Essays on Financial Stability

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Stephan Luck

aus Leverkusen

Bonn 2016

Dekan: Prof. Dr. Rainer Hüttemann
Erstreferent: Prof. Martin F. Hellwig, PhD
Zweitreferent: Prof. Dr. Hendrik Hakenes

Tag der mündlichen Prüfung: 06.11.2015

Diese Dissertation ist elektronisch publiziert auf dem Hochschulschriftenserver der ULB Bonn: http://hss.ulb.uni-bonn.de/diss_online

Working on this thesis has been a long journey on which I have met many interesting people. Many individuals and institutions deserve my gratitude for their support and encouragement. I would like to use these pages to thank them.

First, and foremost, I would like to thank my supervisor Martin Hellwig. His unconditional support, his intellectual guidance and his eye for interesting questions have clearly shaped me. I feel extremely lucky and privileged to be his student.

My friend and co-author, Paul Schempp, also deserves my deepest gratitude. Over the years of working together, we have become close friends and exchanged views not only on questions of economics but on a broad number of topics. This friendship and collaboration has been a gift that I would like to thank you for!

I would also like to particularly thank my advisors at my time at Princeton University, Markus Brunnermeier and Stephen Morris, who have been extremely supportive at later stages of the dissertation. Moreover, I am deeply indebted to the advise and help of Hendrik Hakenes for his support over the years. Finally, all chapters of this thesis have benefited from various discussions and comments. I would like to particularly thank Tobias Adrian, Tobias Berg, Christoph Bertsch, Felix Bierbrauer, Patrick Bolton, Florian Buck, Jean-Edouard Colliard, Brian Cooper, Olivier Darmouni, Christoph Engel, Maryam Farboodi, Falko Fecht, Douglas Gale, Valentin Haddad, Rainer Haselmann, Christian Hellwig, Zongbo Huang, Nobuhiro Kiyotaki, Sam Langfield, Adrien Matray, Konstantin Milbradt, John Moore, Sebastian Pfeil, Jean-Charles Rochet, Eva Schliephake, Isabel Schnabel, Anatoli Segura, and Ansgar Walther. Michael Evers deserves special thanks for advising me in the early stages of my graduate studies.

Throughout my graduate studies in the Bonn Graduate School of Economics, I received funding from Deutsche Forschungsgemeinschaft, as well as from the Max Planck Institute for Research on Collective Goods and the Alexander von Humboldt Foundation. I would like to thank the BGSE and the MPI for offering such a great learning and research environment. Also, I would like to thank my fellow PhD students as well as my colleagues, in particular Benjamin Enke, Andreas Kleiner, Ralph Lütticke, and Frederik Schwerter for this pleasant time!

I am also thankful for the support by the administration of the BGSE and the MPI. Monika Stimpson and Heidi Morgenstern deserve special thanks for helping me out on various administrative matters!

Last but by no means least, I would like to thank my beloved parents and my beloved sister. Knowing that my family always stands behind me is an invaluable gift to my life.

Contents

List of Figures

List of Abbreviations

ABCP Asset-Backed Commercial Papers
ABS Asset-Backed Securities
ASF Available Stable Funding
DGS Deposit Guarantee Scheme
DI Deposit Insurance
FCIC Financial Crisis Inquiry Commission
FSB Financial Stability Board
HQLA High Quality Liquid Assets
LCR Liquidity Coverage Ratio
LoLR Lender of Last Resort
MBS Mortgage-Backed Securities
MMLR Market Maker of Last Resort
MMF Money Market (Mutual) Fund
NAV Net Asset Value
NSFR Net Stable Funding Ratio
RMBS Residential Mortgage-Backed Securities
RSF Required Stable Funding
SIV Structured Investment Vehicle
SoC Suspension of Convertibility
SRM Single Resolution Mechanism
SSM Single Supervisory Mechanism

Introduction

This thesis consists of three chapters that analyze the fragility of financial institutions. All chapters are based on joint work with my friend and colleague Paul Schempp. The analysis is concerned with self-fulfilling crises that are associated with maturity transformation and liquidity provision by financial intermediaries. The need for such services arises because, on the one hand, investors such as consumers or firms have a demand for liquid assets and, on the other hand, investment projects require stable long-term funding. Banks are institutions that can efficiently intermediate in such an environment. However, when illiquid long-term projects are financed with short-term debt such as demand-deposits, strategic complementarity may arise between those agents that hold short-term claims. Thus, maturity mismatch makes banks potentially prone to self-fulfilling liquidity crises, i.e., panic-based bank runs or rollover freezes may become possible. The important questions that arise in this context are: Under which conditions does the risk of self-fulfilling liquidity crises arise? And, how can such fragility be best dealt with?

In this thesis, we show that the government has a distinct role in ensuring the functioning of efficient maturity transformation and liquidity provision (Chapter 1). We also show that if a single government's fiscal power is limited, supranational agreements can help to mitigate this limitation (Chapter 2). Finally, we show that in the presence of regulatory arbitrage, certain types of governmental interventions become ineffective altogether (Chapter 3).

Chapter 1 is based on Luck and Schempp (2014b) and discusses the optimal provision of liquidity. It asks whether financial intermediaries can optimally provide liquidity, or whether the government has a role in creating liquidity by supplying government securities. We discuss a model in which intermediaries optimally manage liquidity with outside rather than inside liquidity: instead of holding liquid real assets that can be used at will, banks sell claims on long-term projects to investors when liquid funds are needed. While increasing efficiency, liquidity management with private outside liquidity is associated with a rollover risk. This rollover risk either keeps intermediaries from providing liquidity optimally, or it makes the economy inherently fragile. In contrast to privately produced claims, government bonds are not associated with coordination problems unless there is a prospect of the government defaulting. Therefore, efficiency and stability can be enhanced if liquidity management relies on public outside liquidity.

The main results of the first chapter are derived under the assumption that a government cannot default because it can commit future liquidity via taxation. It is thus the only institution that can credibly promise to provide liquidity in the future. However, this ability depends on the assumption that the government has access to a sufficiently large and stable tax base. In the next chapter, we assume that the tax base is endogenous and depends to some degree on the performance of the financial sector.

Chapter 2 is based on Luck and Schempp (2014c) and provides a model that unifies the notion of self-fulfilling banking crises and sovereign debt crises. In this model, a bank run can be contagious by triggering a sovereign default, and vice versa. A deposit insurance scheme can eliminate the adverse equilibrium only if the government can repay its debt and credibly insure deposits, irrespective of the performance of the financial sector. Moreover, we analyze how banking crises and sovereign defaults can be contagious across countries. We give conditions under which the implementation of a banking union, including a supranational deposit insurance, prevents crises effectively and at no cost. Finally, we discuss the current proposals for a banking union in the euro area and argue that it should be extended by a supranational Deposit Guarantee Scheme.

Throughout the first two chapters, we abstract from regulatory requirements and the incentive to circumvent such. This assumption is relaxed in Chapter 3, which is based on Luck and Schempp (2015). We study a banking model of maturity transformation in which regulatory arbitrage induces the coexistence of unregulated shadow banking and regulated commercial banking. As in Chapter 1, optimal intermediation relies on outside liquidity. While a deposit insurance and a capital requirement can ensure the stability of the commercial banking sector in this setup, panic-based runs are possible in the shadow banking sector. We then emphasize a new channel through which these runs are contagious, affecting the regulated banking sector: Because a run on shadow banks induces fire sales, the wholesale funding conditions for regulated banks may also deteriorate via a binding cash-in-the-market constraint. We use the model to argue that regulatory arbitrage poses a threat to the stability of regulated banks, even in the absence of explicit or implicit contractual linkages between regulated and non-regulated banking. This is important, as most reforms after the 2007-09 financial crisis have targeted explicit or implicit contractual linkages between the two sectors, often under the premise that this is an effective measure for shielding regulated commercial banks from turmoil in the shadow banking sector. We indicate that more regulatory measures may be desirable, particularly restrictions on wholesale funding. We discuss the liquidity regulation proposed in Basel III and argue that a more macro-prudential set of rules would be an improvement.

All three chapters are concerned with the self-fulfilling elements of financial crises. We find that the government has a distinct role in dealing with such fragilities. During the 2007-09 financial crisis, fiscal authorities and central banks tried to stabilize the financial system by engaging in bail-outs and by providing guarantees and liquidity assistance for distressed institutions. While such ex-post measures can be useful once a financial system is in a state of crisis, this thesis contributes to the understanding of how ex-ante measures can or cannot prevent such crises in the first place. This thesis presents and analyzes models which indicate that a fiscally strong government can ensure efficient liquidity provision by issuing government bonds. If a single country is fiscally weak, we show how a banking union that includes a supranational Deposit Guarantee Scheme may be mutually beneficial for its participants. Finally, we show how regulatory arbitrage can reintroduce panic-based runs in an economy in which the banking sector is covered by a safety net. We show that panics in the shadow banking sector may affect the regulated banking sector even the absence of contractual linkages between the sectors. This indicates that regulatory arbitrage may pose a severe restriction on the government's ability to stabilize the banking system.

1
Outside Liquidity, Rollover Risk, and Government Bonds

1.1 Introduction

Empirical evidence suggests that investors value the liquidity of government bonds (see, e.g., Longstaff, 2004; Krishnamurthy and Vissing-Jorgensen, 2012). There are various explanations for why incomplete financial markets and financial frictions give rise to a demand for liquidity, and for government securities as means to provide such liquidity. Government bonds may be valuable to investors as a simple medium of transfer across time, e.g., to enhance risk-sharing (see, e.g., Gale, 1990) or to improve investment by alleviating frictions (see, e.g., Woodford, 1990; Saint-Paul, 2005). Demand for government securities may especially arise when private liquidity provision is limited, e.g., if moral hazard and commitment problems restrict the pledgeable income of private agents. Publicly issued claims may guarantee the provision of liquidity and reduce the need to set liquid real assets aside (Holmström and Tirole, 1998, 2011). Moreover, it lies in the nature of government bonds that they mitigate the adverse selection problems typically associated with liquidity provision because they are free from private information (Gorton and Pennacchi, 1990; Gorton and Ordoñez, 2013).

This chapter provides a simple but novel explanation for why government securities are especially suited to manage liquidity needs: government securities are less prone to coordination failures than privately issued claims, i.e., less exposed to rollover risk.

In the run-up to the recent financial crisis, financial intermediaries satisfied liquidity needs by transforming long-term real investments into liquid claims instead of setting liquid real assets aside. However, when the crisis unfolded as a consequence of various shocks in the housing market, privately produced assets stopped being liquid – leaving financial markets and intermediaries in turmoil.[1] The crisis ultimately appears as an inability of the private sector to provide liquidity efficiently to the economy.

In our model, financial intermediaries[2] optimally provide liquidity not through holding liquid real assets that can be used at will (inside liquidity). Instead, they optimally rely on liquidity that investors provide in exchange for claims on future returns of long-term real investments (referred to as private outside liquidity). The key friction of our model is that at the time of initial investment, it impossible to contract with the potential providers of private outside liquidity such as wholesale funding. While the reliance on outside liquidity increases profitable long-term investment, it may be also associated with a rollover risk. We argue that this rollover risk is inherent in liquidity management with privately produced claims. We show that the rollover risk may either make intermediaries refrain from providing liquidity optimally in the first place, or it may make the economy inherently fragile. In turn, under the assumption that the government never defaults, public claims are free from such risk. Satisfying liquidity needs by selling government securities in exchange for outside liquidity (referred to as public outside liquidity) may thus enhance efficiency and stability.

We derive our results from a banking model in the tradition of Diamond and Dybvig (1983, henceforth D&D). Demand for liquid assets arises from an idiosyncratic liquidity risk on the part of consumers. Financial intermediaries provide optimal risk-sharing to consumers by offering demand-deposit contracts. However, we alter the D&D setup by assuming that banks can sell claims on their future returns to investors in the interim period in exchange for outside liquidity. Banks use the proceeds to serve early withdrawing consumers. This model feature is reminiscent of Holmström and Tirole (1998, 2011) and Bolton et al. (2011).

The model's implications are the following: First, the presence of investors who may buy claims on future returns generally allows a reduction of the holdings of liquid real assets in order to manage liquidity. Banks can conduct more productive, but illiquid long-term investments. Second, we find that intermediaries might not be able to manage liquidity optimally with privately produced claims. Relying on outside liquidity by investors in exchange for privately produced claims exposes an intermediary to the risk of a rollover freeze. There is strategic complementarity between investors in their decisions

[1] See, e.g., Hellwig (2009), Brunnermeier (2009), Krishnamurthy (2010), and Caballero (2010).

[2] We use the terms "bank" and "financial intermediary" interchangeably throughout the chapter.

to purchase claims on intermediaries' future returns. If no investor purchases claims, the intermediary will be forced to conduct costly liquidation. This in turn may make it optimal to refuse a rollover. Importantly, the rollover risk – unlike the classical bank run problem – cannot be eliminated by a classic deposit insurance or by a suspension of convertibility. This caused by the friction that outside liquidity is not contractible in the initial period. The potential rollover freeze in turn may make intermediaries either reluctant to implement the first-best, or it may make the economy inherently fragile.

As a third result, we show that in the presence of potential coordination failures between investors, the existence of public claims increases welfare. These claims allow intermediaries to implement the optimal allocation without exposing the economy to the risk of a rollover freeze. The reason is simple: under the assumption that the government never defaults, government securities are never subject to a coordination problem, i.e., there is no strategic complementarity between the investors in their decisions to purchase government bonds. In contrast to privately produced assets, the value of government securities is independent of the decision of investors to purchase the security or not. By using government bonds to manage liquidity, banks can reduce inefficient reliance on inside liquidity while avoiding rollover risk. Consequently, government borrowing may have non-Ricardian effects (see, e.g., Barro, 1974).

Finally, we discuss the assumption that the government can always repay its debt. We show that once the government's ability to repay depends on the banking sector, a run on the banking sector may be complemented by a run on government debt if there is public supply of liquidity. In this case, the positive effects of public liquidity provision may vanish. We analyze the interplay of sovereign defaults and banking crises in more depth in Chapter 2.

We use the term "outside liquidity" in the sense of Holmström and Tirole (2011), Bolton et al. (2011), and Gourinchas and Jeanne (2012). The concept of inside and outside liquidity is to some degree reminiscent to the definition of inside and outside money (see, e.g., Lagos, 2006), but there are subtle differences. Outside money is money that is not anyone's liability, and that is thus a net asset for the private sector. In contrast, inside money is created within the private sector, and is thus some private agent's liability. Similarly, inside liquidity is the liquidity that is created within a specified sector, while outside liquidity is supplied by agents or institutions outside this sector. In contrast to the definition of outside money, outside liquidity is mostly defined "from the point of view of the financial sector".[3] In Bolton et al. (2011), inside liquidity denotes the intermediary's cash reserves, whereas the intermediary

[3]Definition in Gourinchas and Jeanne (2012); other definitions are similar.

can raise outside liquidity by selling assets to long-term investors (hedge funds and pension funds). Thus, outside liquidity is the label for liquidity that investors supply to banks (and thus to consumers).

