW_0}=\frac{c_0^*(1-c_0^*)}{W_0}\geq 0.$ ■

Proposition 44 *When capital requirements bind only in the initial period, the tighter the requirement the less deposits are taken from the public.*

Proof. Comparative statics in the Appendix show that $\frac{dD_0}{dc_0}\leq 0$. Also, as D_1 is constant, $\frac{dD_1}{dc_0}=0$. ■

Proposition 45 *Among banks for which the capital requirement binds only in the initial period, the better capitalised banks raise more deposits and take on more risk.*

Proof. Comparative statics in the Appendix show that $\frac{dD_0}{dW_0}\geq 0$ and $\frac{dR}{dW_0}\geq 0$ in this case. ■

Therefore, the consequences of a first period binding capital requirement on the equilibrium risk-taking of banks are mixed. While reducing risk below the unregulated solution – even to the social optimum, for sufficiently tight regulation – fewer deposits are taken, which reduces financial intermediation. Capital regulation is usually considered as a natural counterpart to deposit insurance, attempting to control for moral hazard in an industry that is implicitly receiving a subsidy from the government. However, proposition 45 indicates that capital requirements are not sufficient to control moral hazard because, among the binding banks, the better capitalised ones raise relatively more insured deposits and take on relatively more risk (figure 4-6).

4.4.2 Intermediate Period Binding Requirement

Figure 4-7 shows the dynamic of D_1 and R for the chosen parameters of the numerical example presented in this chapter, when capital requirements bind only in the interim period. In this case, comparative statics do not give conclusive results. However, when capital regulation just becomes binding tightening the requirement will increase bank risk-taking, most likely above the unregulated solution for all values of c_1. In such a case, risk would be decreasing on equity $\left(\frac{dR}{dW_0} \leq 0\right)$, making risk-taking even more aggressive for poorly capitalised banks (figure 4-8). Therefore, capital requirements will not only worsen the risk choice of banks but also make the smaller banks weaker.

The effects on financial intermediation are not clear because intermediate period deposits fall with the requirement $\left(\frac{dD_1}{dc_1} \leq 0\right)$ but, provided $\frac{dR}{dc_1} \geq 0$, first period deposits increase $\left(\frac{dD_0}{dc_1} = \frac{1}{C'(D_0)} \frac{dR}{dc_1} \geq 0\right)$ in order to finance the risky portfolio.

4.4.3 Binding Requirements in Both Periods

The previous results are useful to identify and separate the effects of binding regulation in different periods. However, as regulators tend to apply constant capital requirements it might happen that, for some values, they bind in both periods. Depending on the relationship between the thresholds, one of the following two situations is possible: either $\widetilde{c_0} < \widetilde{c_1}$ or $\widetilde{c_0} \geq \widetilde{c_1}$.[13]

Figure 4-9 depicts both cases. The blank areas show the regions of values of c_0 and c_1 where capital requirements are slack. The horizontally dashed areas show the regions for which only period 0 capital requirements bind. The vertical dashed areas depict the regions for which only period 1 regulation binds. Finally, the diagonal dashed areas show the values for which capital requirements bind in both periods. The solid black line represents constant capital requirements ($c_0 = c_1 = c$), which are in fact the framework regulators applied under Basel I. For values of the requirement above this line, period 1

[13] Although figure 4-3 shows $\widetilde{c_0} < \widetilde{c_1}$ for all values of W_0, this might be highly dependent on the functions chosen for that example. In principle, nothing precludes $\widetilde{c_0} > \widetilde{c_1}$ for some (high) values of W_0.

effects would be stronger, while a decreasing risk effect would be more likely for values below the identity line.

Proposition 46 *When a constant capital requirement binds in both periods, the tighter the regulation the fewer deposits are taken from the public in both periods, though better capitalised banks take on relatively more deposits.*

Proof. Comparative statics in the Appendix show that $\frac{dD_t}{dc} \leq 0$ and $\frac{dD_t}{dW_0} \geq 0$, for $t = 1, 2$. ■

Proposition 47 *With a constant binding capital requirement, R^r equals R^* when c equals* 1.

Proof. With $c = 1$, as capital requirements bind in both periods $W_0 = I_0$ and $W_1 = I_1$. Equation 4.6 becomes:

$$[p'(R)R + p(R)]\, W_0\overline{R} = p'(R)\left\{\left(C\left(D_0\right) - D_0\right)\overline{R} + C\left(D_1\right) - D_1\right\} = 0,$$

because

$$\left.\begin{array}{lll} \lambda_t > 0 & \Rightarrow & D_t = \left(\frac{1-c}{c}\right) W_t = 0 \\ \lambda_t = 0 & \Rightarrow & C'\left(D_t\right) = 1 \Rightarrow D_t = 0 \end{array}\right\} \Rightarrow C\left(D_t\right) = D_t \quad t = 1, 2$$

Therefore $R^r\left(c = 1\right) = R^*$.

Indeed, the regulated solution approaches the optimum "from above", i.e., $R^r \geq R^*$ in the neighbourhood of $c = 1$, because $\left.\dfrac{dR}{dc}\right|_{c=1} < 0$ (see comparative statics in the Appendix). ■

Proposition 48 *If $\widetilde{c_0} < \widetilde{c_1}$, $R^r = R^u$ for all $c < \widetilde{c_0}$, R decreases in c for all $\widetilde{c_0} \leq c < \widetilde{c_1}$ and increases right afterwards. Moreover, there exists $c^* \leq 1$ at which $R^r = R^*$.*

Proof. This is situation (a) in figure 4-9. When $c < \widetilde{c_0}$ none of the requirements bind and so $R^r = R^u$. Afterwards, only period 0 requirements bind and in that case it has already been proved that $\frac{dR}{dc} \leq 0$.

When c reaches $\widetilde{c_1}$, period 1 requirements start to bind, which in terms of the multipliers means $\mu_1 = 0$ and $\lambda_1 > 0$. So $C'(D_1) = \overline{R}$ and $D_1 = \left(\frac{1-c}{c}\right) W_1$, where

$$W_1 = W_0\left(\frac{R+c-1}{c}\right) + D_0 - C(D_0).$$

Implicit differentiation of these expressions gives:

$$\frac{dD_1}{dc} = -\frac{W_1}{c^2} - \left(\frac{1-c}{c}\right)\left(R - C'(D_0)\right)\frac{W_0}{c^2} + \left(\frac{1-c}{c^2}\right) W_0\frac{dR}{dc} = 0 \Longrightarrow \frac{dR}{dc} \geq 0.$$

Therefore, at $\widetilde{c_1}$ the risk chosen by the unregulated bank starts to increase.

Comparative statics do not give a clear sign for $\frac{dR}{dc}$ (see Appendix), but from proposition 47 we know that R^r approaches the optimum "from above". So, this case presents two possibilities: either $c_0^* < \widetilde{c_1}$ or $\widetilde{c_1} \leq c_0^*$. In the first case, when period 1 requirement starts binding $R^r(\widetilde{c_1}) < R^*$ and afterwards R increases above R^*, then decreases and reaches R^* once again when $c = 1$. Otherwise, $R^r(\widetilde{c_1}) \geq R^*$ which means R equals R^* only when $c = 1$ (figure 4-10). ■

Proposition 49 *If $\widetilde{c_0} > \widetilde{c_1}$, $R^r = R^u$ for all $c < \widetilde{c_1}$, R increases right after $\widetilde{c_1}$ and decreases right after $\widetilde{c_0}$. Moreover, only at $c = 1$, $R = R^*$.*

Proof. This is situation (b) in figure 4-9. When $c < \widetilde{c_1}$ none of the requirements bind and so $R^r = R^u$. Afterwards, only period 1 requirements bind and in that case it has already been proved that in the neighbourhood of $\widetilde{c_1}$, $\frac{dR}{dc} > 0$.
When c reaches $\widetilde{c_0}$, period 0 requirements start to bind, which in terms of the multipliers means $\mu_0 = 0$ and $\lambda_0 > 0$. So $C'(D_0) = R$ and $D_0 = \left(\frac{1-c}{c}\right) W_0$.
Implicit differentiation of these expressions gives:

$$\frac{dD_0}{dc} = \frac{1}{C''(D_0)} \frac{dR}{dc} = -\frac{1}{c^2} W_0 = 0 \Longrightarrow \frac{dR}{dc} < 0.$$

Comparative statics do not give a clear sign for $\frac{dR}{dc}$ (see Appendix), but from proposition 47 we know that R^r approaches the optimum "from above" (see figure 4-11). ■

Therefore, the dynamic of risk depends strongly on the relationship between the threshold values of the requirement in each period, which in turn depend on the level of initial equity of the bank.

4.5 Policy Implications

The results contained in the previous section indicate that where $\widetilde{c_0} < \widetilde{c_1}$, capital requirements can reduce risk-taking below the unregulated solution and even achieve the social optimum for a sufficiently tight regulation. However, the effects of capital requirements vary among banks with different values of equity. So, how much should the regulator want to reduce risk-taking? Or, in other words, given constant capital requirements, which type of bank should the regulator aim for?

Consider a system with three banks $\{S, M, B\}$ with increasing levels of equity $\left(W^S < W^M < W^B\right)$ and a constant capital requirement c in both periods, imposed on all banks. Assume $\widetilde{c_0}\left(W^j\right) < \widetilde{c_1}\left(W^j\right) \forall j \in \{S, M, B\}$, and $c^*\left(W^S\right) < \widetilde{c_1}\left(W^S\right)$. Section 4.3 showed that, in an unregulated banking industry, smaller banks take on more risk $\left(\frac{dR^u}{dW_0} < 0\right)$. So, if the regulator wanted to take the smallest bank to the social optimum, a solution like the one depicted in figure 4-12 could occur, where the capital requirement binds in the first period for the three banks. While risk-taking is overall reduced, the distribution of risks among banks has been modified towards a scenario where the bigger banks (both in term of equity and deposits) take on more risk (proposition 45).[14] Recalling that bigger banks are of systemic importance, because their failure might either bring the system down or trigger expensive rescue packages, the new scenario is more dangerous in terms of financial stability.

Then, one could think of a situation where the regulator wanted to take the biggest bank to the social optimum.[15] Assuming again $\widetilde{c_0}\left(W^j\right) < \widetilde{c_1}\left(W^j\right)\ \forall j \in \{S, M, B\}$, a situation like the one presented in figure 4-13 could happen, where the smaller banks invest sub-optimally.

Another option could be to reduce risk-taking towards the unregulated level of bank B. However, this would need different capital requirements for S and M, and counter-intuitively, the higher the equity the tighter the requirement: $c^S < c^M$ (figure 4-14). In this last situation one could think a better idea would be to introduce a minimum equity requirement equal to W^B. In this model, a minimum equity requirement would also reduce financial intermediation, but less than a binding capital requirement. This is shown in figure 4-15, where the region of deposits which are not taken by the banks (the shaded areas) is smaller in case (b), with a minimum equity requirement.[16] However, this would reduce competitiveness even more because it would introduce a new and high entry barrier in the market.[17]

In practice, capital requirements hardly ever bind and, when they do, they are regarded as regulatory thresholds for prudential intervention (capitalisation plans, administration take-over, or asset liquidation; depending on the severity of the problem).[18] Then, a possibility would be to combine low levels of capital requirements with a low minimum equity requirement (Rochet, 1992) set, for example, such that $R^u\left(W^e\right) = R^e$, where $p\left(R^e\right)$ is a tolerance level for the probability of failure, defined by the social planner. Notice that in order to achieve the same level of risk (R^e) through capital regulation, an increasing requirement in equity would be needed.

But if capital requirements are designed for them not to bind, then they cannot curb risk-taking and so, why introduce them? The current financial crisis (the so-called subprime crisis, because it originated in that segment of the mortgage credit market in the USA) enables us to appreciate how useful capital requirements can be for banks to

[14] A worse situation could be imagined, where $\widetilde{c_0}\left(W^B\right) > \widetilde{c_1}\left(W^B\right)$ and the risk choice of the largest bank could increase above the unregulated solution.

[15] Notice that, in this model, the social optimum is independent of the bank's level of equity.

[16] $\widetilde{W_0}\left(c\right)$ is defined as the inverse function of $\widetilde{c_0}\left(W_0\right)$, that is, the maximum value of the initial equity for which a constant capital requirement c binds. Given that $\widetilde{c_0}\left(W_0\right)$ is increasing in W_0, $\widetilde{W_0}\left(c\right)$ exists and it is increasing in c.

[17] Remember I am assuming banks behave as a local monopoly.

[18] See Bhattacharya et al. (2002), and Decamps et al. (2004).

build buffers against financial shocks, contributing to a more resilient banking system.

As a consequence of this crisis, many agents in the market have started to call for increased regulation. If that were translated into tighter capital requirements ($c_1 > c_0$ in this model), those banks experiencing the higher losses would have lower values of equity in the next period, making it more likely for capital regulation to bind for them at that stage. According to section 4.4.2, the effect would be the opposite from what was desired, so risk-taking would be more aggressive in those banks in the current period. In the light of these results, any modification of the current regulatory framework should be studied carefully.

This model has assumed initial equity to be public information. However, if the model were to be replicated infinite times, the initial equity would be the result of past periods' decisions, and so would be private information. In this sense, effective monitoring and improved private disclosure (both elements of Basel II) might help in revealing W_t. Think of this case in order to analyse the effects of macro-financial shocks on the risk choice of a regulated banking system with constant capital requirements. If a recessionary shock were anticipated, we would be in a situation similar to the one described above. The expected value of W_{t+1} would be smaller, making more likely for regulation to bind at $t+1$, therefore increasing risk-taking at t. In such a case, the appropriate policy would depend on the social planner's objective under that scenario, because higher risk-taking might lead to a credit expansion, acting to counter the cycle. On the other hand, if the shock were not anticipated, at t the bank would realise that its initial equity is smaller than expected, making it more likely for regulation to bind in the current period and so reducing risk-taking. However, depending on the severity of the shock, risk could be reduced to suboptimal levels, exacerbating the downturn.

4.6 Concluding Remarks

In this chapter I have explored some features of Blum's (1999) model neglected in his original work. I developed threshold values for which capital regulation becomes binding in each period and studied how the regulated equilibrium is compared to the unregulated solution and to the social optimum. I also studied the effects of regulation over financial intermediation and its impact on the distribution of risk among banks.

In particular, I have shown that for unregulated banks risk-taking is decreasing on the level of initial equity, and converges on the social optimum when equity is sufficiently high. When introducing capital requirements, I show that there exist critical threshold values in each period for which regulation starts binding. When the requirement binds in the initial period only, risk can be reduced below the unregulated solution – even to the social optimum for sufficiently tight regulation – but fewer deposits are taken, which reduces financial intermediation. Moreover, capital requirements are not sufficient to control moral hazard because, among the binding banks, the better capitalised raise relatively more insured deposits and take on relatively more risk. When the requirement binds only in the interim period, bank risk-taking increases, most likely above the unregulated solution for all values of the requirement. In that case, risk would be decreasing on equity, making risk-taking even more aggressive for poorly capitalised banks. Therefore, interim period binding capital requirements will not only worsen the risk choice of banks, but make smaller banks weaker. When a constant capital requirement binds in both periods, the tighter the regulation the fewer deposits are taken from the public, though better capitalised banks raise relatively more deposits. The dynamics of risk in this case depend strongly on the relationship between the threshold values of the requirement in each period, which in turn depend on the level of initial equity of the bank. Finally, a policy recommendation discussed here is to combine a small capital requirement, in order to build a buffer against financial shocks, with a minimum equity requirement, which has the advantage of reducing risk-taking with a smaller welfare loss in terms of financial intermediation.

These results are of extreme interest in the current situation, where the crisis in the international financial markets has called for tighter regulation in the banking industry. The discussion in this chapter shows that an anticipated increase in capital requirements in the next period, combined with a shock reducing the expected return of the risky technology, increases the likelihood of a more aggressive risk-taking response by the bank. In the light of these results, any amendment to the current regulatory framework should be carefully analysed.

This model has the advantage of enabling the study of asset substitution in a simple way, while the assumption of universal risk neutrality isolates the risk choice of the bank from any other decision in the game. However, all of Blum's original disclaimers still apply. This is not a general model, because the results rely strongly on the assumptions made on the probability function and the cost of rising deposits; and the assumption of universal risk neutrality, while useful, introduces an atmosphere of "too much risk-taking". Also, assuming $\overline{R}$ fixed in the intermediate period may be restrictive, although one could think of it as the net present value of all future profitable investment decisions, as estimated by the bank manager at each time period. While all these simplifications facilitated the development of the main conclusions, they also imply that they are incomplete. Nevertheless, the analysis is suggestive and it does help to understand the effects and limitations of capital requirements in an intertemporal framework.

A remark concerning pro-cyclicality and Basel II. Let me explain briefly how risk related weights modify the solution presented before. Consider a weight factor $\delta > 0$, applied to the risky portfolio I_t. This factor changes the computation of the capital to investment ratio, and therefore the effective value of the requirement faced by the regulated bank, in the following way:

$$\frac{W_t}{\delta I_t} \geq c_t \Leftrightarrow \frac{W_t}{I_t} \geq \delta c_t$$

If $\delta < 1$, the effective requirement is smaller, so less likely to be binding. The opposite is true for $\delta > 1$.

Pillar I of Basel II increases the risk sensitivity of capital requirements. Therefore, in a downturn, when risks are more likely to materialise, capital requirements would be higher. However, as discussed in section 4.5, the net effect is not necessarily pro-cyclical and will depend on whether the shock is anticipated or not, its size and the level of initial equity of the bank. Therefore, policy recommendations are not obvious, and should be decided on a case by case basis.

The regulator's commitment to monitor the bank is an issue that has not been addressed in this chapter. The computation of the threshold levels of capital requirements have assumed the bank is willing to comply with the requirement, which only happens when there is an effective threat of punishment for not doing so. This could be the subject of future research.

Finally, an interesting extension would be to explore capital requirements in a model à la Goldstein and Pauzner (2000), keeping the assumption of universal risk neutrality and introducing a second risky technology with the same mean but a higher variance, for example, in order to study risk shifting by the manager, and the optimal response functions of depositors and the regulator.

4.7 Appendix

4.7.1 The Regulated Bank

$$(P) \quad \max_{R,D_0,I_0,W_1,D_1,I_1} \quad p(R)\left\{I_1\left(\overline{R}-1\right)+W_1+D_1-C(D_1)\right\}$$

st.

$$\begin{array}{lll}
(1) & I_0(R-1)+W_0+D_0-C(D_0)-W_1 \geq 0 & (\theta) \\
(2) & W_0+D_0-I_0 \geq 0 & (\lambda_0) \\
(3) & W_0-c_0 I_0 \geq 0 & (\mu_0) \\
(4) & W_1+D_1-I_1 \geq 0 & (\lambda_1) \\
(5) & W_1-c_1 I_1 \geq 0 & (\mu_1)
\end{array}$$

The Lagrangian for this problem is:

$$\mathcal{L}=p(R)\left\{I_1\left(\overline{R}-1\right)+W_1+D_1-C(D_1)\right\}$$
$$+\theta\left\{I_0(R-1)+W_0+D_0-C(D_0)-W_1\right\}$$
$$+\lambda_0(W_0+D_0-I_0)+\mu_0(W_0-c_0 I_0)+\lambda_1(W_1+D_1-I_1)+\mu_1(W_1-c_1 I_1)$$

$$FOC(R): p'(R)\left\{I_1\left(\overline{R}-1\right)+W_1+D_1-C(D_1)\right\}+\theta I_0=0$$
$$FOC(D_0): \theta\left(1-C'(D_0)\right)+\lambda_0=0$$
$$FOC(D_1): p(R)\left(1-C'(D_1)\right)+\lambda_1=0$$
$$FOC(I_0): \theta(R-1)-\lambda_0-c_0\mu_0=0$$
$$FOC(I_1): p(R)\left(\overline{R}-1\right)-\lambda_1-c_1\mu_1=0$$
$$FOC(W_1): p(R)-\theta+\lambda_1+\mu_1=0$$

Solving simultaneously

$$\lambda_0=\theta\left(C'(D_0)-1\right)$$
$$\lambda_1=p(R)\left(C'(D_1)-1\right)$$
$$\theta(R-1)-\theta\left(C'(D_0)-1\right)-c_0\mu_0=0 \Rightarrow \theta\left(R-C'(D_0)\right)=c_0\mu_0$$

If $c_0=0$, equation (3) of program (P) becomes $W_0>0$, so by complementary slackness $\mu_0=0$.

$$p(R)\left(\overline{R}-1\right)-p(R)\left(C'(D_1)-1\right)-c_1\mu_1=0 \Rightarrow p(R)\left(\overline{R}-C'(D_1)\right)=c_1\mu_1$$

If $c_1=0$, equation (3) of program (P) becomes $W_1>0$, so by complementary slackness $\mu_1=0$. Provided $p(R)>0$, $\overline{R}=C'(D_1)$.

$\theta=p(R)+\lambda_1+\mu_1=\frac{p(R)}{c_1}\left\{\overline{R}+(c_1-1)\,C'(D_1)\right\}=p(R)\left\{C'(D_1)+\frac{\overline{R}-C'(D_1)}{c_1}\right\}$, that converges on $p(R)\overline{R}$ as $c_1 \to 0$.

While I have not ruled out short sale of assets explicitly in the constraints, the non-negativity of the multipliers imply $D_0, D_1 \geq 0$. By definition of $p(.)$, $R \geq 1$, and $\{I_t\}_{t=1,2}$ equal either $\frac{W_t}{c_t}$ or W_t+D_t, both being non-negative numbers.

Given that solutions where $\mu_t = \lambda_t = 0$ are ruled out, there are only nine possible combinations for the sign of the multipliers:

	μ_0	λ_0	μ_1	λ_1
1)	$=0$	>0	$=0$	>0
2)	>0	>0	$=0$	>0
3)	>0	$=0$	$=0$	>0
4)	$=0$	>0	>0	>0
5)	$=0$	>0	>0	$=0$
6)	>0	>0	>0	>0
7)	>0	>0	>0	$=0$
8)	>0	$=0$	>0	>0
9)	>0	$=0$	>0	$=0$

a) Unregulated case (slack capital requirements): $\mu_0 = \mu_1 = 0$, and $\lambda_0 > 0, \lambda_1 > 0$ (case 1).

b) Binding requirement at $t = 0$ only: $\mu_1 = 0, \lambda_1 > 0$ (cases 2 and 3).

c) Binding requirement at $t = 1$ only: $\mu_0 = 0, \lambda_0 > 0$ (cases 4 and 5).

d) Both periods binding requirements (cases 6 to 9).

a) Unregulated case:

$W_0 > c_0 I_0 \Rightarrow \mu_0 = 0 \Leftrightarrow \theta\left(R - C'(D_0)\right) \Rightarrow C'(D_0) = R$

$W_1 > c_1 I_1 \Rightarrow \mu_1 = 0 \Rightarrow C'(D_1) = \overline{R} = cnst$

$\lambda_0 > 0 \Rightarrow I_0 = W_0 + D_0$

$\lambda_1 > 0 \Rightarrow I_1 = W_1 + D_1$

$\theta = p(R)\overline{R}$

$W_1 = (W_0 + D_0)\, R - C(D_0)$

$p'(R)\left\{\left[(W_0 + D_0)\, R - C(D_0) + D_1\right] \overline{R} - C(D_1)\right\} + p(R)\overline{R}\,(W_0 + D_0) = 0$

Hence, R and D_0 are determined simultaneously from

$\mathcal{L}_R : p'(R)\left\{\left[(W_0 + D_0)\, R - C(D_0) + D_1\right] \overline{R} - C(D_1)\right\} + p(R)\overline{R}\,(W_0 + D_0) = 0$

$\mathcal{L}_{D_0} : p(R)\overline{R}\left[R - C'(D_0)\right] = 0$

The second order condition reduces to prove that the following matrix of second derivatives for R and D_0 is negative semi-definite, this is, that its leading principal minors alternate in sign:

$$L_a^2 = \begin{bmatrix} p''(R)W_2^a + 2p'(R)(W_0 + D_0)\overline{R} & p(R)\overline{R} \\ p(R)\overline{R} & -p(R)\overline{R}C''(D_0) \end{bmatrix}$$

where $W_2^a = [(W_0 + D_0)\,R - C(D_0) + D_1]\,\overline{R} - C(D_1) \geq 0$.

$$\det\left[L^2\right]_1 = p''(R)W_2^a + 2p'(R)(W_0 + D_0)\overline{R} < 0$$
$$\det\left[L^2\right]_2 = -p(R)\overline{R}C''(D_0)\left(p''(R)W_2^a + 2p'(R)(W_0 + D_0)\overline{R}\right) - \left(p(R)\overline{R}\right)^2$$
$$= -p(R)\overline{R}\left\{C''(D_0)\left[p''(R)W_2^a + 2p'(R)(W_0 + D_0)\overline{R}\right] + p(R)\overline{R}\right\} \geq 0$$
$$\Leftrightarrow C''(D_0)\left[p''(R)W_2^a + 2p'(R)(W_0 + D_0)\overline{R}\right] + p(R)\overline{R} \leq 0.$$

b) Binding requirement at $t = 0$ only

$\mu_0 > 0 \Rightarrow W_0 = c_0 I_0 \Rightarrow I_0 = \frac{W_0}{c_0}$
$W_1 > c_1 I_1 \Rightarrow \mu_1 = 0 \Rightarrow C'(D_1) = \overline{R} = cnst$
$\lambda_1 > 0 \Rightarrow I_1 = W_1 + D_1$
$\theta = p(R)\overline{R}$
$W_1 = \frac{W_0}{c_0}\,(R - 1 + c_0) + D_0 - C(D_0)$
If $\lambda_0 > 0: I_0 = W_0 + D_0 \Rightarrow D_0 = \left(\frac{1-c_0}{c_0}\right) W_0 = cnst$
If $\lambda_0 = 0: C'(D_0) = 1 \Rightarrow D_0 = cnst$

Hence, R is determined alone from
$\mathcal{L}_R$: $p'(R)\left\{\left(\frac{W_0}{c_0}\,(R - 1 + c_0) + D_0 - C(D_0) + D_1\right)\overline{R} - C(D_1)\right\} + p(R)\overline{R}\frac{W_0}{c_0} = 0$,
and the relevant second order condition is
$\mathcal{L}_{RR}$: $p''(R)\left\{\left(\frac{W_0}{c_0}\,(R - 1 + c_0) + D_0 - C(D_0) + D_1\right)\overline{R} - C(D_1)\right\} + 2p'(R)\overline{R}\frac{W_0}{c_0} < 0$.

c) Binding requirement at $t = 1$ only

$\mu_1 > 0 \Rightarrow W_1 = c_1 I_1 \Rightarrow I_1 = \frac{W_1}{c_1}$
$W_0 > c_0 I_0 \Rightarrow \mu_0 = 0 \Leftrightarrow \theta\,[R - C'(D_0)] = 0 \Rightarrow C'(D_0) = R$
$\lambda_0 > 0 \Rightarrow I_0 = W_0 + D_0$
$\theta = \dfrac{p(R)}{c_1}\left\{\overline{R} + (c_1 - 1)\,C'(D_1)\right\}$
$W_1 = (W_0 + D_0)\,R - C(D_0)$

If $\lambda_1 > 0: I_1 = W_1 + D_1 \Rightarrow D_1 = \left(\frac{1-c_1}{c_1}\right)[(W_0 + D_0)\,R - C(D_0)]$
If $\lambda_1 = 0: C'(D_1) = 1 \Rightarrow D_1 = cnst$

In both cases, R and D_0 are determined simultaneously from the system:

$$\mathcal{L}_R : p'(R)\left\{\frac{[(W_0+D_0)\,R - C(D_0)]}{c_1}\left(\overline{R}+c_1-1\right) + D_1 - C(D_1)\right\}$$
$$+\frac{p(R)}{c_1}\left(\overline{R}+ (c_1-1)\,C'(D_1)\right)(W_0+D_0) = 0$$
$$\mathcal{L}_{D_0} : \frac{p(R)}{c_1}\left(\overline{R} + (c_1 - 1)\,C'(D_1)\right)[R - C'(D_0)] = 0.$$