This chapter is closely related to the literature on the government's role in providing safe assets for the purpose of liquidity management. As in the seminal paper by Holmström and Tirole (1998), we allow the economy to reduce the holdings of real assets and to issue claims on future returns in order to manage liquidity needs. In contrast to Holmström and Tirole, the limitation of private liquidity supply originates not from agency problems, but from coordination problems. In terms of our results, this chapter is close to a series of recent papers (Greenwood et al., 2012; Gourinchas and Jeanne, 2012; Gorton and Ordoñez, 2013). With Gorton and Ordoñez (2013), we share the notion that government bonds are more liquid than privately produced assets and make the economy more stable. However, their reasoning is based on the information sensitivity of assets.[4] They show that liquidity provision by privately produced assets may make an economy fragile, as seemingly safe assets may become illiquid when they become information-sensitive. Government bonds in turn are less information-sensitive and thus more liquid. With the paper by Greenwood et al. (2012) we have in common that the government has a comparative advantage in bearing refinancing risk relative to the private sector, and thus public provision of liquidity is welfare-enhancing. However, their focus is on the maturity of different securities. Finally, Gourinchas and Jeanne (2012) provide a macroeconomic model with inside and outside liquidity. As in our setup, a crisis occurs when private liquidity provision is insufficient and the role of public securities for financial stability is emphasized.

The results of this chapter can also be interpreted in the light of the theory of liquidity mismatch. Brunnermeier et al. (2013) argue that maturity transformation and the associated maturity mismatch are not problematic per se. Fragility arises only if maturity transformation also induces a liquidity mismatch. While financing a 20 year government bond with demand deposits is an extreme form of maturity mismatch, it does not constitute a liquidity mismatch as long as there is a liquid market for government bonds. In our model, the government bonds on the banks' balance sheets neither change the mechanism of maturity transformation nor the liquidity mismatch, but it substantially reduces the liquidity mismatch.

We also relate our results to recent empirical findings. In our model, liquidity benefits from government bonds have real effects, consistent with the

[4]See Dang et al. (2013a) and Dang et al. (2013b) on information (in)sensitivity of assets and financial crises.

evidence that investors value these attributes (Krishnamurthy and Vissing-Jorgensen, 2012). Moreover, public provision of liquidity reduces the fragility in our setup, which is in line with the finding that financial crises are more likely when little public debt is available (Krishnamurthy and Vissing-Jorgensen, 2013) and financial crises seem to be related to excessive private debt rather than public debt (Jordà et al., 2013; Schularick, 2014).

This chapter is also very closely related to theories of banking, in which intermediaries optimally rely less on inside liquidity and more on sales of claims on long-term investments, such as the model by Bolton et al. (2011). This model is concerned with the timing of trade in the presence of uncertainty and asymmetric information, while we focus on the coordination failures that may be associated with outside liquidity.

Finally, this chapter contributes to the literature on liquidity provision by financial intermediaries. D&D have argued that financial intermediaries can provide optimal risk-sharing to consumers and allow them to benefit from profitable long-term investments by offering demand-deposit contracts.[5] In contrast, we argue that the ability of financial intermediaries to provide liquidity is limited. We are far from being the first to address the problems of liquidity provision by intermediaries. The banking literature has already produced various arguments. It has been argued that the ability of banks to provide risk-sharing in the presence of financial markets is very limited (Jacklin, 1987; Farhi et al., 2009).[6] Especially when consumers are able to adjust their portfolio, liquidity provision may be harmed (von Thadden, 1998). Moreover, banks may be unable to implement the first-best through demand-deposit contracts in the presence of macroeconomic interest rate risk (Hellwig, 1994).[7] Under aggregate risk and in the presences of moral hazard, financial intermediaries may not be able to insure firms against liquidity shocks either (Holmström and Tirole, 1998). The creation of liquidity through interbank trade may also be limited if banks are unable to diversify the liquidity risk of their consumers (Bhattacharya and Gale, 1987).

Our argument, however, is neither based on agency problems nor on aggregate uncertainty. We argue that liquidity management with privately issued claims creates a coordination problem between those investors who could provide liquidity. A memorable insight from the seminal contributions by Bryant (1980) and D&D is that liquidity provision may be associated with the exis-

[5] On the optimality of intermediaries of liquidity provider, see also, e.g., Gorton and Pennacchi (1990), Calomiris and Kahn (1991), Diamond and Rajan (2001), and Kashyap et al. (2002).

[6] See also Diamond (1997) and Fecht (2004).

[7] See also Allen and Gale (1998) on this point.

tence of run equilibria and make an economy inherently fragile.[8] Importantly, the rollover problem in our setup differs from the classical bank run problem. We show that the coordination problem cannot be eliminated by a deposit insurance nor by a suspension of convertibility. Ultimately, the rollover risk associated with optimal private liquidity provision may prevent the implementation of the optimal allocation in the first place. This chapter stands in contrast to models arguing that banks are especially suited to provide liquidity *because* of their fragile capital structure. Amongst others, Calomiris and Kahn (1991) and Diamond and Rajan (2001, 2005) argue that the fragile nature of bank balance sheets disciplines bank managers and thus allows overcoming commitment problems associated with liquidity provision. In contrast, we argue that the potential rollover risk may cause banks to refrain from supplying liquidity in an optimal fashion in the first place.

We proceed as follows: In Section 1.2, we introduce the general setup and derive the first-best allocation and show how it can be implemented by banks. In Section 1.3, we investigate how the first-best and its implementation change if we introduce outside liquidity. Section 1.4 shows how the rollover risk associated with privately liquidity supply influences the stability and efficiency of banks. Finally, in Section 1.5, we demonstrate why the provision of public liquidity by the government is superior to the private case.

1.2 Intermediation with Inside Liquidity

Consider an economy that goes through a sequence of three dates, $t \in \{0, 1, 2\}$. There is a single good that can be used for consumption as well as for investment. Moreover, there are two investment technologies that we refer to as assets. The economy is populated by risk-averse consumers who face an idiosyncratic liquidity risk.

Consumers

There is a continuum of *ex ante* identical consumers with mass one. Each consumer is endowed with e_0 units of the good in $t = 0$. There are two types of consumers, denoted by $\theta_i \in \{0, 1\}$. The type determines the consumer's

[8] Following the seminal contributions by Bryant and Diamond and Dybvig, a vast literature on bank runs evolved. See, e.g., the literature regarding information-based runs (Jacklin and Bhattacharya, 1988), models with positive probability of bank runs (Postlewaite and Vives, 1987; Chari and Jagannathan, 1988; Allen and Gale, 1998; Rochet and Vives, 2004; Goldstein and Pauzner, 2005), models with interbank contagion (Allen and Gale, 2000; Dasgupta, 2004; Uhlig, 2010), runs in repurchase agreements (Martin et al., 2014a), and dynamic runs (He and Xiong, 2012).

intertemporal preference for consumption in periods one and two. With probability π, consumer i is an "impatient consumer" who needs to consume in $t = 1$, denoted by $\theta_i = 1$. With probability $(1 - \pi)$, she is a "patient consumer" who is indifferent between consumption at both dates, denoted by $\theta_i = 0$. Initially, consumers do not know their type; their probability of being type 1 is identical and independent. In period one, each consumer privately learns his type. This private revelation can be considered as a liquidity shock.

A consumption profile (c_1, c_2) gives a consumer i a utility of

$$U(c_1, c_2, \theta_i) = \theta_i u(c_1) + (1 - \theta_i) u(c_1 + c_2), \tag{1.1}$$

where the "baseline" utility $u : \mathbb{R}^+ \rightarrow \mathbb{R}$ is an increasing and strictly concave function that is twice continuously differentiable and satisfies Inada conditions, $u'(0) = +\infty$ and $u'(+\infty) = 0$. For each consumer, the ex-ante expected utility is given by $\mathrm{E}U(c_1, c_2) = \pi u(c_1) + (1 - \pi) u(c_1 + c_2)$.

Notice that the attributes "patient" and "impatient" characterize the consumer's exogenous type which determines his preference, denoted by θ_i. In contrast, the attributes "late" and "early" will characterize the timing of consumption which is endogenous: An "early consumer" consumes in $t = 1$, while a "late consumer" consumes in $t = 2$.

Assets

There are two different assets (investment technologies) available in $t = 0$: a short asset (storage technology), and a long asset (production technology). The short asset transforms one unit of the good at time t into one unit of the good at $t + 1$, effectively storing the good. The long asset promises a higher expected return in the long run. However, this asset is considered to be illiquid as it can only be liquidated with a substantial discount in $t = 1$.

The long asset is represented by a continuum of investment projects. An investment project is a metaphor for an entrepreneur who is endowed with a production technology but has no endowment of goods for investment. Each consumer has access to exactly one project (or equivalently is matched with exactly one entrepreneur). Each investment project yields a stochastic return of R_i units in $t = 2$ for each unit invested in $t = 0$. The return R_i is the realization of an independently and identically distributed random variable $\tilde{R}$, characterized by a probability distribution F. F is continuous and strictly increasing on a compact interval with minimum $\underline{R} > 0$ and maximum $\overline{R}$, with $\mathrm{E}[R_i] = R > 1$. We assume that the realization of an investment project's long-term return R_i is privately revealed to the project's financier in $t = 1$. As we will shortly see, the idiosyncratic risk implies that financial intermediaries dominate a financial markets solution in terms of welfare.

Finally, an investment project may be physically liquidated prematurely at a rate $\ell \in (0, 1/R)$ in $t = 1$, yielding a liquidation return of ℓR_i units. The liquidation return of a project thus depends on the project's stochastic long-term return. However, the ratio of liquidation return to long-term return is constant and equal to ℓ.

1.2.1 First-Best Allocation

The allocation of consumption across different consumer types and different periods is denoted by $\{c_1(\theta), c_2(\theta)\}_{\theta\in\{0,1\}}$. The unconstrained optimum results from the social planner's first-best problem, which is given by

$$\max_{\{c_1(\theta),c_2(\theta)\}_{\theta\in\{0,1\}}} \pi u(c_1(1)) + (1-\pi)u(c_1(0) + c_2(0)) \tag{1.2}$$

subject to

$$\pi\left(c_1(1) + \frac{c_2(1)}{R}\right) + (1-\pi)\left(c_1(0) + \frac{c_2(0)}{R}\right) \leq e_0. \tag{1.3}$$

Equation (1.3) is the feasibility condition, resulting from the initial investment constraint in $t = 0$ and the two budget constraints in period one and two.

In the first-best, it holds that

$$c_2(1) = c_1(0) = 0. \tag{1.4}$$

The late consumption levels of patient and the early consumption level of impatient consumers are then given by the following first-order condition and budget constraint:

$$u'(c_1(1)) = Ru'(c_2(0)), \tag{1.5}$$

$$\pi c_1(1) + (1-\pi)\frac{c_2(0)}{R} = e_0. \tag{1.6}$$

The optimal allocation is thus characterized by the trade-off between insurance against liquidity risk (investment in storage) and productive investment (investment in the long assets).

1.2.2 Diamond & Dybvig (1983)

The model described above resembles the essential features of the framework of the D&D model. Therefore, we briefly review the key results of the seminal D&D model. The first important result is that a competitive financial market generally fails to implement the first-best allocation. In contrast, a competitive banking sector or a representative bank can implement the first-best. It is assumed that the law of large numbers applies on the bank level. That is, there is neither uncertainty on the fraction of consumers being impatient, π, nor on the return of the portfolio of long assets, R.

A bank that aims at maximizing consumers' expected utility thus needs to maximize (1.2) subject to the feasibility constraint, and because the type of consumers is private information, the constrained efficient program contains two additional restrictions. The allocation of consumption must be such that no consumer has an incentive to misreport his type in the interim period:

$$u(c_1(1)) \geq u(c_1(0)), \tag{1.7}$$

$$u(c_1(0) + c_2(0)) \geq u(c_1(1) + c_2(1)). \tag{1.8}$$

Constraint (1.7) ensures that a impatient consumer has no incentive to misreport, while (1.8) ensures that a patient consumer does not want to misreport. Adding constraints (1.7) and (1.8) to the first-best problem, however, does not change the solution because the constraints are not binding in the first-best. This implies that the first-best is in fact implementable given the friction of unobservable types. The second-best thus coincides with the first-best.

The proposed mechanism, a bank representing a contestable banking sector, proceeds as follows (see also Figure 1.1): In $t = 0$, the endowment of all consumers is collected. In exchange the bank offers a demand-deposit contract that allows a consumer to withdraw c_1^* units in $t = 1$ and c_2^* units in $t = 2$. The bank chooses c_1^{DD} and c_2^{DD} such that $u'(c_1^{DD}) = Ru'(c_2^{DD})$ which is the FOC for the first-best allocation, see Equation (1.5). $R > 1$ and concavity of u imply that $c_1^{DD} \leq c_2^{DD}$, and thus Equations (1.7) and (1.8) are satisfied and it is incentive-compatible for patient consumers to withdraw only in $t = 2$. The bank invests $e_0 - I = \pi c_1^{DD}$ in the storage technology and the remaining funds $I = e_0 - \pi c_1^{DD}$ in the long asset, which implies that Equation (1.6) holds. The representative bank is thus able to implement the first-best allocation.[9]

The second important result is that there is also a second type of equilibrium in $t = 1$. If all patient consumers desire withdrawing at once, the bank will

[9]Note that the optimality of the banking solution relies on a no-trading restriction of consumers in $t = 1$; see Jacklin (1987) and more recently Farhi et al. (2009).

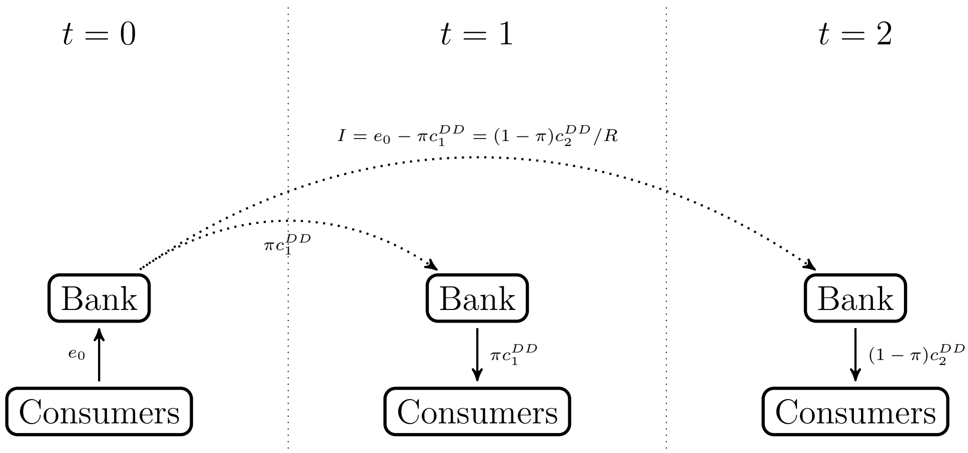

Figure 1.1: Diamond and Dybvig (1983). The graph illustrates investment and the flow of goods (solid arrows). Initially, consumers deposit their endowment at the bank. Banks invest the endowment in the long and in the short asset. In periods one and two, early and late consumers are served with the returns of the short and the long asset, respectively.

be left with assets of $\pi c_1^{DD} + \ell(1 - \pi c_1^{DD}) < 1$ which is typically strictly less than its total liabilities in $t = 1$, which amount to c_1^{DD}.[10] The bank will therefore be insolvent in $t = 1$ and no funds for patient consumers will be left over in $t = 2$. It is thus optimal for all patient consumers to withdraw and a bank run may constitute an equilibrium in the interim period. In fact, there are two subgame-perfect Nash-Equilibria in $t = 0$, one in which the bank is established, and a second one in which consumers refuse to deposit funds in the banks as they expect a bank run in $t = 1$.

Finally, the third result of the D&D model is that the adverse run equilibrium can be eliminated by two policy measures: either the banks should commit to suspending convertibility after paying out an overall amount of πc_1^{DD}, or the government should provide a deposit insurance which guarantees c_1^{DD} units for each consumer in $t = 1$, irrespective of the banks being solvent or not. In both cases, the adverse equilibrium can be eliminated at

[10] The run equilibrium exists whenever $c_1^{DD} \geq 1$. This condition is typically satisfied through the assumption that the coefficient of relative risk-aversion is larger than one, i.e., $-cu''(c)/u'(c) > 1$ for every c.

no cost.[11]

1.3 Intermediation with Outside Liquidity

We now introduce a new type of agents whom we refer to as "investors". Investors can provide banks or consumers with liquidity in the interim period which we refer to as "outside liquidity". In the following, we analyze how the optimal allocation is altered by allowing for interim outside liquidity, and how the new first-best allocation may be implemented by a banking sector.

Assume that there is a continuum of investors with mass $\alpha > 1$. Investors have no endowment initially, but with a probability $1/\alpha$ an investor j has an endowment of $e_{1,j} = e_1$ in period one. Otherwise, her endowment is zero. Therefore, the mass of investors that has a positive endowment is equal to one, and the overall endowment of all investors is $e_1 = \int e_{1,j} \mathrm{d}j$. We assume that $e_1 > \pi R e_0$. As we will see, this condition assures that the supply of outside liquidity is never limited by a binding resource constraint.

The key friction of private outside liquidity is the following: We assume that it is the investor's private information whether she has a positive endowment in period one. Thus, investors cannot write enforceable contracts in $t = 0$, which are contingent on whether they have a positive endowment in $t = 1$. Because investors cannot contract in $t = 0$, we will only consider their behavior form period $t = 1$ onwards. Furthermore, we have to consider only those investors who have a positive endowment.

We assume that investors have no market power, and that they are indifferent between consuming in periods one and two. Their utility is given by $v(\hat{c}_1 + \hat{c}_2)$, where $\hat{c}_t$ is her consumption in period t and $v : \mathbb{R}^+ \rightarrow \mathbb{R}$ is a strictly increasing function. Consequently, they are willing to invest their complete endowment e_1 as long as the gross return in $t = 2$ is at least e_1.

1.3.1 First-Best with Outside Liquidity

We now derive the new first-best allocation, given that outside liquidity is available in the interim period. The social planner's objective function, specified by the maximization problem (1.2), remains unchanged. The objective is maximizing the consumers' welfare, whereas investors' utility does not enter

[11] Observe that for suspension of convertibility to be an effective measure there must not be aggregate uncertainty about the actual fraction of consumers who withdraw early. Moreover, suspension of convertibility is also ineffective if withdrawing depositors are paid out by new depositors, see the extension of the D&D setup to an overlapping generation setting by Qi (1994).

our welfare measure. However, we assume that even the social planner cannot transfer funds from investors to consumers without restrictions. Because investors' welfare does not directly enter into the objective function, we require that investors must be willing to participate.[12] The aggregate transfer from investors to consumers in period one is denoted by $d_1 \leq e_1$, and d_2 denotes the reverse transfer in period two. Investors' participation constraint is given by $d_2 \geq d_1$. It is straightforward that this constraint will be binding in the optimum, i.e., $d_1 = d_2$ will hold in the following.

The first-best program with outside liquidity is slightly different from the one in the previous section. We now explicitly consider the budget constraints in each period. The variable $I \in [0, e_0]$ again denotes the investment in the long asset, and an amount $e_0 - I$ is invested in storage. Let d denote the amount of interim liquidity (i.e., the amount of liquidity that is transferred between investors and consumers), where $d = d_1 = d_2 \leq e_1$.

The budget constraints for the two periods are given by

$$e_0 - I + d \geq \pi c_1(1) + (1 - \pi)c_1(0), \tag{1.9}$$

$$RI \geq \pi c_2(1) + (1 - \pi)c_2(0) + d. \tag{1.10}$$

Constraint (1.9) ensures that, in $t = 1$, the payments to consumers do not exceed the sum of inside liquidity (storage) and interim outside liquidity. Constraint (1.10) ensures that, in $t = 2$, the sum of payments to consumers and the repayment of interim outside liquidity does not exceed the return from investment in the long asset.

Proposition 1.1 (First-best)**.** *In the presence of outside liquidity, the consumers' first-best consumption allocation is given by* $c_1(1) = c_2(0) = Re_0$*, and* $c_1(0) = c_2(1) = 0$*. It is attained by choosing* $I = e_0$ *and* $d = \pi Re_0$*.*

The fundamental insight of Proposition 1.1 is that the D&D allocation, in which consumption levels in both periods are strictly less than Re_0, can strictly be improved upon.[13] The social planner can make full use of the productive long asset because the supply of liquidity in the interim period removes the need to invest in storage. In the model with outside liquidity, the

[12]In our setup, a comparison of consumers' and investors' utility does not appear meaningful. We are neither interested in the allocation of risk, nor in redistribution of wealth between the two groups of agents. We interpret the investor's participation constraint rather as a resource constraint that as a friction. It thus appears adequate to refer to the optimum as the "first-best".

[13]This result is reminiscent of the finding by Qi (1994), who shows that storage may be redundant in a overlapping-generation version of the D&D model.

trade-off between liquidity insurance (provided through the short-asset, i.e., inside liquidity) and return (long-asset) can be solved such that consumption is perfectly smoothed by making full use of the productive technology.

Observe that the first-best allocation is not unique if the endowment of investors strictly exceeds the amount that is given to impatient consumers in the interim period, $e_1 > \pi Re_0$. Because early and late consumption are perfect substitutes for patient consumers, impatient consumers could receive some positive payment in $t = 1$, as long as their total amount of consumption remains unchanged. Without loss of generality, we focus on the solution presented in Proposition 1.1 in the following, i.e., impatient consumers only consume late.[14]

1.3.2 Efficient Banking

We now show that the first-best allocation in the model with outside liquidity can be implemented by an institution that is reminiscent of a financial intermediary that signs demand-deposit contracts with consumers in period zero. In contrast to the situation without outside liquidity, a bank now only invests in the long asset and raises liquid funds in the interim period. It raises funds by issuing claims and selling them to investors. We will refer to those claims as debt.[15]

As in the D&D setup, one may think of the banking sector as a contestable market. The assumption of free entry and the resulting perfect competition imply that financial intermediaries implies contracts that maximize the expected utility of consumers.[16] Again, the law of large numbers is assumed to apply on the bank level, resulting in a gross return of R on the long asset with certainty, and a fraction of early impatient consumers of exactly π.

Therefore, banks can implement the first-best in the following way (see also Figure 1.2): banks collect the total endowment e_0 of all consumers as deposits in period zero against the promise that consumers can withdraw Re_0 units at any time. In order to serve their obligations, banks invest all of the economy's $t = 0$ endowment in the long assets, transforming them in Re_0 units in $t = 2$. In the interim period, banks sell claims d on their portfolio of long assets to investors. Because banks are assumed to be able to diversify the liquidity

[14]Notice that this allocation can be attained by choosing any $d \in [\pi Re_0, e_1]$.

[15]Notice that the bank could likewise issue equity claims. As we frame the problem as one of rollover, however, we refer to the claims as debt claims without giving a specific microeconomic reasoning why debt is preferred over equity.

[16]Alternatively, one may assume that the banking sector is a mechanism or a coalition of consumers that maximizes the consumers' expected utility.

and the return risk, there is no adverse selection in the market for claims in the interim period. Therefore, the investors' participation constraint implies that banks can sell their claims at par. Banks will sell claims with a total value of πRe_0, and only impatient consumers withdraw early. The issuance of claims is equivalent to a rollover of debt, as the liability towards depositors is replaced by liabilities towards investors.

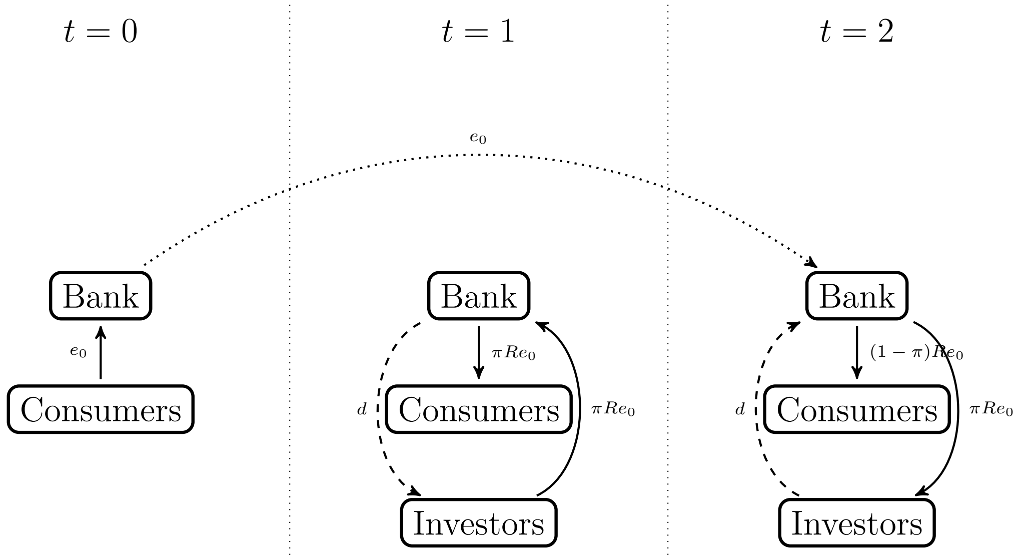

Figure 1.2: Private Outside Liquidity. The graph illustrates investment and the flow of goods (solid arrows) and claims (dashed arrows). Claims associated with demand-deposit contracts are not depicted. In $t = 0$, consumers deposit their endowment e_0 at the banks. Banks invest the endowment in the long asset, transforming e_0 units of the good into Re_0 units. In the interim period, early consumers are served by selling claims d to investors. In $t = 2$, banks redeem the claims of investors and repay investors and late consumers using the returns of the long asset.

Proposition 1.2 (Implementation of the first-best)**.** *The first-best allocation $c_1(1) = c_2(0) = Re_0$ can be implemented in a demand deposit economy. Banks invest only in the long asset. Banks serve withdrawing consumers in the interim period by issuing claims on future returns and selling these to investors in exchange for outside liquidity.*

The implementation of the first-best allocation thus involves privately produced assets. Instead of investing in storage in $t = 0$, financial intermediaries issue claims on their future returns in the interim period. The proceeds from selling these to investors are used to serve withdrawing consumers. This allows intermediaries to increase the investment in the long asset and thus to promise higher payments to consumers.

As in the D&D setup, the first-best allocation cannot be implemented by trade in a financial market in the interim period. The reason is somewhat different, however. In our model, the main reason why financial markets cannot implement the first-best is that consumers cannot insure themselves against idiosyncratic return risk. For such an insurance, contingent contracts between consumers or between consumers and investors are required, which are not feasible because of the unobservability of consumers' returns, liquidity needs and investors' endowments. For a detailed analysis, see the Appendix 1.A. Because the idiosyncratic return risk restricts the benefits of side-trading, it can be interpreted as trading restriction in the sense of (Jacklin, 1987; von Thadden, 1998).

The fragility associated with this implementation will be discussed in the next section.

1.4 Private Outside Liquidity and Rollover Risk

In this section, we show that a bank may face a rollover freeze if it relies on raising liquidity by issuing claims in the interim period. It turns out that the efficient private provision of liquidity is inherently fragile. Therefore, banks might refrain from relying on outside liquidity. Instead, they might rely on the inefficient storage technology and offer the D&D contract.

1.4.1 Rollover Freeze

Consider the subgame starting in the interim period, given that banks have invested the complete endowment in the long asset. In this subgame, consumers are endowed with a demand-deposit contract promising Re_0 units in either period, and banks issue claims on their future returns in order to serve withdrawing consumers. Consumers have the choice to withdraw early or to wait, and investors have the choice whether to buy claims on bank assets. In the previous section, we saw that there exists an efficient equilibrium of this subgame in which consumers do not run on banks and investors roll over the banks' debt. However, there are strategic complementarities between agents, giving rise to multiple equilibria. As in the D&D model, there is a strategic complementarity between consumers whether or not to withdraw early. In the model with outside liquidity, an additional strategic complementarity arises between investors concerning their decision whether to buy claims on bank assets and thus to roll over the banks' debt. Furthermore, there is also strategic complementarity across these two groups of agents. In the following, however, we will only focus on the strategic complementarity between investors.

In order to understand the rollover freeze, consider a situation in which no investor is willing to purchase bank claims. Let us first assume that only impatient consumers withdraw early. In this case, banks will need to liquidate a positive fraction $z = \min[\pi/\ell, 1]$ of their long assets at the inefficient rate ℓR in order to serve impatient consumers with an amount of πRe_0 units. This liquidation implies that the bank will only be left with $(1 - z)Re_0$ in $t = 2$. Therefore, the bank will not have sufficient funds at hand in order to serve its patient customers or any investors in $t = 2$. Therefore, an individual investor will not provide any liquidity in $t = 1$, as banks will be insolvent in $t = 2$. This implies that even if consumers behave diligently and do not run on the bank, a rollover freeze always constitutes an equilibrium.

This consideration also leads to the insight that the standard measures to prevent inefficient liquidation and thus financial crises, such as deposit insurance (DI) or suspension of convertibility (SoC), become ineffective. The reason is that these policies are only targeted at breaking the strategic complementarity between depositors – they are concerned with the demand for liquidity, but not with its supply. The DI may keep patient consumers from running on banks, but a bank run is not the only way a bank can become illiquid and insolvent once a bank relies on outside liquidity. Banks may in fact experience a rollover freeze as the deposit insurance does not alter the strategic complementarity between investors. Moreover, SoC is also ineffective. By suspending convertibility, banks can limit the amount they pay out to early consumers, which induces stability in the D&D model because it eliminates the need for liquidation. However, if banks rely on outside liquidity, this measure does not prevent liquidation in case of a rollover freeze, inducing consumers to run.

Lemma 1.1. *In the $t = 1$ subgame, a rollover freeze by investors constitutes a Nash equilibrium. Moreover, a rollover freeze may occur independently of whether there is a bank run or not, and irrespectively of the existence of a credible deposit insurance or of banks committing to suspend convertibility.*

In the following, we show that the fact that there may be a rollover freeze in $t = 1$ makes banks either refrain from providing the efficient level of liquidity or it will expose the economy to the rollover risk. In the latter case, the economy will be fragile despite DI or SoC.

1.4.2 Inefficient Liquidity Provision

Because the subgame of the interim period has an efficient as well as an adverse equilibrium, the whole game (starting in period zero) has at least one additional, inefficient subgame-perfect Nash equilibrium. While there might be a continuum of equilibria, we are interested in the generic case where investors coordinate on a rollover freeze. In a subgame-perfect Nash equilibrium, consumers and banks anticipate not being able to raise any outside liquidity in the interim period. Given that outside liquidity is not available in the interim period, banks have to rely on storage again. The constraint efficient allocation is given by the Diamond-Dybvig allocation described in Section 1.2.2.

Proposition 1.3. *The model has a subgame-perfect Nash equilibrium in which investors do not roll over bank debt, consumers do not run on banks, and banks implement the Diamond-Dybvig consumption allocation, given by* $c_1(1) = c_1^{DD}$ *and* $c_2(0) = c_2^{DD}$*, and* $c_1(0) = c_2(1) = 0$*.*

We have seen that if banks rely on raising outside liquidity by issuing claims on their future return, they are exposed to the risk of investors coordinating on a rollover freeze. The most efficient allocation entails full exposure to the rollover risk resulting from the coordination problem. If banks fear a rollover freeze, they might completely shy away from relying on outside liquidity, rather implementing the less efficient D&D allocation.