The second order conditions reduce to prove that the following matrix of second derivatives for R and D_0 is negative semi-definite, this is, that its leading principal minors alternate in sign:

$$L_c^2 = \begin{bmatrix} \mathcal{L}_{RR} & \frac{p(R)}{c_1}\left(\overline{R}+ (c_1 - 1)\,C'(D_1)\right) \\ \frac{p(R)}{c_1}\left(\overline{R}+ (c_1 - 1)\,C'(D_1)\right) & -\frac{p(R)}{c_1}\left(\overline{R}+ (c_1 - 1)\,C'(D_1)\right)C''(D_0) \end{bmatrix},$$

where

$$\mathcal{L}_{RR} = p''(R)W_2^c + 2p'(R)\left(\tfrac{W_0+D_0}{c_1}\right)\left(\overline{R}+ (c_1-1)\,C'(D_1)\right)$$
$$-p(R)\,(W_0+D_0)^2\,\Phi(\lambda_1, R, W_0, c_1),$$
$$\Phi(\lambda_1, R, W_0, c_1) = \begin{cases} C''(D_1)\left(\frac{1-c_1}{c_1}\right)^2 & \lambda_1 > 0 \\ 0 & \lambda_1 = 0 \end{cases} \geq 0,$$
$$W_2^c = [(W_0 + D_0)\,R - C(D_0)]\left(\tfrac{\overline{R}+c_1-1}{c_1}\right) + D_1 - C(D_1) \geq 0.$$

$$\det\left[L^2\right]_1 = \mathcal{L}_{RR} \leq 0$$
$$\det\left[L^2\right]_2 = -\tfrac{p(R)}{c_1}\left(\overline{R}+ (c_1 - 1)\,C'(D_1)\right)\left\{\mathcal{L}_{RR}C''(D_0) + \tfrac{p(R)}{c_1}\left(\overline{R}+ (c_1-1)\,C'(D_1)\right)\right\} \geq 0$$
$$\Leftrightarrow \left\{p''(R)W_2^c + 2p'(R)\left(\tfrac{W_0+D_0}{c_1}\right)\left(\overline{R} + (c_1-1)\,C'(D_1)\right)\right\}C''(D_0)$$
$$+p(R)\left[\tfrac{\left(\overline{R}+(c_1-1)C'(D_1)\right)}{c_1} - (W_0 + D_0)^2\,C''(D_0)\Phi(\lambda_1, R, W_0, c_1)\right] \leq 0.$$

d) Both periods binding requirements ($c_0 = c_1 = c$)

$\mu_0 > 0 \Rightarrow W_0 = cI_0 \Rightarrow I_0 = \frac{W_0}{c}$

$\mu_1 > 0 \Rightarrow W_1 = cI_1 \Rightarrow I_1 = \frac{W_1}{c}$

$W_1 = \frac{W_0}{c}\,(R + c - 1) + D_0 - C\,(D_0)$

$\theta = \frac{p(R)}{c}\left\{\overline{R} + (c - 1)\,C'(D_1)\right\}$

If $\lambda_0 > 0 \Rightarrow I_0 = W_0 + D_0 \Rightarrow D_0 = \left(\frac{1-c}{c}\right)W_0 = cnst$

If $\lambda_1 > 0 \Rightarrow I_1 = W_1 + D_1 \Rightarrow D_1 = \left(\frac{1-c}{c}\right)\left[\frac{1}{c}W_0R - C(\left(\frac{1-c}{c}\right)W_0)\right]$

If $\lambda_1 = 0 \Rightarrow C'(D_1) = 1 \Rightarrow D_1 = cnst$

If $\lambda_0 = 0 \Rightarrow \theta\,(1 - C'\,(D_0)) = 0 \Rightarrow C'\,(D_0) = 1 \Rightarrow D_0 = cnst$

If $\lambda_1 > 0 \Rightarrow I_1 = W_1 + D_1 \Rightarrow D_1 = \left(\frac{1-c}{c}\right)\left[\frac{W_0}{c}\,(R + c - 1) + D_0 - C(D_0)\right]$

If $\lambda_1 = 0 \Rightarrow C'(D_1) = 1 \Rightarrow D_1 = cnst$

In all cases, R is determined alone from

$$\mathcal{L}_R : p'(R)\left\{\left[\tfrac{W_0}{c}(R+c-1)+D_0-C(D_0)\right]\left(\tfrac{\overline{R}+c-1}{c}\right)+D_1-C(D_1)\right\}$$
$$+\frac{p(R)}{c^2}\left(\overline{R}+(c-1)\,C'(D_1)\right)W_0=0,$$

and the relevant second order condition is:

$$p''(R)W_2^d+2p'(R)\tfrac{W_0}{c^2}\left(\overline{R}+(c-1)\,C'(D_1)\right)-p(R)\tfrac{W_0^3}{c^2}\Phi(\lambda_1,R,W_0,c)\leq 0,$$

which always holds true, for

$$W_2^d=\left[\tfrac{W_0}{c}(R+c-1)+D_0-C(D_0)\right]\left(\tfrac{\overline{R}+c-1}{c}\right)+D_1-C(D_1)\geq 0.$$

4.7.2 Comparative Statics

So far, for given values of the parameters W_0, c_0 and c_1, I have derived a set of first order conditions determining the optimal solution for R, D_0, D_1, I_0, I_1, and W_1; and each possible combination of the sign of the multipliers. I have also shown that D_1, I_0, I_1 and W_1 can be written in terms of W_0, c_0, c_1, R and D_0, where either D_0 is constant, or R and D_0 are determined simultaneously from the system:

$$\mathcal{L}_R=\frac{\partial\mathcal{L}}{\partial R}=0$$

$$\mathcal{L}_{D_0}=\frac{\partial\mathcal{L}}{\partial D_0}=0$$

From the implicit function theorem it is known that

$$\begin{bmatrix}\frac{dR}{d\varphi}\\ \\ \frac{dD_0}{d\varphi}\end{bmatrix}=\underbrace{\begin{bmatrix}\mathcal{L}_{RR} & \mathcal{L}_{RD_0}\\ \mathcal{L}_{D_0R} & \mathcal{L}_{D_0D_0}\end{bmatrix}}_{L^2}^{-1}\begin{bmatrix}\mathcal{L}_{R\varphi}\\ \\ \mathcal{L}_{D_0\varphi}\end{bmatrix},$$

or simply

$$\frac{dR}{d\varphi}=\frac{\mathcal{L}_{D_0\varphi}\mathcal{L}_{RD_0}-\mathcal{L}_{R\varphi}\mathcal{L}_{D_0D_0}}{\mathcal{L}_{RR}\mathcal{L}_{D_0D_0}-\mathcal{L}_{RD_0}\mathcal{L}_{D_0R}},$$

$$\frac{dD_0}{d\varphi}=\frac{\mathcal{L}_{R\varphi}\mathcal{L}_{D_0R}-\mathcal{L}_{D_0\varphi}\mathcal{L}_{RR}}{\mathcal{L}_{RR}\mathcal{L}_{D_0D_0}-\mathcal{L}_{RD_0}\mathcal{L}_{D_0R}},$$

where $\varphi\in\{W_0,c_0,c_1\}$.

$$sign\left[\frac{dR}{d\varphi}\right]=sign\left[\frac{1}{\det L^2}\right]\times sign\left[\mathcal{L}_{D_0\varphi}\mathcal{L}_{RD_0}-\mathcal{L}_{R\varphi}\mathcal{L}_{D_0D_0}\right],$$
$$sign\left[\frac{dD_0}{d\varphi}\right]=sign\left[\frac{1}{\det L^2}\right]\times sign\left[\mathcal{L}_{R\varphi}\mathcal{L}_{D_0R}-\mathcal{L}_{D_0\varphi}\mathcal{L}_{RR}\right].$$

But since SOCs imply $\det L^2\geq 0$,

$$sign\left[\frac{dR}{d\varphi}\right]=sign\left[\mathcal{L}_{D_0\varphi}\mathcal{L}_{RD_0}-\mathcal{L}_{R\varphi}\mathcal{L}_{D_0D_0}\right],$$
$$sign\left[\frac{dD_0}{d\varphi}\right]=sign\left[\mathcal{L}_{R\varphi}\mathcal{L}_{D_0R}-\mathcal{L}_{D_0\varphi}\mathcal{L}_{RR}\right].$$

a) Unregulated case:

$\mathcal{L}_R : p'(R)\left\{\left[(W_0 + D_0)\,R - C(D_0) + D_1\right]\overline{R} - C(D_1)\right\} + p(R)\overline{R}\,(W_0 + D_0) = 0$

$\mathcal{L}_{D_0} : p(R)\overline{R}\,[R - C'(D_0)] = 0$

$\varphi = W_0$

$$\begin{bmatrix} \mathcal{L}_{RR} & \mathcal{L}_{RD_0} \\ \mathcal{L}_{D_0R} & \mathcal{L}_{D_0D_0} \end{bmatrix} = L_a^2 = \begin{bmatrix} p''(R)W_2 + 2p'(R)(W_0 + D_0)\overline{R} & p(R)\overline{R} \\ p(R)\overline{R} & -p(R)\overline{R}C''(D_0) \end{bmatrix}$$

$\mathcal{L}_{RW_0} = (p'(R)R + p(R))\,\overline{R}$

$\mathcal{L}_{D_0W_0} = 0$

$$sign\left[\frac{dR}{dW_0}\right] = sign\left[(p'(R)R + p(R))\,p(R)\overline{R}^2C''(D_0)\right]$$

$$sign\left[\frac{dD_0}{dW_0}\right] = sign\left[(p'(R)R + p(R))\,\overline{R}^2p(R)\right]$$

As I have assumed that for the unregulated bank $p'(R)R + p(R) < 0$, then

$$\frac{dR}{dW_0} < 0 \text{ and } \frac{dD_0}{dW_0} < 0.$$

b) Binding requirement at $t = 0$ only

$\mathcal{L}_R$: $p'(R)\left\{\left(\frac{W_0}{c_0}(R - 1 + c_0) + D_0 - C(D_0) + D_1\right)\overline{R} - C(D_1)\right\} + p(R)\overline{R}\frac{W_0}{c_0} = 0$

$\mathcal{L}_{D_0} : D_0 - fnct\,(W_0, c_0) = 0$

In this case D_0 is independent of R, therefore comparative statics reduce to

$$\frac{dR}{d\varphi} = -\left[\mathcal{L}_{RR}\right]^{-1}\mathcal{L}_{R\varphi}$$

But since $\mathcal{L}_{RR} < 0$,

$$sign\left[\frac{dR}{d\varphi}\right] = sign\left[\mathcal{L}_{R\varphi}\right]$$

Notice that for $c_0 \geq \widetilde{c}_0$:

$$\frac{\partial D_0}{\partial c_0} = \begin{cases} -\frac{W_0}{c_0^2} & \lambda_0 > 0 \\ 0 & \lambda_0 = 0 \end{cases} \leq 0$$

$$\frac{\partial D_0}{\partial W_0} = \begin{cases} \frac{1-c_0}{c_0} & \lambda_0 > 0 \\ 0 & \lambda_0 = 0 \end{cases} \geq 0$$

Then

$\mathcal{L}_{RW_0} = \frac{\overline{R}}{c_0}\left[p'(R)\,(R - (1 - c_0)\,C'(D_0)) + p(R)\right] \geq 0,$

$\mathcal{L}_{Rc_0} = -\overline{R}\frac{W_0}{c_0^2}\left[p'(R)\,(R - C'(D_0)) + p(R)\right] \leq 0,$

$\mathcal{L}_{Rc_1} = 0.$

This comes from the first order condition on R ($\mathcal{L}_R = 0$):

$$p'(R)\,(R - C'(D_0)) + p(R)$$
$$= -\tfrac{c_0}{\overline{R}W_0}p'(R)\left\{C'(D_0)W_0\overline{R} + [C'(D_0)D_0 - C(D_0)]\,\overline{R} + \left[D_1\overline{R} - C(D_1)\right]\right\} \geq 0.$$

This is because $C'(0)$ bounded and $C(.)$ convex, imply $C'(D_t)D_t \geq D_t$ for all D_t.[19]

Also

$$p'(R)\,(R - (1 - c_0)\,C'(D_0)) + p(R)$$
$$= -\tfrac{c_0}{\overline{R}W_0}p'(R)\left\{[C'(D_0)D_0 - C(D_0)]\,\overline{R} + \left[D_1\overline{R} - C(D_1)\right]\right\} \geq 0$$

Hence, for $c_0 \geq \widetilde{c_0}$:

$$\frac{dR}{dW_0} \geq 0,$$
$$\frac{dR}{dc_0} \leq 0.$$

c) Binding requirement at $t = 1$ only

$$\mathcal{L}_R\text{: } p'(R)\left\{\frac{[(W_0+D_0)\,R - C(D_0)]}{c_1}\,(\overline{R}+c_1-1) + D_1 - C(D_1)\right\}$$
$$+\frac{p(R)}{c_1}\,(\overline{R} + (c_1-1)\,C'(D_1))\,(W_0+D_0) = 0$$

$$\mathcal{L}_{D_0} : \frac{p(R)}{c_1}\,(\overline{R} + (c_1 - 1)\,C'(D_1))\,[R - C'(D_0)] = 0$$

$$\begin{bmatrix} \mathcal{L}_{RR} & \mathcal{L}_{RD_0} \\ \mathcal{L}_{D_0R} & \mathcal{L}_{D_0D_0} \end{bmatrix} = L_c^2$$

$$= \begin{bmatrix} \mathcal{L}_{RR} & \frac{p(R)}{c_1}\,(\overline{R} + (c_1 - 1)\,C'(D_1)) \\ \frac{p(R)}{c_1}\,(\overline{R} + (c_1 - 1)\,C'(D_1)) & -\frac{p(R)}{c_1}\,(\overline{R} + (c_1 - 1)\,C'(D_1))\,C''(D_0) \end{bmatrix},$$

where

$$\mathcal{L}_{RR} = p''(R)W_2^c + 2p'(R)\left(\tfrac{W_0+D_0}{c_1}\right)(\overline{R} + (c_1-1)\,C'(D_1))$$
$$-p(R)\,(W_0 + D_0)^2\,\Phi(\lambda_1, R, W_0, c_1).$$

Notice that for $c_1 \geq \widetilde{c_1}$:

$$\tfrac{\partial D_1}{\partial c_1} = \begin{cases} \left(-\tfrac{1}{c_1^2}\right)[(W_0+D_0)\,R - C(D_0)] & \lambda_1 > 0 \\ 0 & \lambda_1 = 0 \end{cases} \leq 0,$$

$$\mathcal{L}_{RW_0} = [p'(R)R + p(R)]\left(\tfrac{\overline{R}+(c_1-1)C'(D_1)}{c_1}\right) - p(R)R\,(W_0 + D_0)\,\Phi(\lambda_1, R, W_0, c_1),$$

$$\mathcal{L}_{Rc_0} = 0,$$

$$\mathcal{L}_{Rc_1} = p'(R)\left[\tfrac{(W_0+D_0)R - C(D_0)}{c_1^2}\right](C'(D_1) - \overline{R}) + p(R)\left(\tfrac{W_0+D_0}{c_1^2}\right)(C'(D_1) - \overline{R})$$
$$+p(R)\,(W_0 + D_0)\left[\tfrac{(W_0+D_0)R - C(D_0)}{(1-c_1)c_1}\right]\Phi(\lambda_1, R, W_0, c_1),$$

$$\mathcal{L}_{D_0W_0} = -p(R)R\,(R - C'(D_0))\,\Phi(\lambda_1, R, W_0, c_1) = 0,$$

$$\mathcal{L}_{D_0c_0} = 0,$$

[19] Define $F(D) = C'(D)D - C(D)$. As $C'(0)$ is bounded, $F(0) = 0, F'(D) = C''(D)D \geq 0$. Hence $F(D) \geq 0\ \forall D$, i.e. $C'(D)D \geq C(D)$.

$$\mathcal{L}_{D_0c_1} = -\tfrac{p(R)}{c_1^2}\left(\overline{R}+(c_1-1)C'(D_1)\right)(R-C'(D_0))$$
$$+\tfrac{p(R)}{c_1}\left[C'(D_1)+(c_1-1)\,C''(D_1)\tfrac{\partial D_1}{\partial c_1}\right](R-C'(D_0))=0,$$

because a non binding requirement at $t=0$ implies $C'(D_0)=R$.