The reasoning in Proposition 1.3 is in fact very similar to the argument in the D&D model that if a bank run was expected in $t = 1$, consumers would not be willing to deposit their endowment in the bank in $t = 0$. However, it is important to notice that the adverse equilibrium cannot be eliminated by the standard measures (DI or SoC) in our model. This is due to the key friction of non-contractible private outside liquidity. One may assume that the government offers a credible DI or banks may commit to SoC. In our setup, this will not eliminate the fragility associated with the efficient provision of liquidity. In fact, if there is no credible DI, a third equilibrium may exist in which investors would not roll over bank debt and investors would run on the bank, which is why no bank is founded in the first place. In turn, if there is a credible DI, this equilibrium does not exist. However, the DI is tested in the equilibrium of Proposition 1.3 and may be costly for the institution providing it.

Finally, Proposition 1.3 can be seen as an argument for why liquidity provision by banks may be limited in general. We argue hat efficient liquidity provision rests on reliance on outside liquidity. However, privately produced assets may not be able to ensure the provision of outside liquidity. Due to the rollover risk associated with privately produced assets, financial intermediaries may thus not be able to implement the optimal allocation. This line of argument stands in contrast to models arguing that banks are especially suited to provide banking services because of their fragile capital structure (Calomiris and Kahn, 1991; Diamond and Rajan, 2001). In our setup, the fragility in the interim period can cause banks to refrain from supplying liquidity in an optimal fashion.

1.4.3 Fragility

Until this point, we have tied our hands by assuming that investors cannot coordinate their behavior on something that is not observed or not initially contractible. Formally, this means that investors cannot play a strategy by which they condition their action on a public signal that is only revealed in the interim period. This implies that a rollover freeze will never occur in a subgame-perfect Nash equilibrium in pure strategies. Either banks successfully rely on outside liquidity because they know that a rollover will be successful, or they anticipate a rollover freeze and rely on the storage technology. In equilibrium, the rollover "risk" is degenerate, as it either occurs with a probability of zero, or it occurs with a probability of one, but has no effect.[17]

We now want to consider a setup where investors can coordinate on a rollover freeze. The notion of coordination problems in the tradition of the D&D model is that depositors decide in the interim period whether to withdraw, thus coordinating on whether to run on the bank only *after* the investment decision has been made. Formally, the concept of subgame perfection requires agents to choose a strategy in period zero. Therefore, uncertainty about the action in $t = 1$ can only prevail if there exists a public signal upon which agents can condition their action. A popular illustration of such a coordination device is the concept of sunspots.[18]

We adopt this notion and assume that with some exogenous probability $p \in (0, 1)$ a sunspot occurs, and investors play a strategy that prescribes not

[17] There exists no equilibrium in which investors play mixed strategies and banks rely on rollover. While this might seem strange, it is worth mentioning that investors play a weakly dominated strategy in the "rollover equilibrium". As soon as we introduce marginal net profits for investors, rollover stops being weakly dominated and an equilibrium in mixed strategies arises.

[18] See Cooper and Ross (1998) for an analysis of the D&D setup with sunspots.

to roll over the banks' liabilities in case of this sunspot.[19] We restrict our attention to the two extremes where the probability of a rollover freeze is either close to one or close to zero.

Proposition 1.4. *As the probability of a rollover freeze converges to one, i.e.,* $p \to 1$*, the optimal investment converges towards the Diamond-Dybvig case,* $I(p) \to I^{DD}$*. Furthermore, there exists a threshold* $p_\ell \in (0,1)$ *such that for* $p \leq p_\ell$*, it is optimal fully to rely on private outside liquidity storage,* $I^*(p) = e_0$.

For the proof of Proposition 1.4, see the Appendix. The result, however, is very intuitive as it rests on the insight of the following trade-off: On the one hand, efficiency can be attained by choosing high investment in the illiquid but profitable long-term technology. Because banks thereby rely on outside liquidity, this is associated with a high rollover risk. On the other hand, stability can be attained if banks are not exposed to rollover risk. To this end, banks make use of the storage technology and thus rely on inside liquidity. In other words, the trade-off is between strong maturity mismatch and narrow banking.

If the sunspot probability p is sufficiently high, banks will implement the D&D allocation. Banks and consumers know that each unit of early consumption that is not covered by investment in the storage technology has to be raised by liquidating long assets in case of a rollover freeze. Therefore, banks will finance every unit of early consumption by using inside liquidity and the optimal allocation under this constraint is the D&D allocation. In contrast, if the probability of a rollover freeze is sufficiently small, banks choose full exposure to rollover risk by only investing in the long asset, and implement a consumption level of Re_0 for both consumer types. This implies that banks have to engage in substantial liquidation in case of a rollover freeze, but given that this risk is very low, they are willing to accept this risk.

It is worth noticing that even if a rollover freeze occurs with a positive probability, it may still be optimal that banks fully rely on outside liquidity. In this case, the economy is inherently fragile and a financial crisis may unfold in equilibrium if investors coordinate on a rollover freeze.

[19] We do not model the underlying reason for the occurrence of these sunspots, and if we did, their occurrence would probably depend upon the banks' behavior.

1.5 Public Outside Liquidity

In Section 1.3, we showed that optimal liquidity management does not rely on inside liquidity, but rather on outside liquidity. Building on this, Section 1.4 revealed that the efficient allocation can be implemented by banks issuing private claims and relying on the rollover of debt. However, outside liquidity in exchange for privately produced assets is associated with a rollover risk. The anticipation of a rollover freeze can lead to inefficient investment choices ex-ante. We now analyze how this friction could be overcome. In particular, we ask whether the government can mitigate the problem by providing liquidity.

In general, a government has the ability – unlike private agents – to commit future income via taxation or money creation. This makes claims against a public authority inherently safer than claims produced by the private sector. In the first part of this section, we therefore assume that the government never defaults. In this case, we show that the government can increase welfare by issuing a public claim that can be used by banks to manage liquidity. In the second part of this section, we relax the assumption that the government cannot default and show that the benefits from public provision of liquidity may vanish.

1.5.1 Liquidity Management and Government Bonds

Let us assume that the government never defaults. Consider the following mechanism in which the government provides liquidity by issuing a government bond (Figure 1.3): In $t = 0$, consumers deposit their endowment e_0 at a bank in exchange for a demand-deposit contract allowing the consumer to withdraw Re_0 in either period. Banks receive government bonds that promise a payment of b units by the government in $t = 2$. In exchange for the b government bonds, banks write a debt contract with the government, promising to pay d units to the government in $t = 2$. Banks and government will lend and borrow such that $d = b \geq \pi Re_0$. Effectively, banks are expanding their balance sheets by an amount of b.

In $t = 1$, banks sell πRe_0 units of government bonds to the investors and use the resulting liquidity to serve withdrawing consumers. In $t = 2$, the government has due gross liabilities of b units, necessary to redeem the government bonds. An amount of πRe_0 is paid to investors. The difference of $b - \pi Re_0$ units is a gross liability towards banks, resulting from the government bonds they did not sell to investors in $t = 1$. However, the banks also owe $d = b$ units to the government. Therefore, they have a net liability of πRe_0 units towards the government. The banks have an overall return of Re_0 from the long assets which is used to pay out $(1-\pi)Re_0$ units to the patient consumers and πRe_0 to the government.

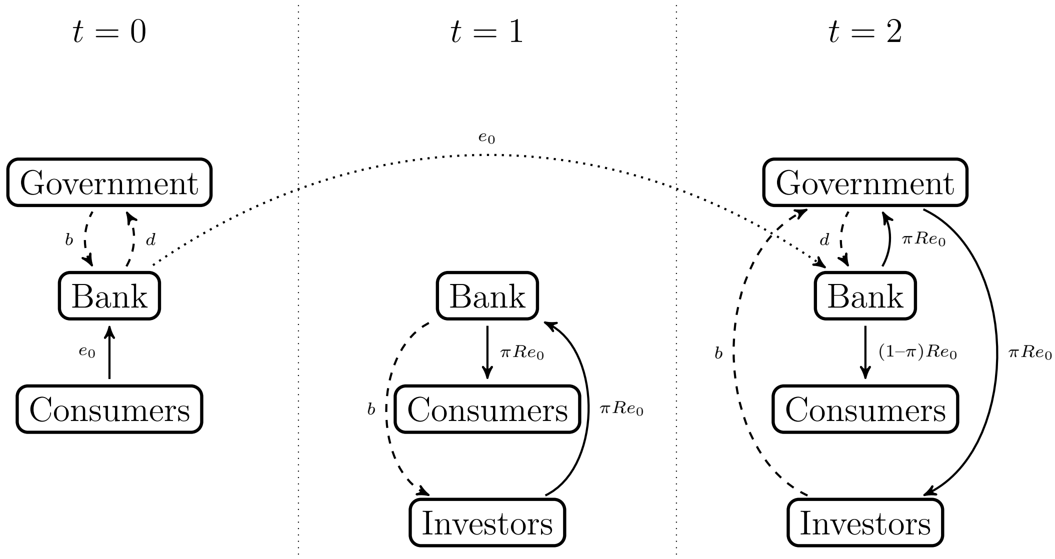

Figure 1.3: Public Outside Liquidity. The graph illustrates investment and the flow of goods and claims. In this graph, we assume that $b = d = \pi e_0 R$. The dotted arrow denotes the investment in assets and thus the "transfer" of goods between periods. The solid arrows denote the flow of goods. The dashed arrows denote the flow of net claims. For simplicity, the claims of consumers towards banks (resulting from the demand-deposit contracts) are left out.

Proposition 1.5. *If the government provides government bonds in $t = 0$, the banks are able to implement the first-best consumption allocation, given by $c_1(1) = c_2(0) = Re_0$ and $c_1(0) = c_2(1) = 0$. Furthermore, the rollover risk is eliminated and the economy has a unique equilibrium.*

By expanding their balance sheet, the banks implement the first-best allocation while the rollover risk is completely eliminated. Even if all other investors refused to buy government bonds, this would not influence the incentives of an individual investor. Public outside liquidity eliminates the coordination problem concerning the supply of liquidity by investors. Note though that public liquidity provision as described does not address the coordination problem between consumers concerning their withdrawal decisions. However, given that there is public outside liquidity, a bank run equilibrium can be eliminated by the standard measures, i.e., by introducing a deposit insurance or allowing banks to suspend convertibility. This is important as we saw in the previous section that these measures are ineffective as long as the rollover problem is not addressed, but in this context they are effective at eliminating the coordination problem.

The central reason for the stability is that by assumption the government's solvency, unlike that of a bank, does not depend on the behavior of investors. This eliminates any strategic considerations of investors when deciding to purchase government bonds in the interim period. Therefore, multiplicity of equilibria vanishes once government securities are used for liquidity management.[20]

Observe that there are alternative implementations of the first-best allocation to the one shown in Figure 1.3. There are two obvious alternatives. First, the government could insure all current and future bank liabilities ex-ante. Second, the government could provide liquidity itself in the interim period. Both mechanisms are equivalent in terms of the results in our setup, as they also eliminate the fragility and thereby enhance efficiency. However, we argue that both alternatives may not be equally desirable as they may be more problematic in a richer setup in which other issues such as agency problems may arise. Insuring bank liabilities may give rise to certain risks on behalf of the bank (e.g., risk-shifting) and creditors (e.g., weak disciplining effects). Moreover, if the government actively manages liquidity by lending directly to banks when they need funds, i.e., in a crisis, this may lead to excessive maturity mismatch as in, e.g., Farhi and Tirole (2012). We discuss these two issues in more depth in Section 6.

1.5.2 Government Solvency

So far, we have made the extreme assumption that the government is always able to repay its debt, irrespectively of what investors do and of whether there is a banking crisis. This assumption gives government bonds the important characteristic of being immune against rollover risk. We now relax this assumption in two different ways. First, we allow the government to default with some exogenous probability. We show that, in this case, public liquidity provision may still be optimal. Second, we assume that the government's ability to repay debt is endogenous and depends on the performance of the banking sector. In this case, the benefits from public liquidity provision vanish.

Assume first the government defaults with some positive probability which is given exogenously. This is not necessarily detrimental to efficiency and stability. Under the condition that investors are risk-neutral, and that the

[20]This is reminiscent of how multiplicity of equilibria is eliminated in the Kiyotaki and Moore (1997) model when government bonds are introduced (see p. 515 in Tirole, 2010).

government's solvency is only revealed after $t = 1$, the optimal allocation could still be implemented. Under these conditions, the value of government bonds is still independent of the behavior of the investors. Thus, even if there is an exogenous default probability, there is no risk of a coordination failure between investors. Government bonds would be traded at the fair price (under par), and the government debt must be chosen such that its expected repayment equals the banks' liabilities towards the government. If we require the debt contracts between the banks and the government to be budget-balanced in expectation, the implementation could be as follows: Assume that the government defaults with probability ρ and repays nothing in this case. The government still holds a claim of $d = \pi Re_0$ against the banks, whereas the banks hold claims with a face value of $b = d/\rho = \pi Re_0/\rho$ against banks. When selling these claims to investors, the fair value is given by $d = \pi Re_0$.

Let us now relax the exogeneity assumption and go to the other extreme. Assume that the government can only repay its debt if banks are solvent and thus fully serve their liabilities towards the government d. We thus relax the assumption that the government has access to exogenous funds, e.g., via taxation. In this setup, all features of the setup without government bonds reappear. The equilibrium of the $t = 1$ subgame still exists in which investors roll over debt, and one equilibrium where they do not. The whole game thus still has a subgame-perfect equilibrium that implements the first-best. However, this equilibrium is not unique – there also exists an equilibrium of the $t = 1$ subgame where a rollover freeze is accompanied by a government default. This induces banks to refrain from implementing the optimal allocation and the D&D allocation is implemented instead. In this case, the public provision of liquidity cannot help to overcome the coordination problem. The fragility of an economy in which both the solvency of banks and that of the government are endogenous and interdependent is discussed in more detail in Chapter 2.

1.6 Conclusion

This chapter has two main results: First, liquidity management with privately produced assets is either inefficient or associated with rollover risk, which makes an economy inherently fragile. Second, financial intermediaries can implement the optimal allocation by using government bonds.

In the absence of public liquidity, financial intermediaries face a trade-off between high investment, which goes along with high rollover risk (high level of illiquid, but profitable long-term investments and low level of storage), and low investment, which comes with low rollover risk (low levels of profitable long-term investments and high level of storage). For the case of the 2007-09 financial crisis, our model suggests that intermediaries chose high investment levels that created a strong maturity mismatch. In the run-up to the crisis, financial intermediaries transformed long-term real investments into short-term securities, thereby aiming at making them liquid. E.g., illiquid assets like ABS and MBS (which are securitized long-term real investments) were transformed into short-term securities such as ABCP. In the crisis, however, these short-term securities stopped being liquid and adverse consequences of the large-scale maturity mismatch realized.

Importantly, the rollover risk in our model is different from the traditional bank run problem in the style of Diamond and Dybvig (1983). In D&D, the bank run problem can be addressed by contracting in the initial period, e.g., by implementing a deposit insurance or allowing for a suspension of convertibility. In contrast, the rollover risk in our model cannot be eliminated in such a way because this would require contracting with a party that is not available initially. The problem originates from the friction that investors cannot commit initially to provide liquidity later. Their endowment does not realize before the time at which financial intermediaries need liquid funds. This makes liquidity management with privately produced assets inherently fragile.

In the second part, we demonstrate how liquidity management with government securities can improve the efficiency and stability of an economy. The government has the unique ability to commit future resources via taxation. Therefore, government securities are – in comparison to privately produced assets – less prone to coordination failures, i.e., less exposed to rollover risk. This property makes public outside liquidity superior to private outside liquidity.

This chapter thereby also contributes to the following basic but yet unresolved question: How should a public authority deal with liquidity provision? The traditional view since Bagehot (1873) is that a government should lend to illiquid but solvent institutions at high rates, while refusing to lend to insolvent institutions. The implementation of this principle might not be straight forward. That is, it may generally be problematic to identify whether an institution is illiquid or insolvent (see, e.g., Rochet and Vives, 2004). Moreover, Farhi and Tirole (2012) point out that if a government commits to intervening in case of liquidity needs, a collective moral hazard may give rise to an overall excessive maturity mismatch in an economy.

This chapter discusses a different approach to this problem. We find that the government should ensure efficient liquidity management by issuing government securities that should be held by financial intermediaries. This is a simple way of circumventing the undesired consequences that may arise when the regulator insures bank liabilities or provides emergency liquidity in case of a financial crisis. In our model, financial intermediaries are in fact always willing to hold government bonds to manage liquidity. However, in a richer model, banks may prefer to hold privately produced assets if these assets promise a higher return than government securities. In this case, a regulator might optimally force banks to hold government securities in order to enhance stable liquidity provision.

Appendix 1.A Sub-Optimality of Financial Markets under Outside Liquidity

Let us analyze the competitive equilibrium of an economy where consumers hold assets directly and trade on financial markets. In the D&D model, the first-best cannot be achieved via financial markets because, in equilibrium, the competitive market prices give consumers investment incentives that do not induce the investment profile that optimally trades off early liquidity needs and the returns of the long asset. This inefficiency arises because consumers do not take into account the pecuniary externalities of their investment. This is induced by the friction of unobservable liquidity needs (unobservable types).

In our setup, pecuniary externality do not necessarily arise. The mere unobservability of the liquidity type itself does not impede the implementation of the first-best. In the absence of return risk, each consumer could privately invest his whole endowment in the long asset and sell claims on the asset in the interim period. The key frictions in our case are that, on the one hand, the return risk cannot be diversified on the individual level and, on the other hand, the unobservability of the risky return induces adverse selection, leading to inefficient liquidation. The liquidity type friction only intensifies the return risk friction.

The first-best can only be implemented if all resources are initially invested in the long asset. Furthermore, consumers must be perfectly insured against return risk. However, this insurance is only implementable with commitment in period zero. In a pure financial market economy with spot markets for claims on future returns, contingent contracts are not feasible. Therefore, the financial markets cannot implement an allocation that is efficient ex-ante. The non-diversifiable idiosyncratic return risk might also imply that consumers invest a positive fraction of their endowment in storage.

Moreover, the financial-market allocation might not even be efficient ex post, i.e., given the initial private investment. If all consumers in fact invest in the long asset privately, all impatient consumers have to either sell or liquidate their asset in the interim period. Since the return is unobservable, all assets have to be sold at the same price R^*. All impatient consumers with a return $R_i > R^*/\ell$ have an incentive to liquidate instead of selling claims, reducing the average quality in the market. The liquidation of projects is a form of adverse selection and constitutes an inefficiency ex post. Moreover, patient consumers with a return $R_i < R^*$ have an incentive to sell at price R^* instead of waiting, thus exacerbating the adverse selection and inefficiency.

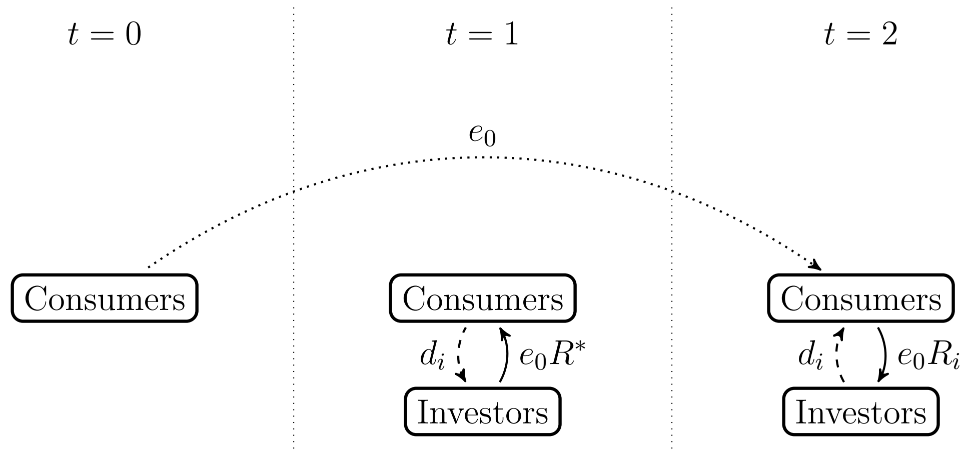

Figure 1.4: Financial Market. The graph illustrates investment and trade under the assumption that full investment in the long asset is chosen. In the interim period, the graph denotes the flow of goods and claims for a consumer who chooses to sell claims on the market. Notice that there are also agents who liquidate their assets or hold on to them until $t = 2$ and thus do not interact with investors.

Appendix 1.B Sunspots: Optimal Contracts under Private Outside Liquidity

Assume that there is a public signal in $t = 1$ which we might call sunspot, following a Bernoulli distribution with success probability p. Assume further that investors play a pure strategy by which they base their rollover decision on this public signal. With probability p, all investors refrain from rolling over, and with probability $1 - p$, all investors engage in rollover. We are now looking for the optimal consumption profiles of consumers, i.e., the optimal investment behavior of banks, and abstract from bank runs. The optimal contract between banks and consumers will be state-contingent, i.e., contingent on the sunspot, or equivalently, on the behavior of investors.

Define the investment threshold I_ℓ such that

$$\ell u'\left(\frac{e_0 - I_\ell}{\pi}\right) = u'\left(\frac{RI_\ell}{1-\pi}\right). \tag{1.11}$$

This threshold has the following interpretation: It is ex-post optimal to liquidate a positive fraction in case of a rollover freeze if and only if investment exceeds this threshold, $I > I_\ell$. Because $\ell < 1/R$, it holds that $I_\ell \in (I^{DD}, e_0)$, i.e., liquidation is not efficient ex post in a setup without outside liquidity (D&D), but it is efficient if banks do not invest in storage and completely rely on rollover, but this investors coordinate on a rollover freeze.

For the contingency of rollover, the optimal contract includes perfect consumption smoothing through outside liquidity and no liquidation, $c_1 = c_2 = e_0 - I + RI$. For the contingency of a rollover freeze, there is positive liquidation if $I > I_\ell$, and there is no liquidation if $I \leq I_\ell$. The optimal investment level I is determined by the probability of the rollover freeze.

We first derive the optimal consumption allocation in case of a rollover freeze for a given investment level I. The optimization problem is given by

$$\max_{c_1,c_2,z\in[0,1]} \pi u(c_1) + (1-\pi)u(c_2), \tag{1.12}$$

$$\text{s.t.} \quad \pi c_1 \leq e_0 - I + z\ell RI, \quad \text{and} \tag{1.13}$$

$$\pi c_1 + (1-\pi)c_2 \leq e_0 - I + z\ell RI + (1-z)RI. \tag{1.14}$$

The aggregate budget constraint (1.14) is always binding. There exists a threshold $I_0 < I^{DD}$ such that for $I < I_0$ it holds that $z = 0$ and the budget constraint for period one is not binding, leading to perfect consumption smoothing. For $I_0 \leq I \leq I_\ell$ it holds that $z = 0$, and the first period budget constraint is binding, implying that $c_2 > c_1$. For $I > I_\ell$ it holds that $z(I) \in (0,1)$ such that

$$\ell u'\left(\frac{e_0 - I + z\ell RI}{\pi}\right) = u'\left(\frac{(1-z)RI}{1-\pi}\right). \tag{1.15}$$

We now have (implicitly) specified the optimal contingent consumption profiles given an investment level I. We now maximize over this investment level. We split the problem by looking at the maximizing level of investment within each of the two intervals $[I_0, I_\ell]$ and $(I_\ell, e_0]$. We can ignore the interval $[0, I_0)$ because it is dominated by I_0.

There exist two thresholds p_0 and p_ℓ, $0 < p_0 < p_\ell < 1$, such that $I(p) = e_0$ iff $p \leq p_0$, and $I^*(p_\ell) = I_\ell$.
If $p \in (p_0, p_\ell)$, the optimal $I^*(p) \in (I_\ell, e_0)$, and I^* and z are determined by

$$\begin{aligned} 0 = \quad & p\left[(z\ell R - 1)u'\left(\frac{e_0 - I^* + z\ell RI}{\pi}\right) + (1-z)Ru'\left(\frac{(1-z)RI^*}{1-\pi}\right)\right] \\ & + (1-p)(R-1)u'(e_0 + (R-1)I^*), \quad \text{and} \end{aligned} \tag{1.16}$$

$$\ell u'\left(\frac{e_0 - I + z\ell RI}{\pi}\right) = u'\left(\frac{(1-z)RI}{1-\pi}\right). \tag{1.17}$$

If $p > p_\ell$, the optimal $I^*(p) \in [I^{DD}, I_\ell)$ and its level is determined by

$$p\left[-u'\left(\frac{e_0 - I^*}{\pi}\right) + Ru'\left(\frac{RI^*}{1-\pi}\right)\right] + (1-p)(R-1)u'(1+(R-1)I^*) = 0. \tag{1.18}$$

For any $p \in (p_\ell, 1)$, it holds that $I^*(p) \in (I^{DD}, I_\ell)$. As $p \to 1$, it holds that $I^*(p) \to I^{DD}$.

2
Sovereign Defaults, Bank Runs, and Contagion

2.1 Introduction

In this chapter, we provide a model that unifies the notion of self-fulfilling banking crises and sovereign debt crises. We show how these crises can be contagious, i.e., how a bank run can trigger a sovereign default, and vice versa (first type of contagion). We discuss under which conditions a government is unable to eliminate self-fulfilling banking crises by implementing a deposit insurance scheme. Moreover, we illustrate how crises can be contagious across countries (second type of contagion), and how contagious crises can be prevented. This allows us to evaluate the efficacy of recent policy proposals for the implementation of banking union in the euro area. We show under which conditions a supranational Deposit Guarantee Scheme can eliminate self-fulfilling crises at not cost.

The sovereign debt crisis in the euro area which has accompanied and followed the recent financial crisis since early 2009 has made the interdependence of sovereign and financial stability a prominent topic in the academic and political debate. Farhi and Tirole (2014) state that danger of the feedback loop between banking crises and sovereign debt crises is an exceptionally uncontroversial economic idea. Several terrifying terms have been invented invented to describe this phenomenon, like “vicious cycle”, “doom loop”, “diabolic loop”, or “deadly embrace”.

However, this phenomenon is anything but new. Historically, sovereign defaults and banking crises have often preceded and accompanied each other (see, e.g., Reinhart and Rogoff, 2009, 2011), but most existing data concerns emerging economies. Furthermore, there have been surprisingly few formal models that help to guide our theoretical understanding of how sovereign defaults and banking crises are interrelated, in particular for the case of developed and highly leveraged economies. Only recent, theoretical models on this topic were provided, e.g., by Acharya et al. (2014), Farhi and Tirole (2014), Leonello (2013), Cooper and Nikolov (2013), and König et al. (2013).

Banking crises and sovereign debt crises have the common feature that they may result from coordination on a bad equilibrium. In a self-fulfilling bank run, depositors desire to withdraw all at once. This is an equilibrium because if all depositors desire to withdraw at once, it forces an otherwise solvent bank to engage in inefficient liquidation, leading to insolvency (see, e.g., Diamond and Dybvig, 1983; Goldstein and Pauzner, 2005). In a self-fulfilling sovereign debt crisis, investors roll over a sovereign's debt only at a high risk premium, or even refuse to do so. This constitutes an equilibrium as the high sovereign risk premium increases the government's debt burden and thereby the likelihood of a default (see, e.g., Calvo, 1988; Cole and Kehoe, 2000).

We present a simple banking model of maturity transformation in the tradition of Diamond and Dybvig (1983). In the first part of this chapter, we consider the case of a closed economy. The model is reduced to a two-period version (we do not model the investment stage) and features consumers, banks, investors, and a government. We make two key assumptions: First, banks hold government bonds that they can sell in a secondary market in order to manage the liquidity needs of consumers. Second, the government's tax base is correlated with the real economic activity which in turn depends on the performance of the financial sector. The model features a strategic complementarity within the consumers' withdrawal decision, within the investors' decision to purchase government bonds, as well as across the decisions of the two types of agents. There exist two types of self-fulfilling equilibria in our model: The first one is a *no-crisis equilibrium*, in which government bonds trade at face value, and the government as well as the banks fulfill their obligations. The second one is a *crisis equilibrium*. In the crisis equilibrium, all consumers withdraw early, causing a bank run. Depending on the fiscal soundness of the government, a bank run can be accompanied by a rollover freeze and a sovereign default. If the government is fiscally weak, a banking crisis and a sovereign default aggravate and reinforce each other in a "vicious circle". Only if a government is fiscally strong, it can eliminate the crisis equilibrium by providing a deposit insurance.

In the second part of this chapter, we extend our model to a multiple country setup where countries are interdependent, and we analyze cross-country effects of banking crises and sovereign debt crises. We assume that countries are interdependent due to banks diversifying their government bond holdings. If countries are sufficiently interdependent, self-fulfilling twin crises are contagious across borders. We show that if one country is fiscally weak while the other country is fiscally sound, it may be beneficial for both countries to pool their funds. The crisis equilibrium and its adverse consequences can be ruled out ex-ante by the following policy: Both countries form a banking union that implements a supranational deposit insurance scheme, and potentially also a fiscal union. By committing to repay the sovereign debt and to provide deposit insurance jointly, their joint promise will never be tested in equilibrium and is thus costless. A crucial insight is that forming such a union is not only beneficial for the fiscally weak country, but also for the fiscally strong country.

Guided by the insights of the model, we discuss two policy implications. The first policy implication concerns the design of the European Banking Union, with a special focus on the deposit insurance. Our model features cross-border costs of banking crises and sovereign defaults and points out channels through which a crisis in one country can trigger a crisis in another country. This in turn allows rationalizing policy responses by countries that are affected by foreign banking crises or sovereign defaults. The model allows us to give conditions under which a banking union (i.e., a joint deposit insurance) or the combination of a banking and a fiscal union can prevent contagious self-fulfilling banking crises and sovereign defaults. The model hence sheds light on the policy debates following the European debt crisis and allows us to investigate the efficacy of recent policy proposals (European Commission, 2013a). These proposals for a banking union focus on the Single Supervisory Mechanism (SSM) and the Single Resolution Mechanism (SRM). A supranational Deposit Guarantee Scheme (DGS) which would take the current national deposit insurance to a supranational level seems to be politically infeasible so far. By considering the self-fulfilling nature of banking crises, we show to what extend a banking union in its current form is ineffective at preventing such crises. Given that there are differences in the fiscal soundness of its member states, we argue that a banking union might only be effective if it comes with a joint deposit insurance.

The second policy implication concerns the regulatory treatment of banks holding government bonds. While there may be good reasons for banks to use government bonds as an instrument to manage liquidity needs,[1] we show that this may also be a considerable source of fragility once there is a prospect of a government default. Fragility arises in our setup whenever the government's ability to repay its debt depends on the performance of the financial sector. This condition may be satisfied in developed economies that have highly leveraged financial systems. This chapter can therefore also be understood

[1]See, e.g., Holmström and Tirole, 1998, Gorton and Ordoñez, 2013, and Chapter 1.

as a contribution to the debate concerning the liquidity regulation of banks. Regulatory frameworks typically facilitate the holding of government debt by intermediaries. The Basel Committee on Banking Supervision initially refrained from imposing any capital requirement for government bond holdings (see, e.g., Goodhart, 2011). Positive risk weights for poorly rated government bonds have been put on the agenda only recently, and were introduced in Basel III (Basel Committee, 2011). Our model provides an argument for why the exposure of banks to sovereign debt is a severe problem that is not adequately dealt with under both the current and the currently planned bank regulation.