Hence,

$$sign\left[\frac{dR}{d\varphi}\right] = sign\left[\mathcal{L}_{D_0\varphi}\mathcal{L}_{RD_0}-\mathcal{L}_{R\varphi}\mathcal{L}_{D_0D_0}\right]$$
$$= -sign\left[\mathcal{L}_{D_0D_0}\right]sign\left[\mathcal{L}_{R\varphi}\right] = sign\left[\mathcal{L}_{R\varphi}\right]$$

From $\mathcal{L}_R=0$,

$$[p'(R)R+p(R)](W_0+D_0)$$
$$= p'(R)C(D_0)+p'(R)\left(\tfrac{C(D_1)-D_1}{\overline{R}+c_1-1}\right)c_1+p(R)\frac{(1-c_1)}{\overline{R}+c_1-1}(W_0+D_0)\left(C'(D_1)-1\right)$$

Then,

$$\mathcal{L}_{Rc_1} = \left\{[p'(R)R+p(R)](W_0+D_0)-p'(R)C(D_0)\right\}\left(\tfrac{C'(D_1)-\overline{R}}{c_1^2}\right)$$
$$+p(R)\left(W_0+D_0\right)\left[\tfrac{(W_0+D_0)R-C(D_0)}{(1-c_1)c_1}\right]\Phi(\lambda_1,R,W_0,c_1)$$
$$= \left\{p'(R)\left(\tfrac{C(D_1)-D_1}{\overline{R}+c_1-1}\right)c_1+p(R)\frac{(1-c_1)}{\overline{R}+c_1-1}(W_0+D_0)\left(C'(D_1)-1\right)\right\}\left(\tfrac{C'(D_1)-\overline{R}}{c_1^2}\right)$$
$$+p(R)\left(W_0+D_0\right)\left[\tfrac{(W_0+D_0)R-C(D_0)}{(1-c_1)c_1}\right]\Phi(\lambda_1,R,W_0,c_1)$$

For all values of c_1 where $\lambda_1>0$:

$$\mathcal{L}_{Rc_1} = -p'(R)\frac{(C(D_1)-D_1)(\overline{R}-C'(D_1))}{(\overline{R}+c_1-1)c_1}$$
$$+p(R)(W_0+D_0)\left(\tfrac{1-c_1}{c_1^2}\right)\left\{[(W_0+D_0)R-C(D_0)]\tfrac{C''(D_1)}{c_1}-\frac{(C'(D_1)-1)}{\overline{R}+c_1-1}\left(\overline{R}-C'(D_1)\right)\right\}$$

The first term is positive, hence

$$\frac{dR}{dc_1}\geq 0 \text{ if } A(c_1)=[(W_0+D_0)R-C(D_0)]\,\tfrac{C''(D_1)}{c_1}-\frac{(C'(D_1)-1)}{\overline{R}+c_1-1}\left(\overline{R}-C'(D_1)\right)\geq 0$$

When the requirement just starts binding (at $c_1=\widetilde{c_1}$), by continuity $\mu_1=0$ and $\lambda_1>0$, which implies that $C'(D_1)=\overline{R}$.

Hence,

$$A(\widetilde{c_1})=[(W_0+D_0)R-C(D_0)]\,\tfrac{C''(D_1)}{c_1}\geq 0,$$

so $\dfrac{dR}{dc_1}\geq 0$, and by continuity this is also true for values of c_1 in the neighbourhood of $\widetilde{c_1}$.

When $c_1=1$ (or if $\lambda_1=0$)

$A(1)=[(W_0+D_0)R-C(D_0)]\,C''(D_1)\geq 0$, and indeed

$\mathcal{L}_{Rc_1}\big|_{c_1=1}=p'(R)\left(\tfrac{C(D_1)-D_1}{\overline{R}}\right)\left(C'(D_1)-\overline{R}\right)=0$, because $C(D_1)=D_1$.

Therefore, $\left.\dfrac{dR}{dc_1}\right|_{c_1=1}=0.$

Also notice that $\lim\limits_{c_1\to 1}A(c_1)=[(W_0+D_0)R-C(D_0)]\,C''(D_1)\geq 0$, which implies that in the neighbourhood of $c_1=1$, $\dfrac{dR}{dc_1}\geq 0$.

$$\mathcal{L}_{RW_0} = [p'(R)R + p(R)]\left(\frac{\overline{R}+(c_1-1)C'(D_1)}{c_1}\right) - p(R)R\,(W_0 + D_0)\,\Phi(\lambda_1, R, W_0, c_1)$$

Until $c_1 = \widetilde{c_1}$ no requirement bind, therefore by assumption $p'(R)R + p(R) < 0$. Hence, when the requirement just starts binding $\left(C'(D_1) = \overline{R}\right)$:

$$\mathcal{L}_{RW_0} = [p'(R)R + p(R)]\,\overline{R} - p(R)R\,(W_0 + D_0)\,\Phi(\lambda_1, R, W_0, c_1) < 0$$

Therefore, $\left.\dfrac{dR}{dW_0}\right|_{c_1=\widetilde{c_1}} < 0.$

Also, replacing $p'(R)R + p(R)$ from $\mathcal{L}_R = 0$:

$$\mathcal{L}_{RW_0} = p'(R)\left[C(D_0) + \left(\frac{C(D_1)-D_1}{\overline{R}+c_1-1}\right)c_1\right]\left(\frac{\overline{R}+(c_1-1)C'(D_1)}{c_1(W_0+D_0)}\right)$$
$$+p(R)\left\{\left(\frac{1-c_1}{c_1}\right)(C'(D_1) - 1)\left(\frac{\overline{R}+(c_1-1)C'(D_1)}{\overline{R}+c_1-1}\right) - R\,(W_0 + D_0)\,\Phi(\lambda_1, R, W_0, c_1)\right\}$$

If $c_1 = 1$:

$$\mathcal{L}_{RW_0} = p'(R)C(D_0)\left(\frac{\overline{R}}{W_0+D_0}\right) < 0.$$

Therefore, $\left.\dfrac{dR}{dW_0}\right|_{c_1=1} < 0.$

Summing up, provided $R > R^u$ (as is likely to happen, given the sign of $\dfrac{dR}{dc_1}$ around the corner values of c_1), $\dfrac{dR}{dW_0} \leq 0.$

d) Both periods binding requirements ($c_0 = c_1 = c$)

$$\mathcal{L}_R\text{: } p'(R)\left\{\left[\tfrac{W_0}{c}\,(R + c - 1) + D_0 - C(D_0)\right]\left(\tfrac{\overline{R}+c-1}{c}\right) + D_1 - C(D_1)\right\}$$
$$+p(R)\left(\overline{R} + (c-1)\,C'(D_1)\right)\tfrac{W_0}{c^2} = 0$$

$$\mathcal{L}_{D_0} : D_0 - fnct\,(W_0, c_0) = 0$$

In this case D_0 is independent of R, therefore comparative statics reduce to

$$\frac{dR}{d\varphi} = -\,[\mathcal{L}_{RR}]^{-1}\,\mathcal{L}_{R\varphi}$$

But since $\mathcal{L}_{RR} < 0$,

$$sign\left[\frac{dR}{d\varphi}\right] = sign\,[\mathcal{L}_{R\varphi}]$$

For all $c \geq \max\{\widetilde{c_0}, \widetilde{c_1}\}$:

$$\frac{\partial D_0}{\partial c} = \begin{cases} -\frac{W_0}{c^2} & \lambda_0 > 0 \\ 0 & \lambda_0 = 0 \end{cases} \leq 0$$

$$\frac{\partial D_1}{\partial c} = \begin{cases} -\frac{1}{c^2}\left\{W_1 + W_0\left(\frac{1-c}{c}\right)[R - C'(D_0)]\right\} & \lambda_1 > 0 \\ 0 & \lambda_1 = 0 \end{cases} \leq 0$$

$$\frac{\partial D_0}{\partial W_0} = \begin{cases} \frac{1-c}{c} & \lambda_0 > 0 \\ 0 & \lambda_0 = 0 \end{cases} \geq 0$$

$$\frac{\partial D_1}{\partial W_0} = \begin{cases} \left(\frac{1-c}{c^2}\right)\left[R - C'(D_0)(1-c)\right] & \lambda_1 > 0 \\ 0 & \lambda_1 = 0 \end{cases} \geq 0$$

$$\mathcal{L}_{Rc}\text{: } p'(R)\left\{\left[-\tfrac{W_0}{c^2}(R-1) + (1-C'(D_0))\tfrac{\partial D_0}{\partial c}\right]\left(\tfrac{\overline{R}+c-1}{c}\right) + (1-C'(D_1))\tfrac{\partial D_1}{\partial c}\right\}$$
$$+p'(R)\left\{-\left[\tfrac{W_0}{c}(R+c-1) + D_0 - C(D_0)\right]\left(\tfrac{\overline{R}-1}{c^2}\right)\right\}$$
$$-2p(R)\frac{W_0}{c^3}\left(\overline{R} + (c-1)C'(D_1)\right) + \frac{p(R)}{c^2}W_0\left\{C'(D_1) + (c-1)C''(D_1)\tfrac{\partial D_1}{\partial c}\right\}$$
$$= p'(R)\left\{-\tfrac{W_0}{c^3}(R - C'(D_0))\left(\overline{R} + (c-1)C'(D_1)\right) - \tfrac{W_1}{c^2}\left(\overline{R} - C'(D_1)\right)\right\}$$
$$-\frac{p(R)}{c^3}W_0\left\{2\overline{R} - (2-c)C'(D_1)\right\}$$
$$+\frac{p(R)}{c^5}W_0\left[W_1 + W_0(1-c)(R - C'(D_0))\right](1-c)C''(D_1)$$

$$\mathcal{L}_{RW_0}\text{: } p'(R)\left\{\left[\left(\tfrac{R+c-1}{c}\right) + (1-C'(D_0))\tfrac{\partial D_0}{\partial W_0}\right]\left(\tfrac{\overline{R}+c-1}{c}\right) + (1-C'(D_1))\tfrac{\partial D_1}{\partial W_0}\right\}$$
$$+\tfrac{p(R)}{c^2}\left\{\left(\overline{R} + (c-1)C'(D_1)\right) + W_0(c-1)C''(D_1)\tfrac{\partial D_1}{\partial W_0}\right\}$$

From the FOC ($\mathcal{L}_R = 0$):

$$[p'(R)R + p(R)]\tfrac{W_0}{c^2}\left(\overline{R} + (c-1)C'(D_1)\right)$$
$$= -p'(R)\left\{[D_0 - C(D_0)]\left(\tfrac{\overline{R}+c-1}{c}\right) + D_1 - C(D_1) + \ldots\right.$$
$$\left.\ldots + \tfrac{W_0}{c^2}(c-1)\left[\overline{R} + c - 1 + R(1 - C'(D_1))\right]\right\},$$

or equivalently

$$[p'(R)R + p(R)]\left(\overline{R} + c - 1\right)\tfrac{W_0}{c^2} = -p(R)(c-1)(C'(D_1) - 1)\tfrac{W_0}{c^2}$$
$$-p'(R)\left\{[D_0 - C(D_0)]\left(\tfrac{\overline{R}+c-1}{c}\right) + D_1 - C(D_1) + \tfrac{W_0}{c^2}(c-1)\left(\overline{R} + c - 1\right)\right\}$$

In particular, when $c = 1$

$[p'(R)R + p(R)]\overline{R}W_0 = 0$, so $R(c=1) = R^*$

and

$$\mathcal{L}_{Rc}\text{: } p'(R)W_0\left\{-(R - C'(0))R - R\left(\overline{R} - C'(0)\right)\right\} - p(R)W_0\left\{2\overline{R} - C'(0)\right\}$$
$$= -[p'(R)R + p(R)]W_0\left(2\overline{R} - C'(0)\right) + p'(R)W_0\overline{R}C'(0) = p'(R)\overline{R}W_0C'(0) < 0.$$

Therefore, $\left.\dfrac{dR}{dc}\right|_{c=1} < 0.$

If $\tilde{c}_0 < \tilde{c}_1$, when c reaches $\tilde{c}_1$, period 0 requirement is binding and period 1 requirement starts to bind, which in terms of the multipliers means $\lambda_1 > 0$ and $\mu_1 = 0$, so $C'(D_1) = \overline{R}$ and $C''(D_1) = 0$.