This chapter is structured as follows: Section 2.2 presents a model of a closed economy, derives the equilibria, and discusses the effect of a deposit insurance. In Section 2.3, the model is extended to a two-country setting with international integration. We analyze contagion across countries and discuss optimal crisis prevention policies. Section 2.4 relates our findings to the current debate about the European Banking Union.

Related Literature

This chapter reaches out to the large literature on self-fulfilling banking crises (see, e.g., Diamond and Dybvig, 1983; Rochet and Vives, 2004; Goldstein and Pauzner, 2005) and self-fulfilling sovereign debt crises (see, e.g., Calvo, 1988; Alesina et al., 1990; Cole and Kehoe, 2000), and attempts to unify some aspects of the two strands.

The first part of this chapter is very closely related to a series of recent papers that model banking crises and sovereign debt crises in unified frameworks (Cooper and Nikolov, 2013; König et al., 2013; Leonello, 2013). Cooper and Nikolov also provide a model with multiple equilibria where the adverse equilibrium is characterized by a vicious cycle in which a government debt crisis and a banking crisis aggravate and reinforce each other. However, their focus is on the pricing of government debt, while emphasize the strategic complementarity of agents. The papers by König et al. and Leonello provide models featuring unique equilibria – reminiscent of Goldstein' (2005) twin crisis model – and they analyze how government guarantees affect financial stability and the government's ability to fulfill its obligation. All three papers have in common that the contagion from a banking crisis to a sovereign default originates from the increased public liabilities that arise from a safety net. In contrast, contagion in our setup arises because a financial crisis reduces the government's tax base and thus decreases its funding instead of increasing its expenditure. The channel from sovereign debt to banking crisis is similar, however, it results from banks hold government bonds.

With Acharya et al. (2014), we share the notion that the government's tax base is limited by a Laffer-curve property. Unlike our approach, they focus on the optimal redistribution (bailout) between a financial sector with debt overhang and a corporate sector. They find that a bailout can lose its bite if it lowers the value of government bonds that are held by the financial sector.

In the second part of this chapter, we analyze how crises can be contagious across countries. This part is related to the literature on financial contagion and the spreading of banking panics (see, e.g., Allen and Gale, 2000; Dasgupta, 2004). In particular, the second part of this chapter relates to Bolton and Jeanne (2011) who analyze the cross-border effects of sovereign defaults in financially integrated areas. In their model, government debt is used as collateral in interbank markets. Economic integration is beneficial as banks can diversify their government bond holdings, which fosters welfare-increasing interbank trade. However, this comes with possible contagion of a sovereign default ex-post, and fiscally strong countries might suffer from fiscal integration. This chapter is concerned with maturity transformation by banks and its inherent fragility, and not with the banks' role in allocating capital. Moreover, government defaults are endogenous in our setup and directly linked to the performance of the banking sector. In contrast to the results of Bolton and Jeanne, we find that fiscally strong countries might actually benefit from fiscal integration if this prevents self-fulfilling crises.

Farhi and Tirole (2014) consider a model featuring fundamental financial and fiscal shocks in which banks hold domestic and foreign government bonds. Banks have an incentive to engage in excessive risk taking, particularly in collective moral hazard because the national government cannot commit to refrain from bailouts. This provides a new argument in favor of a banking union because the government is better off by delegating regulation to a supranational supervisor who takes a tough ex-post regulatory stance.

2.2 Single-Country Model

2.2.1 Setup

Consider an economy that goes through a sequence of two dates, $t \in \{1, 2\}$. The economy is populated by a continuum of consumers of mass one and a continuum of investors of mass one. Moreover, there is a banking sector and a government. There exists a single good that can be used for both consumption and investment, and all units are denoted in terms of this good.

Consumers

Each consumer i is endowed with a demand deposit contract (c_1^*, c_2^*) that allows her either to withdraw c_1^* units from her bank account in $t = 1$ or c_2^* units in $t = 2$. Consumers have preferences as proposed by Diamond and Dybvig (1983). There are two types: a fraction $\pi \in [0, 1]$ of consumers is *impatient*, while the remaining fraction $(1 - \pi)$ is *patient*. Impatient consumers only derive utility from consuming early; their utility is given by $u(c_1)$. Patient consumers are indifferent between consuming early and late; their utility is given by $u(c_1 + c_2)$. Types are private information of each consumer. Consumers face the decision to withdraw and to consume in $t = 1$ or to withdraw and consume in $t = 2$. Notice that the attributes "patient / impatient" characterize the consumer's exogenous types. In contrast, the attributes "late / early" will characterize the endogenous decision of consumers: an "early consumer" withdraws and consumes in $t = 1$, while a "late consumer" withdraws and consumes in $t = 2$. We denote the decision of each consumer i to withdraw as well as to consume early with $\omega_i \in \{0, 1\}$, where ω_i takes the value one if consumer i withdraws in $t = 1$. Let $\omega = \int_0^1 \omega_i di$ be the aggregate mass of early consumers.

Banking Sector

There is a banking sector that has the demand deposit contracts – which are the assets of consumers – as liabilities. It owns two types of assets: it holds government bonds as well as an illiquid portfolio of loans, both maturing at $t = 2$.

Banks are assumed to hold government bonds for the purpose of liquidity management. While we are not giving a micro-foundation for why banks are holding government bonds, we refer to various arguments for why financial intermediaries use government securities for liquidity management.

Government bonds are valuable as a medium of transfer across time (see, e.g., Gale, 1990; Woodford, 1990), and private agents may not be able to provide sufficient pledgable income (Holmström and Tirole, 1998). Furthermore, government securities – unlike private assets – are not subject to adverse selection (Gorton and Ordoñez, 2013), and government securities are simply less exposed to rollover risk than privately produced assets (compare Chapter 1).

In our model, banks are not considered to be agents. They behave mechanically in that they serve early-withdrawing consumers by selling government bonds to investors and by liquidating the illiquid assets if necessary. Having the demand deposit contracts as liabilities, banks need to serve a mass ω of consumers with c_1^* units in $t = 1$ each, and a mass $1 - \omega$ of consumers with c_2^* units in $t = 2$. Banks own a stock of government bonds which mature in

$t = 2$. Bonds are liquid in the sense that they may be sold to investors in $t = 1$. Selling these government bonds allow banks to fulfill their short-term liability, i.e., to serve early consumers. The total amount of government debt in the economy is given by B, and banks own a fraction α of them, i.e., they own $\alpha B < B$ units of government bonds. One unit of the government bond is a promise of the government to repay one unit of the good in $t = 2$. Details of the government bonds will be further specified below.

Moreover, banks also own I units of an illiquid asset to serve their long-term liabilities. The illiquid asset has an after-tax return of $r = (1 - \tau)R > 1$ in period two. The asset can be liquidated in $t = 1$, yielding a return per unit of $\ell < 1$. The fraction of illiquid assets which banks liquidate is denoted by z. The total return of liquidation is thus given by $z\ell I$. As indicated, the illiquid asset can be interpreted as a loan portfolio which pays off in the long run. In the short run, it can be liquidated at a substantial discount. The liquidation value ℓ can be interpreted as the price in the secondary market for the bank's loan portfolios and the discount may result from various frictions we do not model.[2]

Government

There is a government that has an outstanding amount of debt B, maturing in $t = 2$. Like banks, the government is assumed to behave mechanically. The government always repays its debt if possible and defaults otherwise. In $t = 2$, the government has an overall tax revenue of $T(z) = E + \tau(1 - z)RI$ at its disposal. It consists of an exogenous tax revenue of $E \geq 0$, and an endogenous tax revenue $\tau(1 - z)RI$ from taxation of the illiquid technology of the banking sector, where $\tau \in (0, 1)$ is fixed. The tax revenue is used for the repayment of the government's debt.[3]

We interpret the exogenous tax revenue E as the tax revenue that the government generates irrespective of the performance of the banking sector. In turn, the endogenous tax revenue displays the fiscal revenue that depends on the performance of the banking sector and thus decreases in the level of liquidation z. It should thus be interpreted as the taxable economic activity that is generated through successful intermediation by banks. We assume that the government cannot raise any taxes in $t = 1$. This clearly displays an

[2]The assumption of low liquidation values is standard in the banking literature and may result from moral hazard (Holmström and Tirole, 1997), limited commitment of future cash-flows (Hart and Moore, 1994), adverse selection (Flannery, 1996), or uncertainty-averse investors (Uhlig, 2010).

[3]The remaining government budget can be used for other purposes. It could be used to provide a public good, or it could be transferred to the consumers. The exact use of remaining funds is not relevant in our model.

extreme simplification. However, we argue that a government's ability to raise taxes at any point in time has natural limits,[4] and we make the simplifying assumption that it is zero in the short run.

Importantly, we assume that the government repays its debt whenever $B \leq T(z)$. For simplicity we assume that it fully defaults otherwise. With this assumption, we deviate from large parts of the sovereign risk literature and completely abstract from willingness to pay considerations.[5] However, we refer to recent contributions arguing that ability-to-pay constraints dominate willingness-to-pay considerations, especially in advanced economies with a high degree of leverage where defaults may trigger severe financial sector turmoil (Gennaioli et al., 2014; Acharya and Rajan, 2013). If the government cannot default selectively (Guembel and Sussman, 2009; Broner et al., 2010), its incentives to default are generally very weak whenever the costs of defaulting are very high for domestic creditors. Thus, a sovereign default in a leveraged economy is likely to result from a binding ability to pay constraint.

Investors

There is a continuum of investors of mass 1. Each investor j is equipped with one unit of the good in $t = 1$. Investors are risk-neutral and do not discount. Investors buy government bonds from banks whenever their return is non-negative. Formally, the decision of an outside investor j to purchase government bonds from banks at face value or not is denoted $\eta_j \in \{0, 1\}$. It takes the value one if she is willing to buy a government bond at a price of one. Let $\eta = \min[\alpha B, \int_0^1 \eta_j dj]$ be the aggregate mass of outside investors that buy government bonds at face value from banks.

In the following, we will refer to the purchase of government bonds by investors as *rollover*. Note that, in our setup, it does not matter whether the government needs to borrow $t = 1$ in order to repay banks that hold bonds that mature in $t = 1$ or whether banks need to sell government bonds that mature in $t = 2$ in a secondary market in $t = 1$. The first scenario clearly looks like a classical rollover problem. As both scenarios are equivalent, we use the expression rollover in order to simplify the wording.

[4]See, e.g., the Laffer-curve property in Acharya et al. (2014).

[5]The literature on sovereign debt and risk has been shaped by the willingness to pay view, which argues that governments repay their debt when the costs of repayment are lower than the penalty expected for default. In the literature, default penalties have been argued to be, e.g., exclusion from capital markets or trade sanctions (see, e.g., Eaton and Gersovitz, 1981; Bulow and Rogoff, 1989).

Parameters
In the following, we make some restrictions on the model's parameters in order to ensure outcomes and effects in a relevant domain. The first three assumptions guarantee the existence of a no crises equilibrium (also referred to as type I equilibrium), while the last assumption ensures the existence of a crisis equilibrium (type II equilibrium).

Assumption 2.1. $c_2^* \geq c_1^*$

Assumption 2.1 guarantees that it is incentive-compatible for patient consumers to withdraw late and to consume in $t = 2$ conditional on banks being able to pay out their promised payment, i.e., conditional on no liquidation.

Assumption 2.2. $\pi c_1^* = \alpha B \leq 1$ *and* $rI = (1 - \pi)c_2^*$

The first equation of Assumption 2.2 ensures that banks can serve all impatient consumers by selling their government bond holdings at face value. The second equation ensures that all patient consumers can be served by the long-term return of the loan portfolio if they withdraw late. Moreover, $\alpha B \leq 1$ implies that investors have enough funds to purchase all government bonds from banks at face value.

Assumption 2.3. $T(0) = E + \tau RI \geq B$

Assumption 2.3 ensures that the government's tax revenue is sufficient to repay the government's debt given that there is no liquidation by banks.

Assumption 2.4. $(1 - \pi)c_1^* > \ell I$

Assumption 2.4 implies that the banks will be insolvent and illiquid in $t = 1$ in case all consumers withdraw early, irrespective of the government's solvency. The reason is that liquidation is sufficiently inefficient for a panic-based bank run to exist. While the patient consumers' claims might be met by selling the government bonds, Assumption 2.4 implies that if all patient consumers withdraw early, their claims equal to $(1 - \pi)c_1^*$ cannot be met by proceeds of complete liquidation, ℓI. That is, the banking sector will be illiquid and insolvent in $t = 1$ whenever there is complete withdrawal and liquidation.

2.2.2 Outcomes

In the following section, we show that the economy described above has two equilibria in pure strategies: a no crisis (type I) equilibrium and a crisis (type II) equilibrium. In the no-crisis equilibrium, only impatient consumers withdraw early and outside investors roll over the government's debt. In the crisis equilibrium, all consumers withdraw early, causing a bank run. Depending on the fiscal soundness of the government, a bank run can be accompanied by a sovereign default and a rollover freeze.

In order to derive the equilibrium outcomes, we first analyze the banks' liquidation of the loan portfolio for any given level of aggregate withdrawal and any rollover decision. We can then calculate the value of the demand deposit contract, as well as the value of government bonds in $t = 2$, as functions of aggregate withdrawal and rollover. This in turn will pin down the optimal individual withdrawal and rollover decisions in $t = 1$.

Liquidation

Banks have to fulfill their obligations in $t = 1$ whenever possible. Recall that ω denotes the mass of consumers that withdraw early, and η the mass of investors purchasing government bonds at face value. Banks need liquid funds of ωc_1^* in $t = 1$, since they have to pay c_1^* units of the good to a mass ω of consumers. Banks sell η units of the governments bonds to investors. Given ω and η, banks must liquidate a fraction z such that their liquid funds equal the demand for early consumption or engage in complete liquidation, $z = 1$, otherwise. Liquidation z is implicitly given by the budget equation $\omega c_1^* = \eta + z\ell I$ whenever feasible, or explicitly by

$$z(\omega, \eta) = \min\left[1, \frac{[\omega c_1^* - \eta]^+}{\ell I}\right]. \tag{2.1}$$

If banks can serve all withdrawing consumers by selling government bonds, liquidation is unnecessary. However, if the proceeds from selling government bonds are not sufficient to serve all withdrawing consumers, banks will have to engage in inefficient liquidation of the loan portfolio.

Withdrawal and Rollover

The individual decision of patient consumers to withdraw depends on the funds that banks have available in $t = 2$. Similarly, the decision of investors to purchase government bonds depends on the funds that the government has available in $t = 2$. Whenever there is liquidation, the amount left for late consumers and the tax revenue of the government decrease.

The deposit contract (c_1^*, c_2^*) is characterized by *promised* payments. If there is liquidation, *actual* repayments (c_1, c_2) may fall short of the promised levels. In period one, banks have to serve any withdrawing consumer with c_1^* whenever possible. If banks engage in liquidation, this reduces the level of late consumption c_2, and if consumers in addition start to run on the banks, this also reduces c_1. For impatient consumers, it is dominant strategy to withdraw early, implying $\omega \in [\pi, 1]$. Given liquidation $z(\eta, \omega)$, the payments made to each patient consumer who is withdrawing late is given by

$$c_2(z(\omega, \eta), \omega) = (1 - z(\omega, \eta)) \frac{1 - \pi}{1 - \omega} c_2^*. \tag{2.2}$$

A patient consumer only withdraws early if $c_2 < c_1^*$.[6] The optimal withdrawal decision of a patient consumer is therefore given by

$$\omega_i^*(\omega, \eta) = \begin{cases} 0 & \text{if } c_2(z(\omega, \eta)) \geq c_1^* \\ 1 & \text{if } c_2(z(\omega, \eta)) < c_1^*. \end{cases} \tag{2.3}$$

We can derive the optimal rollover decision in a similar fashion. Given $z(\eta, \omega)$, the government has a tax revenue of

$$T(z(\omega, \eta)) = E + \tau(1 - z(\omega, \eta))RI. \tag{2.4}$$

The government repays its debt whenever the tax revenue $T(z(\eta, \omega))$ exceeds the government's outstanding debt B, and defaults otherwise. Investors purchase government debt at face value if the government will be able to repay its debt, and do not purchase if the government is expected to default. An investor's rollover decision is thus given by

$$\eta_i^*(\omega, \eta) = \begin{cases} 1 & \text{if } B \leq T(z(\omega, \eta)) \\ 0 & \text{if } B > T(z(\omega, \eta)). \end{cases} \tag{2.5}$$

The interrelation of the model's key variables is summarized in Figure 2.1. The left cycle is the well-known cycle that lies at the heart of a self-fulfilling bank run, as in the classic bank-run model by Diamond and Dybvig (1983): Increased liquidation lowers the level of funds available for late consumption. This in turn increases the incentive to withdraw early. High early withdrawal, however, further increases liquidation.[7] The right cycle shows how an anticipated sovereign default can be self-fulfilling: The inability of banks to sell

[6] We define $c_2(1, 1) := 0$, i.e., the potential late consumption is zero in case of complete liquidation.

[7] For the sake of completeness, the dotted arrow represents a positive feedback effect of early withdrawal on late consumption: more consumers withdrawing early implies that the remaining available funds are distributed among a smaller mass of late consumers. This channel represents the same effect through which a bank run is welfare-increasing in Allen and Gale (1998).

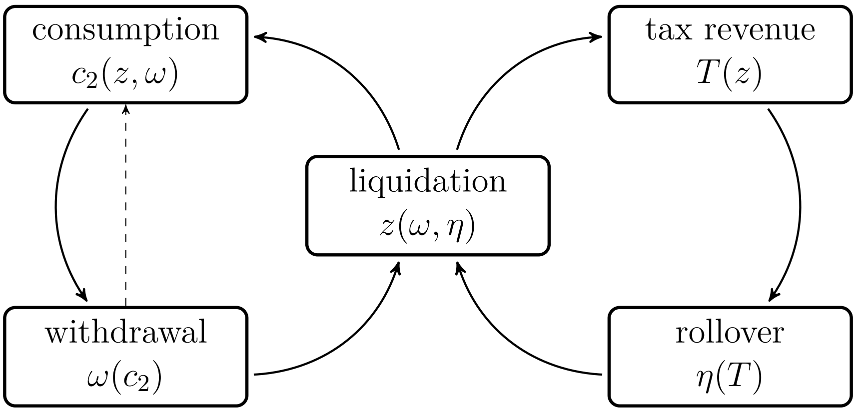

Figure 2.1: Interdependence of Sovereign Debt and Banking.

government bonds forces them to liquidate some of the loan portfolio. Liquidation reduces the tax base and thus future tax revenue. This in turn may reduce the amount the government can repay. Consequentially, investors may become unwilling to purchase government bonds, forcing banks to liquidate even more.

The two cycles are connected through the liquidation of the illiquid loan portfolio. This allows a banking crisis to be contagious by triggering a sovereign debt crisis, and vice versa (1st type of contagion).

Definition 2.1. A Nash equilibrium in pure strategies is given by a set of consumers' withdrawal decisions $\{\omega_i\}$ and outside investors' rollover decisions $\{\eta_j\}$, such that these decisions are best responses, i.e., $\omega_i = \omega_i^*(\omega, \eta)\ \forall i$ and $\eta_j = \eta_j^*(\omega, \eta)\ \forall j$, where $\omega = \int_0^1 \omega_i di$, and $\eta = \min[\alpha B, \int_0^1 \eta_j dj]$.

We are now equipped in order to formulate the first result:

Proposition 2.1. *The model has two Nash equilibria in pure strategies.*

a) ***Type I equilibrium*** $(\omega, \eta) = (\pi, \alpha B)$*:*
Only impatient consumers withdraw early, banks do not liquidate, the tax revenue is sufficient to repay creditors, and investors are willing to buy government bonds.

b) ***Type II equilibrium***

$E < B$ ***Sovereign default and bank run*** $(\omega, \eta) = (1, 0)$*: All consumers withdraw early and there is no rollover, inducing full liquidation. This results in illiquidity and insolvency of both the government and the banking sector.*

$E \geq B$ ***Bank run*** $(\omega, \eta) = (1, \alpha B)$*: Investors roll over government debt, but all consumers withdraw early. Although there is a bank run and full liquidation, the government is still able to fully serve its debt.*

For the proof of Proposition 2.1, see the Appendix. The multiplicity of equilibria arises from the strategic complementarity between agents. There are three different components of strategic complementarity in the model. First, there is a strategic complementarity between consumers in their decision to withdraw: the more consumers withdraw, the higher the incentive for an individual consumer to withdraw as well. Second, there is strategic complementarity between the investors in their decision to purchase government bonds: more investors purchasing government bonds increases the individual incentive to purchase government bonds as well. Third, there is strategic complementarity across the two types of agents: higher levels of withdrawal decrease the incentive to roll over and vice versa.

Note that in the above setup both types of equilibria always exist. The type I equilibrium is always characterized by successful debt rollover and the absence of a panic-based bank run. The type II equilibrium is characterized by either a twin crisis where a sovereign default and a panic-based bank run accompany each other, or by a panic-based bank run without sovereign default. The type II equilibrium is a twin crisis whenever the exogenous tax base E is less than the government's outstanding debt B, or if $E/B < 1$. In this case, banking crises and sovereign debt crises are contagious in the sense that they aggravate and reinforce each other. Whenever E exceeds B, i.e., $E/B \geq 1$, the government will be able to repay its debt irrespective of the occurrence of a banking crisis. In this case, a sovereign default will never occur, but a bank run still constitutes an equilibrium.

The parameter E (or the ratio E/B) can be interpreted as a measure of the government's fiscal stability. If $E/B \geq 1$, the government can raise taxes irrespective of the performance of banks which will suffice to repay the outstanding debt. The taxable economic activity thus does not depend too much on the provision of financial services. If $E/B < 1$, the government's ability to tax and to repay is closely linked to the banking sector, i.e., the taxable economic activity depends strongly on the performance of the banking sector. Therefore, whenever $E/B < 1$, the crisis equilibrium is not only characterized by a banking crisis, but also by a sovereign default.

Generally, $E/B < 1$ is reminiscent of the *crisis zone* in Cole and Kehoe (2000): when the exogenous tax base that is available irrespective of the performance of the banking sector is low, runs become possible. In the following, we will show that this may be true irrespective of the existence of a deposit insurance.

2.2.3 Deposit Insurance Scheme

We now analyze the effect of a deposit insurance. We define deposit insurance to be a guarantee by the government that each consumer receives c_1^* units at a period of his choice. If the deposit insurance is credible, it prevents patient consumers from withdrawing early because in any contingency consumers get at least as much in period two as in period one. In the next paragraph, we will analyze under which conditions a deposit insurance is credible in our setup. We assume that the government uses its tax revenue to repay its bonds first, and only uses its remaining funds to fulfill the DIS afterwards if possible. This ordering might seem odd at first sight because government bonds only get repaid in period two, but the deposit insurance might already be needed in period one. However, since the government does not have funds in period one – recall that we assumed that the short-term tax base is zero – it will have to borrow in order to provide a DIS. The government will only be able to borrow and actually fulfill a deposit insurance if its outstanding debt is not already exceeding its available funds. Therefore, the government bonds are effectively senior to the deposit insurance.

The deposit insurance is credible if the government is able to repay its debt and to pay for the deposit insurance in any contingency. The most adverse contingency is the case in which all consumers withdraw early, and the banks thus have to engage in full liquidation. A sufficient condition for the deposit insurance to be credible is that the government can repay its debt B. Therefore, banks can sell their πc_1^* bonds at face value. The complete liquidation of the illiquid loan portfolio provides the banks with an additional amount of ℓI units. Thus, the deposit insurance has to cover the missing funds in order to serve the each consumer with c_1^* units. Therefore, the maximal amount a deposit insurance might have to cover is given by $DI = c_1^* - \ell I - \pi c_1^*$.

Whenever $E \geq B + DI$, the deposit insurance scheme is credible, because the government can actually provide this amount in any contingency.

Proposition 2.2. *By providing a deposit insurance scheme, the government can eliminate the crisis equilibrium iff $E \geq B + DI$.*

The government is able to eliminate the crisis equilibrium whenever its exogenous tax revenue exceeds the sum of the outstanding debt B and the maximum cost of a deposit insurance DI. In this case, it can repay its debt and credibly insure deposits of all consumers. The deposit insurance is never tested and therefore eliminates the adverse equilibrium at no costs. For $B \leq E < B + DI$, there are multiple equilibria. The government cannot prevent a bank run because the deposit insurance scheme is not credible, but since it can serve its debt, a rollover freeze does not occur in equilibrium. For $E < B$, the government can neither prevent a bank run nor a sovereign default. Figure 2.2 shows which type of equilibria exist for different levels of E under the deposit insurance scheme.

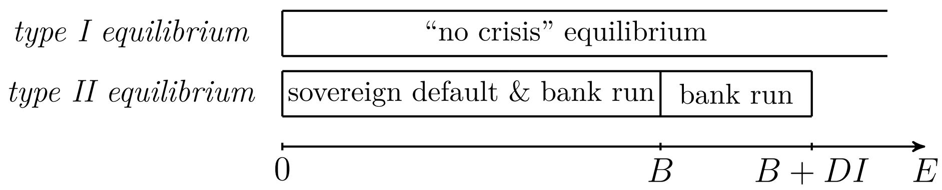

Figure 2.2: Existence of Equilibria under the Deposit Insurance Scheme. The type I equilibrium always exist. The type II equilibrium only exists if the deposit insurance is not credible, i.e., if $E < B + DI$. It is characterized by a bank run for $E \geq B$, and by a twin crisis for $E < B$.

Finally, notice that the deposit insurance does not bail out banks; it only steps in after banks have already defaulted on their liabilities. In fact, it would be more efficient in our model to bail out banks in order to prevent them from engaging in inefficient liquidation. However, in terms of preventing the crisis equilibrium, a bailout mechanism would have exactly the same effects as a DIS. The government could announce that it would bail out the banks in case of a crisis and thereby eliminate the crisis equilibrium if the announcement is credible. For this to be true, the government would need exactly the same budget, i.e., $B + DI$.

2.3 Two-Country Model

We now consider an extended, two-country setting of the model. This allows us to analyze under which conditions a crisis in one country may be contagious, triggering a crisis in another. We will use the setup to investigate which policies eliminate the adverse crisis equilibrium and ensure financial stability. In our model, a country consists of domestic consumers who hold demand deposit contracts with domestic banks. Furthermore, there is a government that taxes domestic economic activity. In our model, investors are not associated with countries.

Assume that there are two countries that are labeled home H and foreign F. Without loss of generality, we take the view of the home country to facilitate the verbal interpretation of our analysis. Both countries are as described in the single-country case and identical to each other, except for some international financial interdependence. Furthermore, we vary the amount of exogenous tax revenue E^H and E^F. A country k is called fiscally sound whenever E^k is very high, and fiscally weak whenever E^k is low. We assume throughout most of this section that each country implements a deposit insurance scheme targeting domestic depositors whenever feasible. We analyze a policy setup where both countries can form a banking union or a fiscal and banking union. The banking union is a supranational policy tool that implements a joint deposit insurance for both countries. When tested, the costs are borne by the two countries jointly. We contrast these policies with a situation of *political autarky* where there is no supranational policy. Throughout the analysis, we maintain Assumptions 2.1 to 2.4 for both countries.

Importantly, we assume that countries are interdependent. We introduce interdependency by assuming that banks of both countries hold government bonds of both countries. While we assume this interdependence, we refer to empirical evidence as well as to theoretical explanations why government bond holdings are diversified.[8]

In a nutshell, we will present two main results: First, crises can be contagious across countries once there is interdependence. A sovereign debt crisis in the foreign country is always costly for the home country, and also triggers a crisis in the home country if the interdependence is sufficiently strong. Second, a fiscal and banking union may eliminate the adverse equilibrium at no costs if joint exogenous tax revenue is sufficiently high.

2.3.1 Setup

Assume that banks in both countries still hold a portfolio of government bonds. However, now this portfolio not only contains bonds of the domestic country, but also bonds of the other government. In both countries, banks hold an amount $(1-\lambda)\alpha B$ of the domestic and $\lambda\alpha B$ of the non-domestic government bonds, where $\lambda \in (0,1)$. The mass of investors who are willing to buy bonds of the respective government is denoted by η^H and η^F. When buy-

[8]For empirical evidence, see Bolton and Jeanne (2011) and Cooper and Nikolov (2013), who describe the cross-country holdings of government bonds in the euro area by using the European Banking Authority Stress Test data. Moreover, cross-country holdings of government bonds can result, e.g., from international activities of banks, or from diversification considerations (see, e.g., Bolton and Jeanne, 2011).

ing government bonds, investors do not discriminate based on the nationality of banks selling the bonds.

As before, banks in the home country need to serve each early consumer with c_1^* units in $t = 1$, potentially forcing them to liquidate a fraction $z^H(\omega^H, \eta^H, \eta^F)$ of its loan portfolio. The budget equation of home banks in $t = 1$ is therefore given by

$$\omega^H c_1^* = (1-\lambda)\eta^H + \lambda\eta^F + z^H \ell I \tag{2.6}$$

whenever possible. In analogy to the single-country case we can express liquidation as

$$z^H(\omega^H, \eta^H, \eta^F) = \min\left[1, \frac{[\omega^H c_1^* - (1-\lambda)\eta^H - \lambda\eta^F]^+}{\ell I}\right]. \tag{2.7}$$

Observe that, in contrast to the single country case, home banks' liquidation is now not only a function of aggregate withdrawal and aggregate rollover in the home country, but also a function of aggregate rollover of the foreign country's sovereign debt. Late consumption and tax revenue are given as above: they are functions of the liquidation fraction, $c_2^H(z^H, \omega^H)$ and $T^H(z^H)$. Therefore, if the foreign country defaults, which goes along a rollover freeze of foreign debt, the consumption and the tax revenue in the home country decreases because the countries are interdependent. We focus on a case where there is a high degree of interdependence between the countries.

Assumption 2.5. $\lambda \geq \frac{c_2^* - c_1^*}{c_2^*} \frac{\ell I}{\pi c_1^*}$

The assumption implies that interdependence, measured by λ, is so strong that whenever there is a sovereign default abroad, there also is a bank run at home – unless home depositors are kept from running by a deposit insurance scheme. A high λ implies that once the foreign government defaults, losses of banks at home on the foreign government bonds are also high. Assumption 2.5 implies a default abroad in fact induces a liquidation that would lower the late consumption to a level below the promised amount of early consumption, $c_2 < c_1^*$, giving patient consumers an incentive to withdraw early. For a formal analysis, see the proof of Proposition 2.3 in the Appendix.

2.3.2 International Contagion

Let us first assume that countries do not intervene abroad, but only provide a deposit insurance scheme to domestic depositors. As mentioned above, we refer to this as a *political autarky*. We analyze how a sovereign default abroad (i.e., $\eta^F = 0$, possible whenever $E^F < B$) may be contagious and affect outcomes in the home country. In doing so, we implicitly characterize the crisis equilibrium of the two-country economy.

Whenever there is a sovereign default abroad, the amount required to make a deposit insurance at home credible is given by $\widetilde{DI} = DI + \lambda\alpha B$. This amount is larger than in the single-country case. In order to make the deposit insurance scheme credible in the two-country case, the home country's government has to be able to cover the losses on foreign government bonds in addition to the cost of the deposit insurance, as specified in the single-country setup.