$\mathcal{L}_{Rc}$: $-\tfrac{W_0}{c^2}\overline{R}\left[p'(R)(R - C'(D_0)) + p(R)\right]$

Also, $\lambda_1 > 0$ implies $D_1 = \left(\frac{1-c}{c}\right)W_1$, therefore $\mathcal{L}_R = 0$ reduces to its form in case (b) and

$$p'(R)\left\{\left(\tfrac{W_0}{c}\left(R - C'(D_0) + C'(D_0) + c - 1\right) + D_0 - C(D_0)\right)\tfrac{\overline{R}}{c} - C(D_1)\right\}$$
$$+ p(R)\overline{R}\tfrac{W_0}{c} = 0$$

$$p'(R)\tfrac{W_0}{c}\left(R - C'(D_0)\right)\tfrac{\overline{R}}{c} + p(R)\overline{R}c\tfrac{W_0}{c^2}$$
$$= -p'(R)\left\{\left(\tfrac{W_0}{c}\left(C'(D_0) + c - 1\right) + D_0 - C(D_0)\right)\tfrac{\overline{R}}{c} - C(D_1)\right\}$$
$$\left[p'(R)\left(R - C'(D_0)\right) + p(R)\right]\overline{R}\tfrac{W_0}{c^2}$$
$$= -p'(R)\left\{\left(\tfrac{W_0}{c}\left(C'(D_0) + c - 1\right) + D_0 - C(D_0)\right)\tfrac{\overline{R}}{c} - C(D_1)\right\} + p(R)\left(1 - c\right)\overline{R}\tfrac{W_0}{c^2}.$$

If $\widetilde{c_0} \geq \widetilde{c_1}$, when c reaches $\widetilde{c_0}$, period 1 requirement is binding and period 0 requirement starts to bind, which in terms of the multipliers means $\lambda_0 > 0$ and $\mu_0 = 0$, so $C'(D_0) = R$ and $D_0 = \left(\frac{1-c}{c}\right) W_0$.

This time, however, neither $\mathcal{L}_{Rc}$ nor $\mathcal{L}_{RW_0}$ have a clear sign.

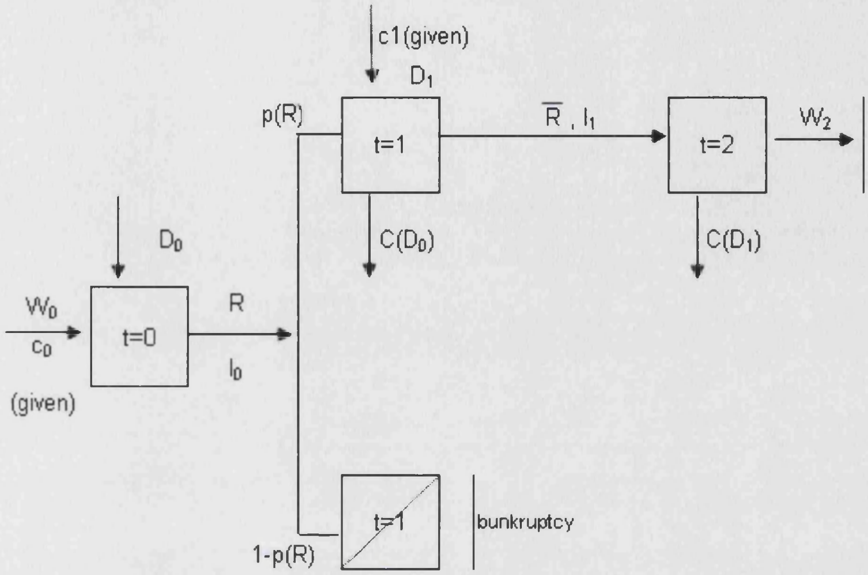

Figure 4-1: Timeline for the optimal decision of the regulated bank.

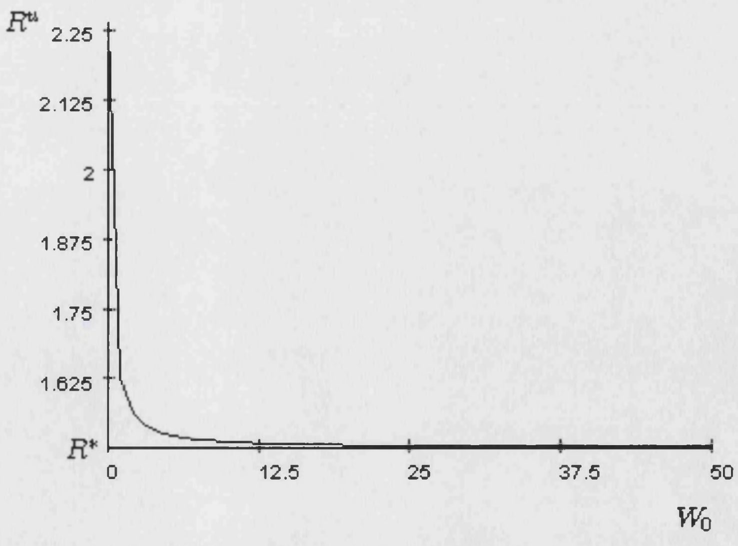

Figure 4-2: Example of the equilibrium relationship between risk-taking and the initial level of equity for the unregulated bank.

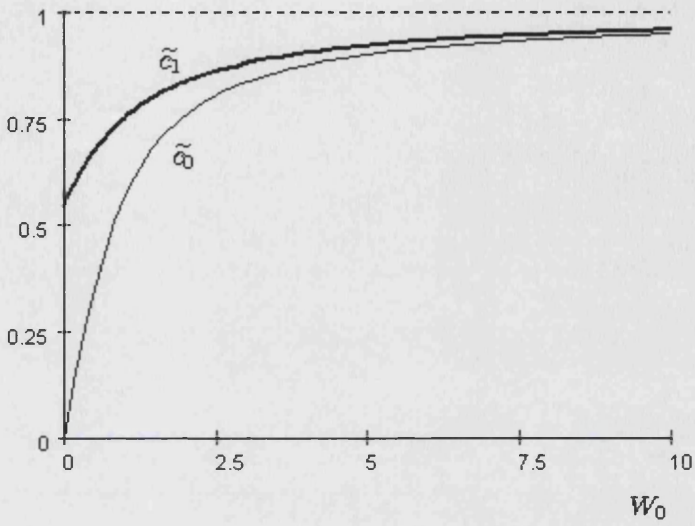

Figure 4-3: Example of binding capital requirement thresholds as a function of initial equity.

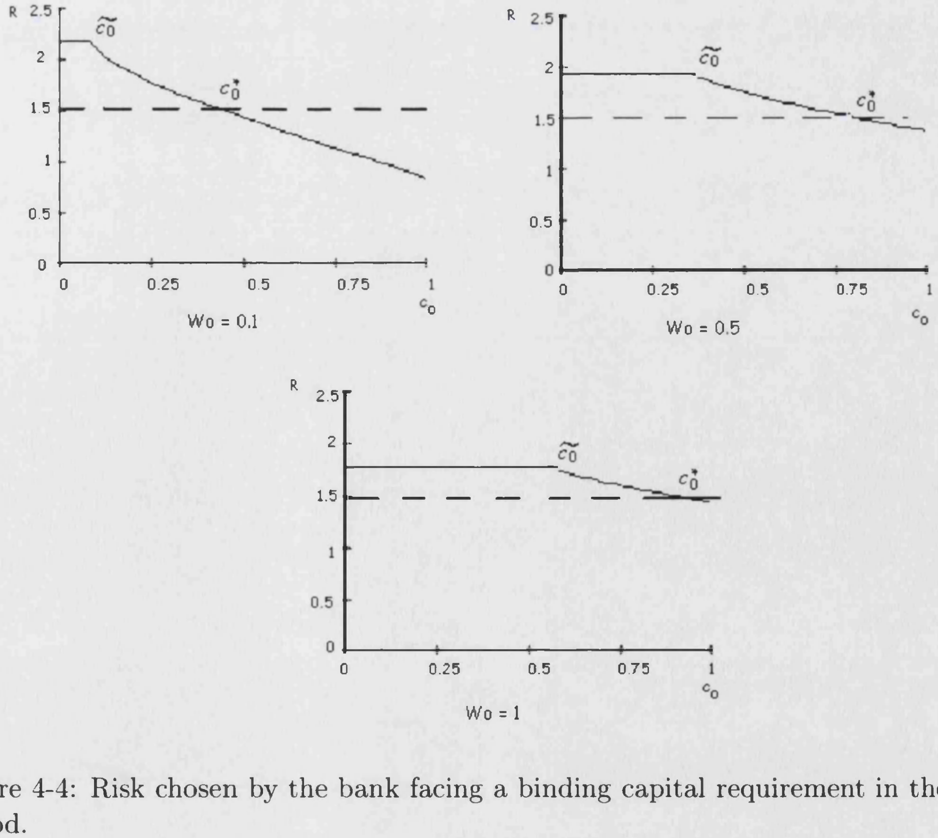

Figure 4-4: Risk chosen by the bank facing a binding capital requirement in the initial period.

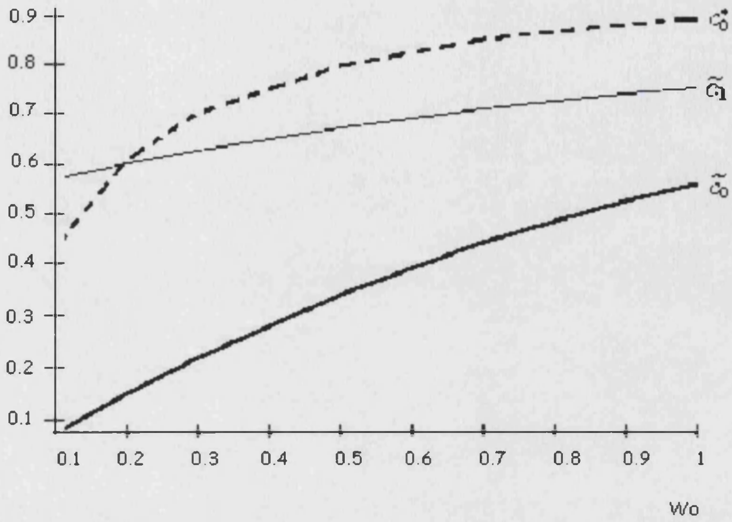

Figure 4-5: Optimal value of the capital requirement in the initial period.

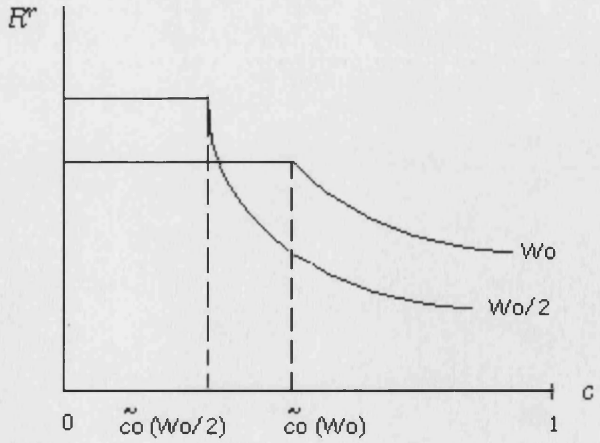

Figure 4-6: Better capitalised banks take on more risk when the capital requirement binds in the initial period only.

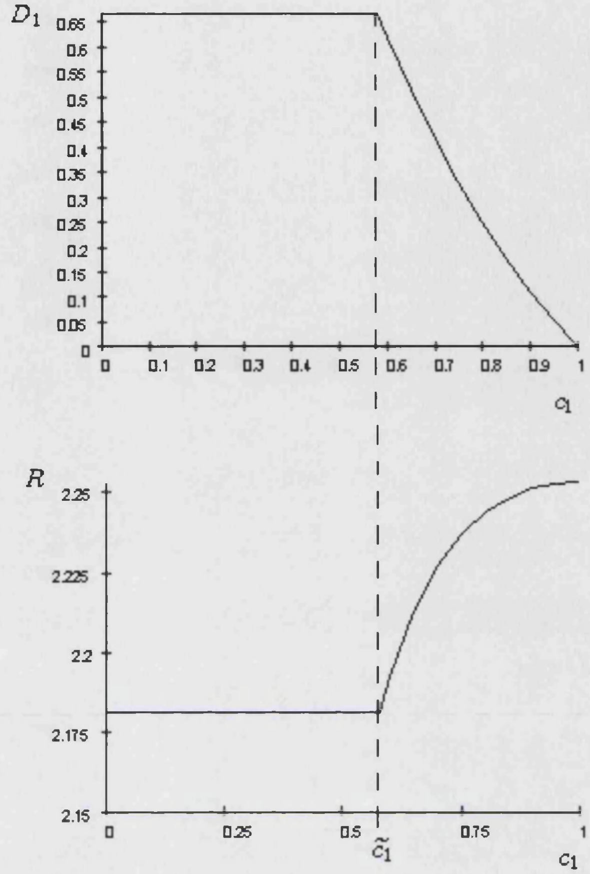

Figure 4-7: Example of the evolution of the equilibrium values of R and D_1, when only the intermediate period requirement binds.

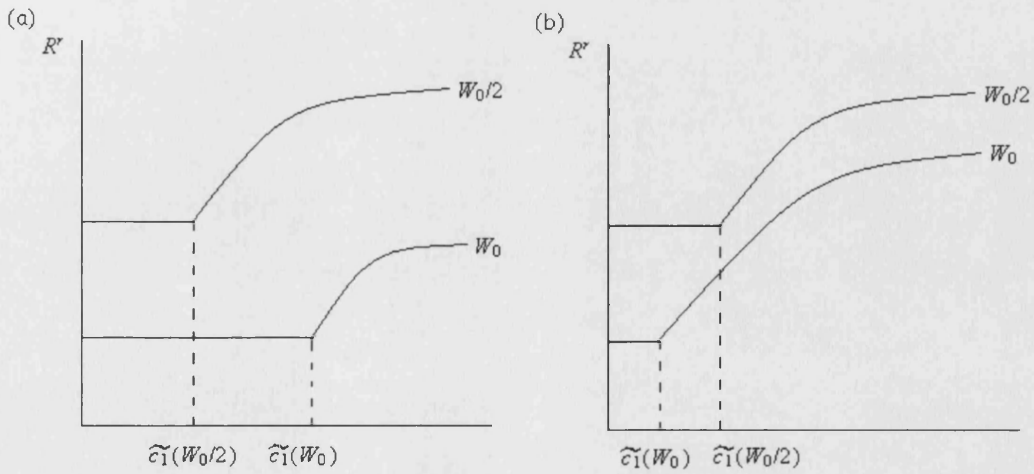

Figure 4-8: Risk-taking in banks with different levels of equity, under an interim period binding capital requirement.

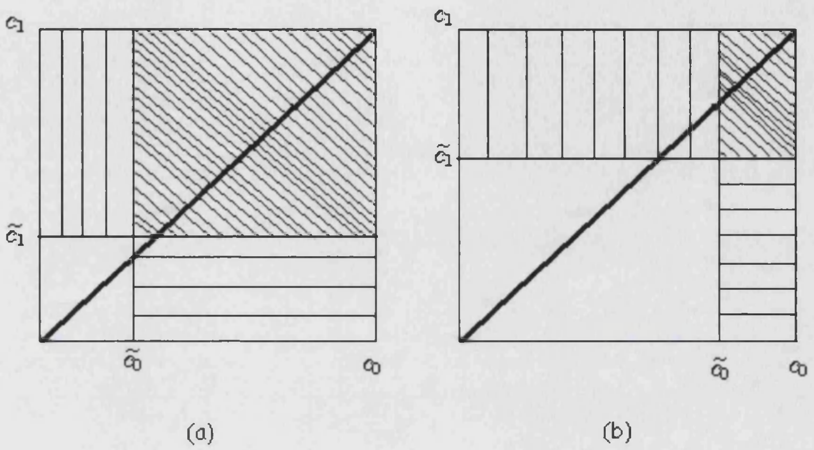

Figure 4-9: Binding capital requirements: (a) $\widetilde{c_0} < \widetilde{c_1}$, (b) $\widetilde{c_0} \geq \widetilde{c_1}$

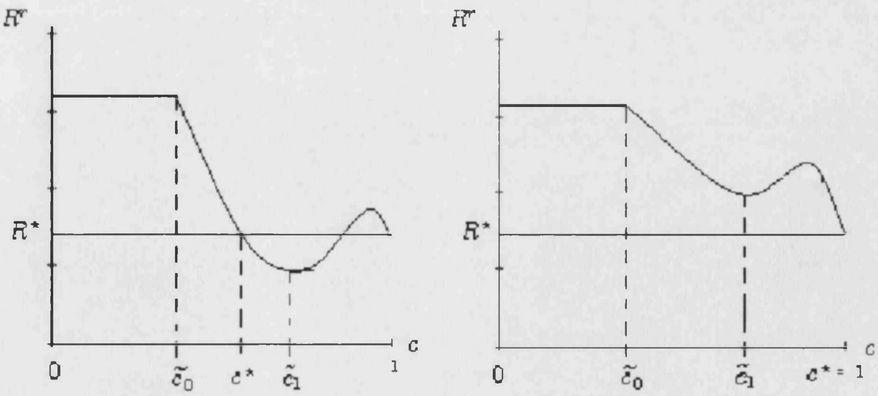

Figure 4-10: Possible paths of R^r as a function of c, with constant capital requirements and $\widetilde{c_0} < \widetilde{c_1}$.

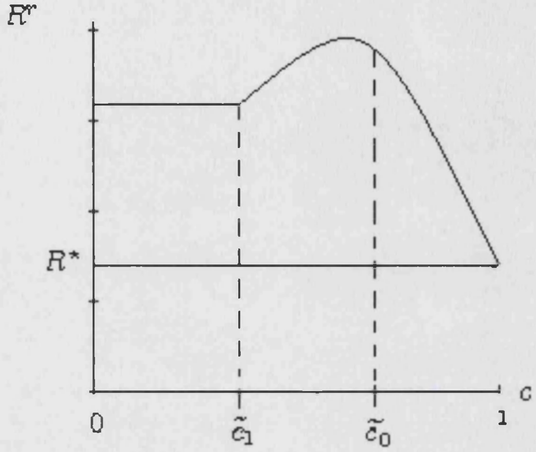

Figure 4-11: Possible path of R^r as a function of c, with constant capital requirements and $\widetilde{c_0} > \widetilde{c_1}$.

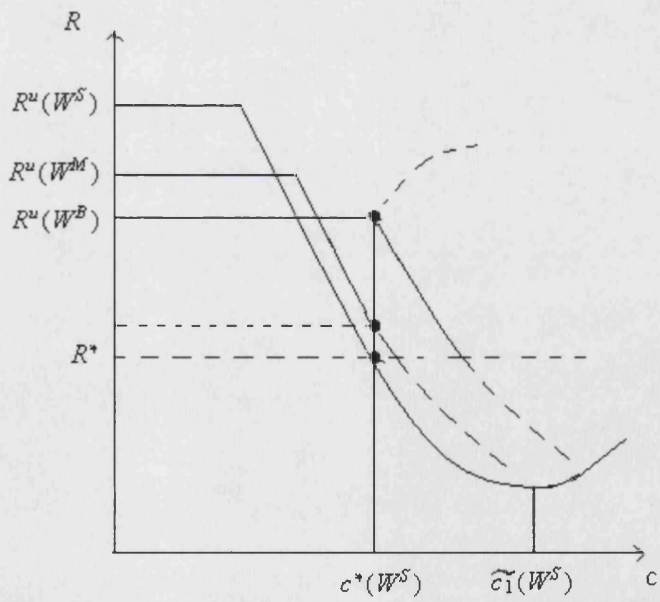

Figure 4-12: Distribution of risks when capital requirements are used to force the smallest bank to the social optimum.

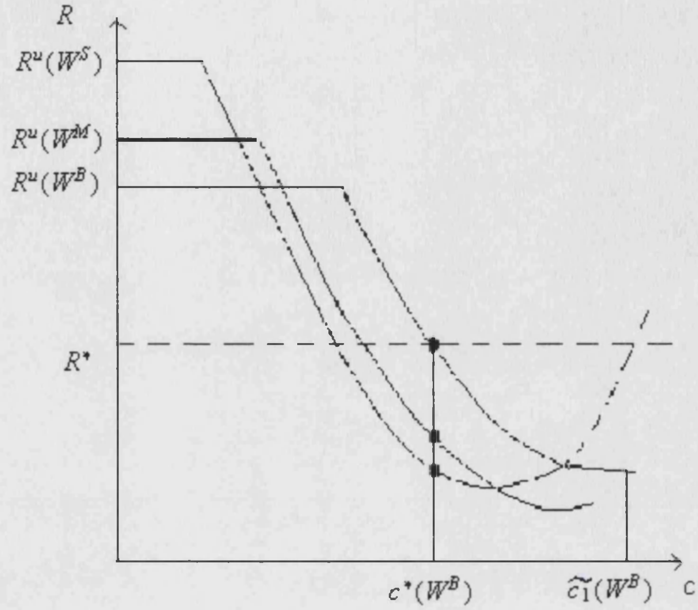

Figure 4-13: Distribution of risks when capital requirements are used to force the biggest bank to the social optimum.

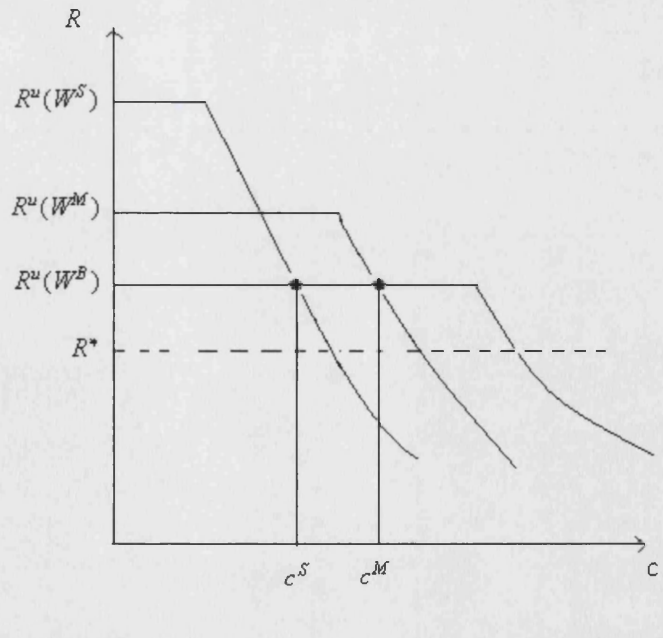

Figure 4-14: Distribution of risks when capital regulation is designed to make the smaller banks converge on the unregulated risk choice of the biggest one.

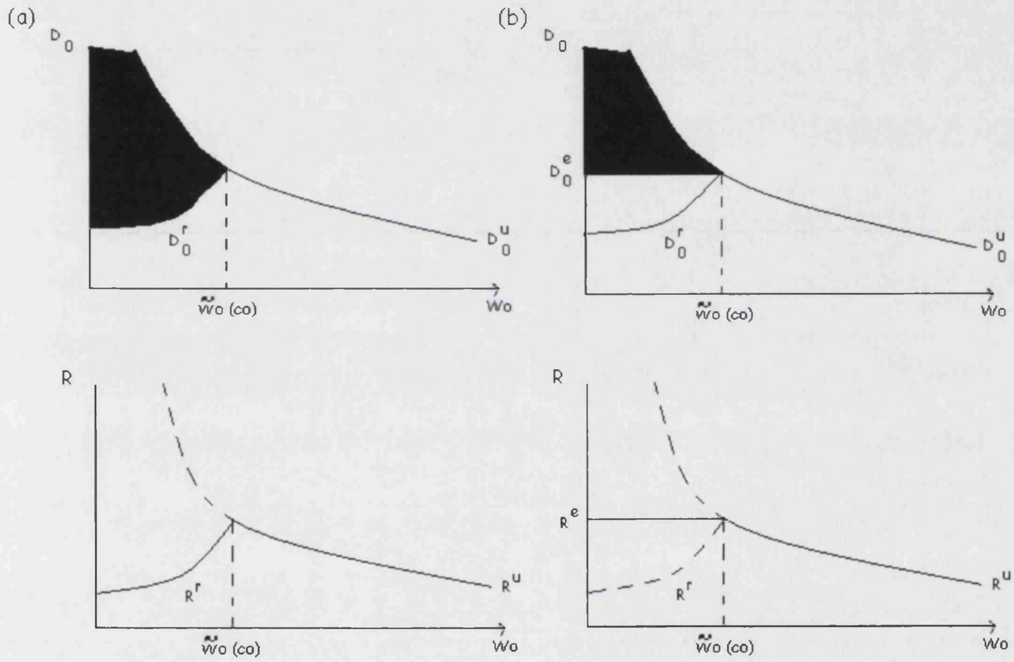

Figure 4-15: Loss in financial intermediation under (a) a first period binding capital requirement and (b) a binding minimum equity requirement.

Bibliography

[1] Acharya, S. and J-F. Dreyfus. 1989. *"Optimal Bank Reorganization Policies and the Pricing of Federal Deposit Insurance"*. Journal of Finance 44(5), 1313-1333.

[2] Acharya, V. 2001. *"A Theory of Systemic Risk and Design of Prudential Bank Regulation"*. London Business School, Working Paper Series. Available at:http://faculty.london.edu/vacharya/pdf/correlation.pdf.

[3] Acharya, V., I. Hasan and A. Saunders. 2006. *"Should Banks be Diversified? Evidence from Individual Banks Loan Portfolios."* Journal of Business 79(3), 1355-1412.

[4] Acharya, V. and T. Yorulmazer. 2007a. *"Cash-in-the-market Pricing and Optimal Resolution of Bank Failures."* Bank of England Working Papers 328.

[5] Acharya, V. and T. Yorulmazer. 2007b. *"Too Many to Fail - an Analysis of Time Inconsistency in Bank Closure Policies."* Journal of Financial Intermediation 16(1), 1-31.

[6] Acharya, V. and T. Yorulmazer. 2008. *"Information Contagion and Bank Herding."* Journal of Money, Credit and Banking 40(1), 215 - 231.

[7] Akhigbe, A. and J. Madura. 2001. *"Why do Contagion Effects Vary Among Bank Failures?"*. Journal of Banking & Finance 25(4), 657-680.

[8] Allen, F. 2000. *"Financial Crises"*. Global Financial Association, Keynote Address. Chicago, Illinois. Available at http://condor.depaul.edu/~afatemi1/globalfinanceassociation0400.pdf.

[9] Allen, F. and D. Gale. 1995. *"A Welfare Comparison of Intermediaries and Financial Markets in Germany and the U.S."* European Economic Review 39(2), 179-209.

[10] Allen, F. and D. Gale. 1998. *"Optimal Financial Crises"*. Journal of Finance, Papers and Proceedings Issue 53(4), 1245-1284.

[11] Allen, F. and D. Gale. 2000a. *"Bubbles and Crisis"*. The Economic Journal 110 (460), 236-255.

[12] Allen, F. and D. Gale. 2000b. *"Financial Contagion"*. Journal of Political Economy 108(1), 1-33.

[13] Allen, F. and D. Gale. 2001. *"Comparative Financial Systems: A Survey"*. In *Financial intermediation,* A. Boot, S. Bhattacharya and A. Thakor (Eds.) Oxford University Press. Available at: http://fic.wharton.upenn.edu/fic/papers/01/0115.pdf.

[14] Alonso, I. 1996. *"On Avoiding Bank Runs"*. Journal of Monetary Economics 37(1), 73-87.

[15] Bagehot, W. 1873. *"Lombard Street: A Description of the Money Market"*. London: Henry S. King and Co.

[16] Bank of England. November 2000. *"Oversight of Payment Systems"*. Available at http://www.bankofengland.co.uk/publications/psor/ops.pdf.

[17] Baring, F. 1797. *"Observations on the Establishment of the Bank of England, and on the Paper Circulation of the Country"*. London 1797, Reed. In Selected Tracts 1694-1804, Ed. Gregg International Publisher Ltd, Bank of England, London, 1968.

[18] Basel Committee on Banking Supervision. 2001. *"Pillar 3: Market Discipline Consultative Document"*. Bank for International Settlements.

[19] Basel Committee on Banking Supervision. 2002. *"Supervisory Guidance on Dealing with Weak Banks"*. Bank for International Settlements.

[20] Basel Committee on Banking Supervision. 2005. *"International Convergence of Capital Measurement and Capital Standards"*. Bank for International Settlements.

[21] Basel Committee on Banking Supervision. 2006. *"International Convergence of Capital Measurement and Capital Standards. A Revised Framework Comprehensive Version"*. Bank for International Settlements.

[22] Beck, T., R. Levine and N. Loayza. 1999. *"Finance and the Source of Growth"*. World Bank Working Papers 2057.

[23] Beck, T. and R. Levine. 2001. *"Stock Markets, Banks, and Growth: Correlation or Causality?"*. World Bank Working Papers 2670.

[24] Besanko, D. and G. Kanatas. 1993. *"Credit Market Equilibrium with Bank Monitoring and Moral Hazard".* Review of Financial Studies, 213-232.

[25] Berger, A., S. Davies and M. Flannery. 1999. *"Comparing Market and Supervisory Assessments of Bank Performance: Who Knows What When?".* Journal of Money, Credit and Banking 32(3), 641-667.

[26] Bernake, B. 1983. *"Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression".* American Economic Review 73(3), 257-276.

[27] Bhattacharya, S. and D. Gale. 1987. *"Preference Shocks, Liquidity and Central Bank Policy".* In *New approaches to monetary economics.* W. Barnett and K. Singleton (Eds.) New York: Cambridge University Press, 69-88.

[28] Bhattacharya, S., M. Plank, G. Strobl and J. Zechner. 2002. *"Bank Capital Regulation with Random Audits".* Journal of Economic Dynamics and Control 26(7-8), 1301-1321.

[29] Bliss R. and M. Flannery. 2000. *"Market Discipline in the Governance of U.S. Bank Holding Companies: Monitoring vs. Influencing".* Federal Reserve Bank of Chicago, Working Papers 2000-3.

[30] Bliss R. 2001. *"Market Discipline and Subordinated Debt: A Review of Some Salient Issues".* Economic Perspectives, Federal Reserve Bank of Chicago, 24-45.

[31] Blum, J. 1999. *"Do Capital Adequacy Requirements Reduce Risk in Banking?".* Journal of Banking and Finance 23(5), 755-771.

[32] Blum, J. 2000. *"The Limits of Market Discipline in Reducing Bank's Risk Taking".* Working Papers 00.08, Study Center Gerzensee, Swiss National Bank.

[33] Bolton, P. and X. Freixas.1997. *"Direct Bond Financing, Financial Intermediation, and Investment".* Mimeo. Universitat Pompeu Fabra.

[34] Bond, E. and K. Crocker. 1993. *"Bank Capitalization, Deposit Insurance, and Risk Categorization".* Journal of Risk and Insurance 60(4), 547-569.

[35] Boot A. and S. Greenbaum. 1993. *"Bank Regulation, Reputation and Rents: Theory and Policy Implications".* In *Capital markets and financial intermediation.* C. Mayer and X. Vives (Eds.) CEPR: Cambridge University Press, 262-285.

[36] Boot, A. and A. Thakor. 1993. *"Security Design".* Journal of Finance 48(4), 1349-1378.

[37] Bryant, J. 1980. *"A Model of Reserves, Bank Runs and Deposit Insurance"*. Journal of Banking and Finance 4(4), 335-344.

[38] Calomiris, C. 1995. *"Financial Fragility: Issues and Policy Implications"*. Journal of Financial Services Research 9(3-4), 241-257.

[39] Calomiris, C. 1996. *"Building an Incentive-Compatible Safety Net: Special Problems for Developing Countries"*. Mimeo, Columbia University.

[40] Calomiris, C. 1997. *"The Postmodern Bank Safety Net. Lessons from Developed and Developing Economies"*. The American Enterprise Institute Press. Washington, D.C. Available at: http://www.aei.org/docLib/20021130_71007.pdf.

[41] Calvo, G. 1998. *"Varieties of Capital-Market Crises"*. In *The debt burden and its consequences for monetary policy.* G. Calvo and M. King (Eds.) London: MacMillan Press.

[42] Calvo, G. 2000. *"Balance of Payment Crises in Emerging Markets: Large Capital Inflows and Sovereign Governments"*. In *Currency crisis.* Paul Krugman, ed. The University of Chicago Press. Series: (NBER-C).

[43] Campbell, T. and W. Kracaw. 1980. *"Information Production, Market Signalling, and the Theory of Financial Intermediation"*. Journal of Finance 35(4), 863-882.

[44] Carlsson, H. and E. van Damme. 1993. *"Global Games and Equilibrium Selection"*. Econometrica 61(5), 989-1018.

[45] Chan, Y-S., S. Greenbaum and A. Thakor. 1992. *"Is fairly priced deposit insurance possible?"*. Journal of Finance 47(1), 227-245.

[46] Chang, R. and A. Velasco. 1998. *"Financial Fragility and the Exchange Rate Regime"*. NBER Working Papers 6469. Available at http://papers.nber.org/papers/w6469.pdf.

[47] Chang, R. and A. Velasco. 1999. *"Liquidity Crises in Emerging Markets: Theory and Policy"*. NBER Working Papers 7272. Available at: http://www.nber.org/papers/w7272.

[48] Chari. V.V. 1989. *"Banking Without Deposit Insurance or Bank Panics: Lessons from a Model of the U.S. National Banking System"*. Federal Reserve of Minneapolis. Quarterly Review 13, 3-19.

[49] Chari, V.V. and R. Jagannathan. 1988. *"Banking Panics, Information, and Rational Expectations Equilibrium"*. Journal of Finance 43(3), 749-761.

[50] Cooper, R. and T. Ross. 1998. *"Bank Runs: Liquidity Costs and Investment Distortions"*. Journal of Monetary Economics 41(1), 27-38.

[51] Corsetti, G., A. Dasgupta, S. Morris and H.S. Shin. 2004. *"Does One Soros Make a Difference? A Theory of Currency Crises with Large and Small Traders"*. Review of Economic Studies 71(1), 87-113.

[52] Corsetti, G., B. Guimaraes and N. Roubini. 2003. *"International Lending of Last Resort and Moral Hazard. A Model of IMF's Catalytic Finance"*. NBER Working Papers 10125.

[53] Covitz, D., D. Hancock, M. Kwast. 2004. *"Market Discipline in Banking Reconsidered: The Roles of Funding Manager Decisions and Deposit Insurance Reform"*. Finance and Economics Discussion Series 2004-53, Federal Reserve System (U.S.).

[54] Curry, T. and L. Shibut. 2000. *"The Cost of the Savings and Loan Crisis: Truth and Consequences"*. FDIC Banking Review 13(2), 26-35.

[55] Daníelsson, J. 2000. *"The Emperor has no Clothes: Limits to Risk Modelling"*. FMG Special Papers 126. London School of Economics. Available at: http://fmg.lse.ac.uk/pdfs/sp126.pdf.

[56] Dasgupta, A. 2002. *"Financial Contagion through Capital Connections: A Model of the Origin and Spread of Bank Panics"*. FMG Discussion Papers 436. London School of Economics. Available at: http://fmg.lse.ac.uk/pdfs/dp436.pdf.

[57] Decamps J.P., J.C. Rochet and B. Roger. 2004. *"The Three Pillars of Basel II: Optimizing the Mix"*. Journal of Financial Intermediation 13(2), 132-155.

[58] Demirguç-Kunt, A. and E. Detragiache. 1999. *"Does Deposit Insurance Increase Banking System Stability? An Empirical Investigation"*. Policy Research Working Paper Series 2247, The World Bank.

[59] Demirguç-Kunt, A and T. Sobaci. 2000. *"Deposit Insurance Around the World: A Data Base"*. Mimeo, The World Bank.

[60] Demirguç-Kunt, A. and H. Huizinga. 2004. *"Market discipline and deposit insurance"*. Journal of Monetary Economics 51(2), 375-399.

[61] Dewatripont, M. and J. Tirole. 1993. *"Efficient Governance Structure: Implications for Banking Regulation"*. In *Capital Markets and Financial intermediation.* C. Mayer and X. Vives (Eds.) Cambridge University Press, 12-35.

[62] Dial, J.B. 1996. *"Plans for Derivatives Regulation"*. In *Fourth International Conference on Derivatives Regulation*, (October) University of London. Available at: http//www.cftc.gov/opa/speeches/opadial-56.htm.

[63] Diamond, D. 1984. *"Financial Intermediation and Delegated Monitoring"*. Review of Economic Studies 51(3), 393-414.

[64] Diamond, D. 1991. *"Monitoring and Reputation: the Choice between Bank Loans and Privately Placed Debt"*. Journal of Political Economy 99(4), 698-721.

[65] Diamond, D. and P.H. Dybvig. 1983. *"Banks Runs, Deposit Insurance, and Liquidity"*. Journal of Political Economy 91(3), 401-419.

[66] Diamond, D. and R. Rajan. 1998. *"Liquidity risk, Liquidity Creation and Financial Fragility: A Theory of Banking"*. Proceedings, Federal Reserve Bank of San Francisco, issue Sep.

[67] Echeverria, C. and L. Opazo. 2007. *"Prominent Features of Development in the Credit Derivatives Market"*. Financial Stability Report, First Half, Central Bank of Chile.

[68] Eisenbeis, R.A. 1997. *"Bank Deposits and Credit As Sources of Systemic Risk"*. Economic Review, Federal Reserve Bank of Atlanta, 1-16.

[69] Evanoff D. and L. Wall. 2000. *"Subordinated Debt and Bank Capital Reform"*. Federal Reserve Bank of Chicago, Working Paper Series 00-07.

[70] Fama, E. 1980. *"Banking in the Theory of Finance"*. Journal of Monetary Economics 6(1), 39-57.

[71] Flannery, M. 1989. *"Capital Regulation and Insured Bank's Choice of Individual Loan Default Risks"*. Journal of Monetary Economics 24(2), 235-258.

[72] Flannery, M. 1994. *"Debt Maturity and the Deadweight Cost of Leverage: Optimally Financing Banking Firms"*. American Economic Review 84(1), 320-31.

[73] Flannery, M. and S. Nikolova. 2003. *"Market Discipline of U.S. Financial Firms: Recent Evidence and Research Issues"*. Forthcoming in Market discipline across

countries and industries, Hunter W., G. Kaufman, C. Borio and K. Tsatsanoris (Eds.) MIT Press.

[74] Folkerts-Landau, D. and C.J. Lindgren. 1998. *"Towards a Framework for Financial Stability"*. IMF World Economic and Financial Surveys.

[75] Frankel, D., Morris, S. and A. Pauzner. 2002. *"Equilibrium Selection in Global Games with Strategic Complementarities"*. Cowles Foundation Discussion Papers 1336. Available at: http://papers.ssrn.com/abstract=291411.

[76] Fratzscher, M. 2000. *"On Currency Crises and Contagion"*. Peterson Institute for International Economics 00-9.

[77] Freixas, X. 1999. *"Optimal Bail Out Policy, Conditionality and Creative Ambiguity"*. FMG Discussion Papers 327. London School of Economics. Available at: http://fmg.lse.ac.uk/pdfs/dp327.pdf.

[78] Freixas, X., Giannini, C. and F. Soussa. 1999. *"Lender of Last Resort: a Review of the Literature"*. Financial Stability Review (November), Bank of England, 151-167.

[79] Freixas, X. and C. Holthausen. 2001. *"Interbank Market Integration under Asymmetric Information"*. European Central Bank, Working Papers 74.

[80] Freixas, F. and B. Parigi. 1998. *"Contagion and Efficiency in Gross and Net Interbank Payment Systems"*. Journal of Financial Intermediation 7(1), 3-31.

[81] Freixas, F., B. Parigi and J.C. Rochet. 2000. *"Systemic Risk, Interbank Relations, and Liquidity Provision by the Central Bank"*. Journal of Money, Credit, and Banking 32(3), 611-638.

[82] Freixas, X. and J.C. Rochet. 1997. *"Microeconomics of Banking"*. The MIT Press.

[83] Freixas, X. and A.M. Santomero. 2002. *"An Overall Perspective on Banking Regulation"*. Federal Reserve Bank of Philadelphia, Working Papers 02-1. Available at: http://www.phil.frb.org/files/wps/2002/wp02-1.pdf.

[84] Friedman, M. and A. Schwartz. 1963. *"A Monetary History of the United States, 1867-1960"*. Princeton, N.J.: Princeton University Press.

[85] Furfine, C. 2001. *"Bank Portfolio Allocation: The Impact of Capital Requirements, Regulation Monitoring, and Economic Conditions"*. Journal of Financial Services Research 20(1), 33-56.

[86] Furlong , F. 1988. *"Changes in Bank Risk-Taking"*. Federal Reserve Bank of San Francisco Economic Review, Spring, 45-56.

[87] Furlong , F. and M.C. Keeley. 1989. *"Capital Regulation and Bank Risk Taking: A Note"*. Journal of Banking and Finance 13(6), 883-91.

[88] Gale, D. 2000. *"Understanding Financial Crisis"*. Later published with F. Allen, Oxford University Press (2007).

[89] Garcia, G.G.H. 2000. *"Deposit Insurance. Actual and Good Practices"*. IMF Occasional Papers 197.

[90] Giannetti, M. 2001. *"Bank-Firm Relationships and Contagious Banking Crisis"*. Available at: http://www.cepr.org/meets/wkcn/5/565/papers/Giannetti.pdf.

[91] Glick, R. and M. Hutchinson. 1999. *"Banking and Currency Crises: How Common are Twins?"*. Pacific Basin Working Paper Series 99-07, Federal Reserve Bank of San Francisco.

[92] Goldstein, I. 2005. *"Strategic Complementarities and the Twin Crises"*. Economic Journal 115(503), 368-390.

[93] Goldstein, I. and A. Pauzner. 2000. *"Demand Deposit Contracts and the Probability of Bank Runs"*. Available at: http://www.ifk-cfs.de/papers/bankruns.pdf. Later published in Journal of Finance 60(3), 1293-1327 (2005).

[94] Goldstein, I. and A. Pauzner. 2004. *"Contagion of Self-Fulfilling Financial Crisis due to Diversification of Investment Portfolios"*. Available at: http://www.tau.ac.il/~pauzner/papers/contagion.pdf.

[95] Gondat-Larralde, C. and E. Nier. 2006. *"Switching Costs in the Market for Personal Current Accounts: Some Evidence for the United Kingdom"*. Bank of England Working Papers 292.

[96] Gonzalez-Hermosillo, B. 1996. *"Banking Sector Fragility and Systemic Sources of Fragility"*. IMF Working Papers 96/12.

[97] Goodhart, C. 1995. *"Why do Banks need a Central Bank?"*. In *The central bank and the financial system.* Cambridge, Mass.: MIT Press.

[98] Goodhart, C., P. Hartman, D. Llewellyn and L. Rojas-Suarez. 1997. *"Financial Regulation: Why, How and Where Now?"*. Central Bank Governor's Meeting, Bank of England.

[99] Goodhart, C. and D. Schoenmaker. 1995. *"Should the Functions of Monetary Policy and Banking Supervision be Separated?"*. Oxford Economic Papers 47(4), 539-560.

[100] Goodfriend, M. and R. King. 1988. *"Financial Deregulation, Monetary Policy and Central Banking"*. In *Restructuring banking and financial services in America.* W. Haraf and R.M. Kushmeider (Eds.) AEI Studies 481. Lanham, Md.:UPA.

[101] Gorton, G. 1988 *"Banking Panics and Business Cycles"*. Oxford Economics Papers 40(4), 751-781.

[102] Gorton, G. and G. Pennacchi. 1990. *"Financial Intermediaries and Liquidity Creation"*. Journal of Finance 45(1), 49-71.

[103] Gray, S. and J. Place. 1999. *"Financial Derivatives"*. Handbooks in Central Banking 17 Centre for Central Banking Studies, Bank of England.

[104] Greenspan, A. 2001. *"The Financial Safety Net"*. Remarks to the 37th Annual Conference on Bank Structure and Competition of the Federal Reserve Bank of Chicago, Chicago, IL (May 10).

[105] Grossman, S. and J. Stiglitz. 1980. *"On the Impossibility of Informationally Efficient Markets"*. American Economic Review 70(3), 393-408.

[106] Gurley, J. and E. Shaw.1960. *"Money in the Theory of Finance"*. Washington: Brookings Institution.

[107] Hellmann, T., K. Murdock and J. Stiglitz. 2000. *"Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?"*. American Economic Review 90(1), 147-165.

[108] Hoelscher, D.S. and M. Quintyn. 2003. *"Managing Systemic Banking Crises"*. IMF Occasional Papers 224.

[109] Hoggarth, G., P. Jackson and E. Nier. 2005. *"Banking Crisis and the Design of Safety Nets"*. Journal of Banking and Finance 29(1), 143-159.

[110] Hortala-Vallve, R. 2002. *"A simple model of bank disclosure and subordinated debt"*. Mimeo Bank of England.

[111] Hovaguimian, C. 2003. *"The Catalytic Effect of IMF Lending: A Critical Review"*. Financial Stability Review. Bank of England, December, 160-169.

[112] Huang, H. and C. Xu. 2000. *"Financial Institutions, Financial Contagion and Financial Crisis"*. William Davidson Institute Working Papers Series 316, University of Michigan.

[113] Hunter, W.C. and D. Marshall. 1999. *"Thoughts on Financial Derivatives, Systemic Risk, and Central Banking: A Review of Some Recent Developments"*. Federal Reserve Bank of Chicago Working Paper Series 99-20.

[114] Jacklin, C.J. and S. Bhattacharya. 1988. *"Distinguishing Panics and Information-Based Bank Runs: Welfare and Policy Implications"*. Journal of Political Economy 96(3), 568-592.

[115] Jackson, P. 1999. *"Capital Requirements and Bank Behavior: The Impact of the Basel Accord"*. Bank for International Settlements, Basel Committee on Banking Supervision Working Papers 1.

[116] Jameson R., Erisk 2003, *Case Studies.* Available at: http://www.erisk.com/Learning/CaseStudies.asp.

[117] Jeanne, O. and C. Wyplosz. 2001. *"The International Lender of Last Resort: How Large is Large Enough?"*. IMF Working Papers 01/76. Available at: http://www.imf.org/external/pubs/ft/wp/2001/wp0176.pdf.

[118] Jones, D. 2000. *"Emerging Problems with the Basel Capital Accord: Regulatory Capital Arbitrage and Related Issues"*. Journal of Banking and Finance 24 (1-2), 35-58.

[119] Kahane, Y. 1977. *"Capital Adequacy and the Regulation of Financial Intermediaries"*. Journal of Banking and Finance 1(2), 207–218.

[120] Kahn, C. and W. Roberds. 2001. *"Real-time Gross Settlement and the Costs of Immediacy"*. Journal of Monetary Economics 47(2), 299-319.

[121] Kaminsky, G. and C. Reinhart. 1996. *"The Twin Crises: The Causes of Banking and Balance of Payments Problems"*. International Finance Discussion Papers 544, Board of Governors of the Federal Reserve System.

[122] Kane, E. 1995. *"Three Paradigms for the Role of Capitalization Requirements in Insured Financial Institutions"*. Journal of Banking and Finance 19(3-4), 431-459.

[123] Kang, J.K. and R. Stulz. 2000. *"Do Banking Shocks Affect Borrowing Firm Performance? An Analysis of the Japanese Experience"*. Journal of Business 73(1), 1-23.

[124] Kareken, J. and N. Wallace. 1978. *"Deposit Insurance and Bank Regulation: A Partial-Equilibrium Exposition"*. Journal of Business 51(3), 413–438.

[125] Ketcha, N. Jr. 1999. *"Deposit Insurance System Design and Considerations"*. BIS Policy Papers 7, 221–239.

[126] Kim, D. and A. Santomero. 1988. *"Risk in Banking and Capital Regulation"*. Journal of Finance 43(5), 1219-1233.

[127] Koehn, M. and A. Santomero. 1980. *"Regulation of Bank Capital and Portfolio Risk"*. Journal of Finance 35(5), 1235-1244.

[128] Krugman, P. 1979. *"A Model of Balance of Payments Crises"*. Journal of Money, Credit, and Banking 11(3), 311-325.

[129] Kupiec, P. and J. O'Brien. 1997. *"The Pre-Commitment Approach: Using Incentives to set Market Risk Capital Requirements"*. Board of Governors of the Federal Reserve System, Finance and Economics Discussions Series 1997/14.

[130] Kwan, S. December 13, 2002. FRBSF Economic Letter 2002-36.

[131] Kwast, M., D. Covitz, D. Hancock, J. Houpt, D. Adkins, N. Barger, B. Bouchard, J. Connolly, T. Brady, W. English, D. Evanoff and L. Wall. 1999. *"Using Sub-debt as an Instrument of Market Discipline"*. Report of a study group on subordinated notes and debentures, Board of Governors of the Federal Reserve System (M. Kwast, chair), Staff Study 172.

[132] Laporta, R., F. Lopez-de-Silanes and A. Schleifer. 1998. *"Law and Finance"*. Journal of Political Economics 106(6), 1113-1155.

[133] Latter, T. 1997. *"The Causes and Management of Banking Crisis"*. Handbook in Central Banking 12, Centre for Central Banks Studies, Bank of England.

[134] Levine, R. 1997 *"Financial Development and Economic Growth: Views and Agenda"*. Journal of Economic Literature 35(2), 688-726.

[135] Levy-Yeyati E., M. Martinez and S. Schmukler. 2004. *"Market Discipline under Systemic Risk: Evidence from Bank Runs in Emerging Economies"*. World Bank Policy Research Working Papers 3440. Available at SSRN: http://ssrn.com/abstract=625318.

[136] Merton, R. 1977. *"An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantees"*. Journal of Banking and Finance 1(1), 3-11.

[137] Miller, V. 1996. *"Speculative Currency Attacks with Endogenously Induced Commercial Bank Crises"*. Journal of International Money and Finance 15(3), 385-403.

[138] Miller, V. 1999. *"The Timing and size of Bank-Financed Speculative Attacks"*. Journal of International Money and Finance 18(3), 459-470.

[139] Miron, J.A. 1986. *"Financial Panics, the Seasonality of the Nominal Interest Rate, and the Founding of the Fed"*. American Economic Review 76(1), 125-140.

[140] Mishkin, F. 1992. *"The Economics of Money, Banking and Financial Markets"*. London, Scott, Foresman.

[141] Morris, S. and H.S. Shin. 2002. *"Global Games: Theory and Applications"*. Eighth World Congress of the Econometric Society.

[142] Morris, S. and H.S. Shin. 2003. *"Catalytic Finance: When does it Work?"*. Yale University, Cowles Foundation Discussion Papers 1400.

[143] Nascimento, A. J. 2000 *"Interbank Markets: Liquidity Provision and Systemic Risks"*. EFMA 2001 Lugano Meetings. Available at: http://ssrn.com/abstract=271074.

[144] Niu, J. 2008. *"Can subordinated debt constrain banks' risk taking?"*. Journal of Banking and Finance 32(6), 1110-1119.

[145] Obstfeld, M. 1994. *"The Logic of Currency Crises"*. Banque de France, Cahiers Économiques et Monétaires 43, 189-213.

[146] Obstfeld, M. 1996. *"Models of Currency Crises with self-fulfilling Features"*. European Economic Review 40(3-5), 1037-1047.

[147] Ongena, S., D.C. Smith and D. Michalsen. 2000. *"Distressed Relationships: Lessons from the Norwegian Banking Crises (1988-1991)"*. Center for Economic Research Discussion Paper 2000-13.

[148] Postlewaite, A. and X. Vives. 1987 *"Bank Runs as an Equilibrium Phenomenon"*. Journal of Political Economy 95(3), 485-491.

[149] Rajan, R.G. 1992. *"Insiders and Outsiders: the Choice between Informed and Arm's-length Debt"*. Journal of Finance 47(4), 1367-1400.

[150] Repullo, R. 2000. *"Who Should Act as Lender of Last Resort? An Incomplete Contract Model"*. Journal of Money, Credit and Banking 32(3), 580-605.

[151] Rochet, J.C. 1992. *"Capital Requirements and the Behaviour of Commercial Banks"*. European Economic Review 36(5), 1137-1170.

[152] Rochet, J.C. 1999. *"Solvency Regulation and the Management of Banking Risks"*. European Economic Review 43(4-6), 981-990.

[153] Rochet, J.C. 2004. *"Rebalancing the Three Pillars of Basel II"*. FRBNY Economic Policy Review 10(2), 7-21.

[154] Rochet, J.C. and J. Tirole. 1996. *"Interbank Lending and Systemic Risk"*. Journal of Money, Credit and Banking 28(4), 733-762.

[155] Rochet, J-C and X. Vives. 2002. *"Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?"*. CEPR Discussion Papers 3233.

[156] Rojas-Suarez, L. and S. Weisbrod. 1995. *"Financial Fragilities in Latin America: The 1980s and 1990s"*. IMF Occasional Papers 132.

[157] Rule, D. 2001a. *"The Credit Derivatives Market: its Development and Possible Implications for Financial Stability"*. Bank of England, Financial Stability Review (June), 117-140.

[158] Rule, D. 2001b. *"Risk Transfer Between Banks, Insurance Companies and Capital Markets: an Overview"*. Bank of England, Financial Stability Review (December), 137-159.

[159] Santos, J. 1999. *"Bank Capital and Equity Investment Regulations"*. Journal of Banking and Finance 23(7), 1095–1120.

[160] Santos, J. 2000. *"Bank Capital Regulation in Contemporary Banking Theory: A Review of the Literature"*. BIS Working Papers 90, Bank for International Settlements.

[161] Schoenmaker, D. 1996. *"Contagion Risk in Banking"*. FMG Discussion Papers 239, London School of Economics.

[162] Sharpe, W. F. 1978. *"Bank Capital Adequacy, Deposit Insurance and Security Values"*. Journal of Financial and Quantitative Analysis 13(4), 701–718.

[163] Sheldon, G. 1996. *"Capital Adequacy Rules and the Risk-Seeking Behavior of Banks: A Firm-Level Analysis"*. Swiss Journal of Economics and Statistics 132(4), 709-734.

[164] Sironi, A. 2003. *"Testing for market discipline in the European banking industry: Evidence from subordinated debt issues"*. Journal of Money and Banking 35(3), 443-472.

[165] Slovin, M., M. Sushka and J. Polonchek. 1993. *"The Value of Bank Durability: Borrowers as Banks Stakeholders"*. Journal of Finance 48(1), 247-266.

[166] Stiglitz, J.E. 1992. *"S&L Bailout"*. In *The Reform of Federal Deposit Insurance.* J.R. Barth and R.D. Brunbaugh Jr. (Eds.), New York, HarperBusiness.

[167] Stiglitz, J.E. 1994. *"The Role of the State in Financial Markets"*. Proceedings of the World Bank Annual Conference on Development Economics. Washington, D.C.

[168] Stiglitz, J.E. 2002. *"Globalization and its Discontents"*. London: Allen Lane.

[169] Thakor, A.V. 1996. *"Capital Requirements, Monetary Policy, and Aggregate Bank Lending"*. Journal of Finance. 51(1), 279–324.

[170] Thornton, H. 1802. *"An Enquiry into the Nature and Effects of the Paper Credit of Great Britain"*. London, Hatchard.

[171] Van Rijckeghem, C. and B.Weder. 2001. *"Sources of Contagion: is it Finance or Trade?"*. Journal of International Economics 54(2), 293-308.

[172] Velasco, A. 1987. *"Financial Crises and Balance of Payment Crises: A Simple Model of the Southern Cone Experience"*. Journal of Development Economics 27(1-2), 263-283.

[173] Von Thadden, E.L. 1995. *"Long-Term Contracts, Short-Term Investment, and Monitoring"*. Review of Economic Studies 62(4), 557-575.

[174] Von Thadden, E.L. 1998. *"Optimal Liquidity Provision and Dynamic Incentive Compatibility"*. Journal of Financial Intermediation 7(2), 177-197.

[175] Williamson, S. 1995. *"Discount Window Lending and Deposit Insurance"*. Working Papers 95-01, University of Iowa, Department of Economics. Available at: http://econwpa.wustl.edu:80/eps/mac/papers/9504/9504001.pdf.