Proposition 2.3. *In a Nash equilibrium in which there is a sovereign default in the foreign country, the following outcomes prevail in the home country:*

$E^H < B$	***Sovereign default and bank run*** $(\omega^H, \eta^H) = (1, 0)$*: All consumers withdraw early and there is no rollover, inducing full liquidation and thus resulting in illiquidity and insolvency of both the government and the banking sector.*
$E^H \in [B, B + \widetilde{DI})$	***Bank run*** $(\omega^H, \eta^H) = (1, \alpha B)$*: Investors purchase government debt, but all consumers withdraw early. Although there is full liquidation, the government is still able fully to serve its debt.*
$E^H \geq B + \widetilde{DI}$	***No bank run, but costly deposit insurance*** $(\omega^H, \eta^H) = (\pi, \alpha B)$*: Investors purchase government debt, and only impatient consumers withdraw early. However, the deposit insurance scheme becomes costly.*

The proof of Proposition 2.3 can be found in the Appendix. Let us discuss these results in some more depth. In the first case, the home country has weak fiscal fundamentals; a sovereign debt crisis abroad will always trigger a twin crisis in the home country as well. In the second case, E^H is in an intermediate range and the home country can repay its debt for sure, but it cannot provide a credible deposit insurance. In this case, banks in the home country make a loss of $\lambda\pi c_1^* = \lambda\alpha B$, forcing them to liquidate a share of their loan portfolio, which triggers a bank run. Finally, in the third case, the fiscal fundamentals are strong and the home country can credibly promise to repay its debt and insure its deposits. Therefore, the home country can rule out a bank run at home once the foreign country defaults. However, the crisis abroad remains contagious in that banks incur a loss of $\lambda\alpha B$. Because the remaining funds of banks in $t = 2$ are smaller than $(1-\pi)c_1^*$ by Assumption 2.5, the deposit insurance scheme has to step in. The results of Proposition 2.3 are depicted in the lower area of Figure 2.3, for $E^F < B$. The three different scenarios are represented by the areas I to III.

2.3.3 Optimal Policies: Supranational Institutions

We have seen that a crisis abroad causes real losses for home banks and is thus contagious under political autarky even if the home government is able to provide a credible deposit insurance. However, it might be possible to prevent the crisis abroad through the implementation of adequate supranational institutions. We are looking for institutions that constitute a Pareto improvement compared to the situation of political autarky, in the sense that both countries weakly benefit from this policy. We focus on two different institutional setups: first, the implementation of a banking union, and second, the joint implementation of a banking union and a fiscal union. In our model, a banking union describes a supranational institution that provides a deposit insurance scheme for both countries and is financed by both countries. Similarly, in a fiscal union, both countries mutualize sovereign debt and promise to repay the debt of both countries together.

Proposition 2.4. *Assume $E^H + E^F \geq 2(B + DI)$ and $E^H > E^F$.*

$E^F < \widetilde{E}$	*A **banking union** is Pareto-efficient only if it is complemented with a **fiscal union**.*
$E^F \in [\widetilde{E}, B + DI)$	*A **banking union** is required for Pareto efficiency, but a fiscal union is not necessary.*
$E^F \geq B + DI$	*Remaining in **political autarky** is Pareto-efficient, there is no need for a banking or fiscal union.*

The threshold $\widetilde{E}$ is defined as $\widetilde{E} = B - \frac{[\ell I - (1-\lambda)\pi c_1^*]^+}{\ell I}\tau RI$.

The assumption of $E^H + E^F \geq 2(B + DI)$ implies that the pooled exogenous tax revenues of both countries suffice to repay the government debt and credibly to insure the depositors of both countries. Let us go backwards to illustrate the results of Proposition 2.4. If $E^F \geq B + DI$, the foreign government is fiscally sound and can prevent a crisis by providing a deposit insurance scheme on its own, so Pareto efficiency is already attained under political autarky. As soon as $E^F < B + DI$, the foreign country cannot provide a credible deposit insurance any more and a bank run can occur. Therefore, a joint deposit insurance is needed. Based on the level of E^F,

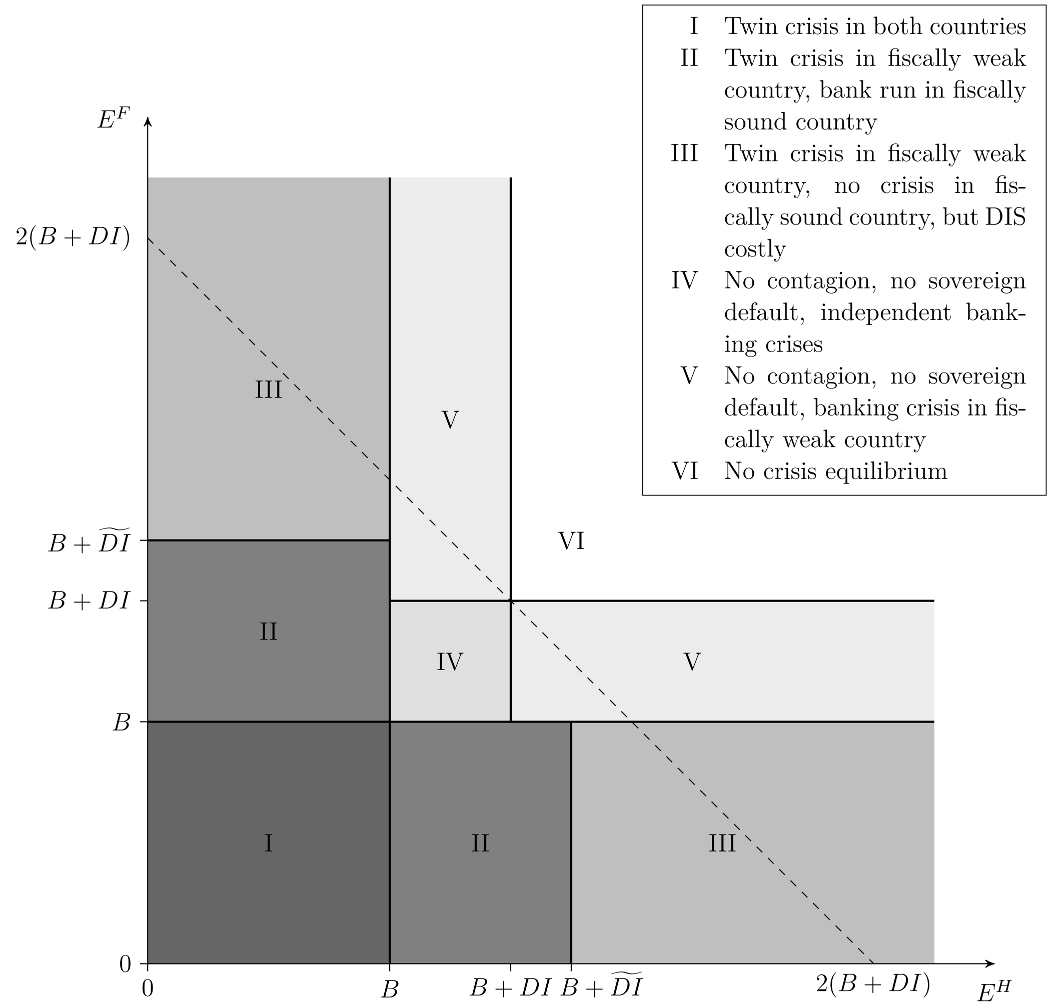

Figure 2.3: Equilibria under Political Autarky. This figure depicts the types of crisis equilibria in the case of political autarky (each government only provides a DIS for domestic depositors) for different values of external tax revenues E^H and E^F. In region I, the crisis equilibrium is a twin crisis (sovereign default and bank run) in both countries. In region II, the fiscally weak country defaults and experiences a bank run, while the fiscally sound country does not default, but experiences a banking crisis. In region III, there is a twin crisis in the fiscally weak country and no crisis in the fiscally sound country, but the DIS is costly. In region IV, one country or both countries experience a banking crisis, but sovereigns do not default and there is no contagion. The banking crises can occur independently of each other. In region V, there is a banking crisis in the fiscally weak country, but no contagion, and no crisis in the fiscally sound country. In region VI, no crisis equilibrium exists.

we have to make one further case distinction. Notice that even though the banking union prevents a bank run, banks might have to liquidate because of a rollover freeze. The rollover freeze can only occur if $E^F < \widetilde{E}$. As long as E^F is above this threshold, the remaining tax revenue after liquidation suffices to repay the government bonds B. Because the rollover freeze is ruled out, the banking union is a sufficient measure. However, if the exogenous tax revenue falls below this threshold, the rollover freeze can only be ruled out by the additional implementation of a fiscal union through which the home government guarantees the repayment of foreign government debt.

The results are depicted in Figure 2.4. Proposition 2.4 is concerned with the area above the dashed line, where $E^H + E^F \geq 2(B + DI)$. In region (i) no union is required. A banking union is strict Pareto improvement in regions (ii) and (iii), whereas in region (iv) the implementation of both a banking and a fiscal union is required.

We conclude that if the countries are sufficiently different with respect to their exogenous tax revenue, it may be beneficial for both countries to form a banking union as this eliminates the adverse crisis equilibrium at no costs.

Discussion

Notice that there is no uncertainty regarding fiscal soundness in our model, i.e., it is clear which country is fiscally weak and which country is fiscally strong. However, both countries have an incentive to form a banking union or even a fiscal union. The type of unions discussed can therefore be understood as something that is different from typical insurance against potential adverse states in the future. A typical insurance would be a contract between agents which is signed before relevant states are realized and which aims at insuring at least one of the contracting parties. Typically, there is ex post one party that makes losses on the contract because it has to transfer net funds to the other party.

In our case, however, the insurance contract can be signed after the values of external tax revenues, E^H and E^F, are realized because there is no transfer of funds from the strong to the weak country. In contrast, both countries benefit from this atypical insurance even ex post, even though the union might be valued more by the fiscally weak than by the fiscally strong country. Because it is effective in preventing self-fulfilling crises, the unions are costless for both countries. This consideration implies that if there was initial uncertainty about which of the two countries is the strong one and which is the weak one, both countries would have an incentive to form the union.

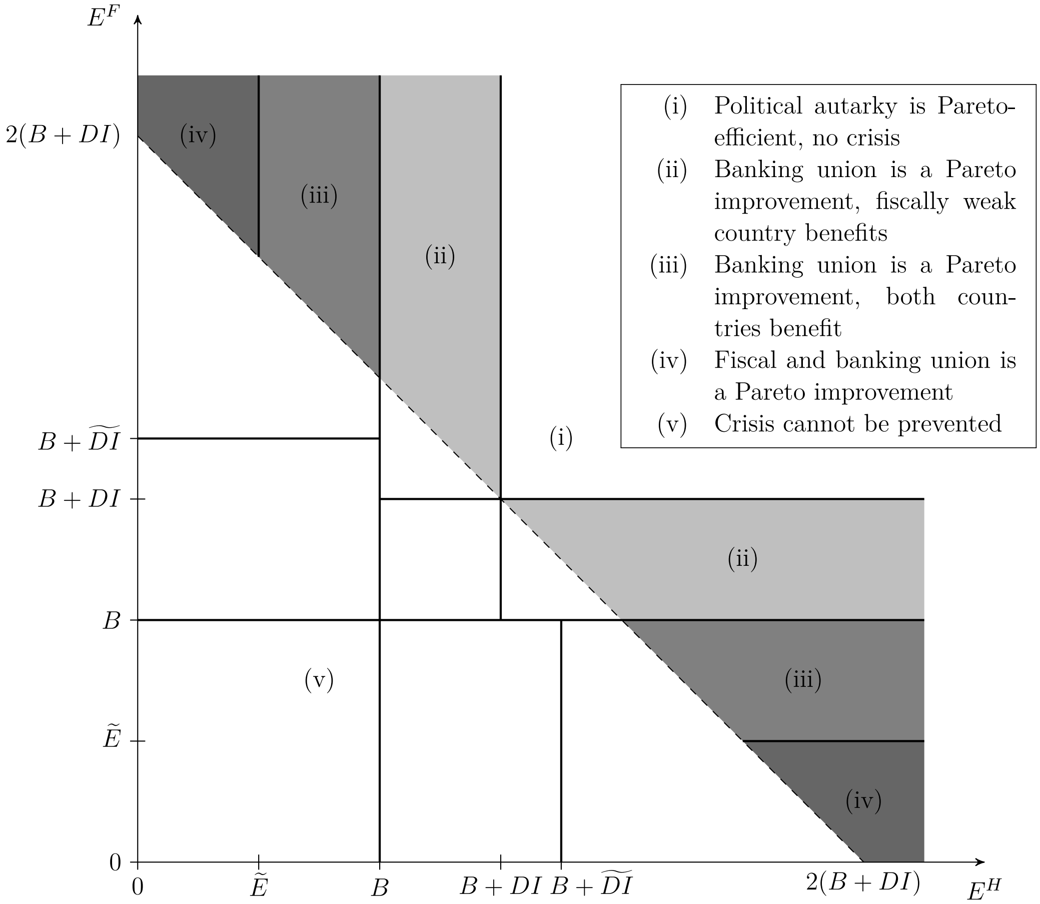

Figure 2.4: Efficient Policy Measures. This figure depicts regions in which the crisis equilibrium can be eliminated by either a banking union or the joint implementation of a Banking and fiscal union for different values of external tax revenues E^H and E^F. In region (i), a crisis equilibrium does not exist even under political autarky, thus a union is not needed. In region (ii), a banking union stabilizes the weaker country by ruling out a bank run. While it does not benefit the stronger country, it does not cost anything either. In region (iii), the banking union rules out a bank run and a sovereign default of the weaker country, thus benefiting both countries. Finally, in region (iv), the banking union is not effective anymore. Here, only the joint implementation of banking and fiscal union can eliminate the crisis, and it is costless for both countries. If the sum of exogenous tax revenues is too small, it is not possible to rule out crisis equilibria by forming a union (region (v)). While the fiscally stronger country might experience neither a sovereign default nor a bank run, it suffers whenever the weaker country experiences a sovereign default.

2.4 The European Banking Union

We now use the insights of our model to investigate the efficacy of recent policy proposals. The proclaimed goal of the proposal for the implementation of a banking union in the euro area is to ensure financial stability and to break the "potentially vicious circle between banks and sovereigns" (European Commission, 2013a).

The current proposals for the formation of a banking union consist of three components. First, the Single Supervisory Mechanism (SSM), which is supposed to be complemented by a single rulebook of the European Banking Authority (Council of the European Union, 2013). Second, the Single Resolution Mechanism (SRM) for the centralization of competencies and resources for managing the failure of banks (European Commission, 2013b). Third, a supranational Deposit Guarantee Scheme (DGS).

Currently, the first two components are already implemented (SSM) or close to being implemented (SRM), but a supranational deposit insurance scheme so far seems to be politically infeasible and is currently off the table (European Commission, 2013a). Hellwig (2014) points out several doubts about the effectiveness of SSM and SRM at dealing with cross-boarder externalities, especially for the case of banks that operate in several countries. Furthermore, he points out that national authorities may be unable or unwilling to provide funding in case of a crisis, calling for a fiscal backstop at the European level. Our model points to a further problem: We show that the lack of a supranational DGS may be a serious shortcoming of the European Banking Union, and may undermine the overall efficacy of the proposed reforms in ensuring financial stability.

Note that due to the stylized nature of our model there is no role for supervision and resolution of banks. Thus, our model remains silent on the efficacy of the components of the banking union that have already been or are about to be implemented (supervision and resolution). Clearly, both components are crucial for harmonizing banking regulation on the European level and may well be considered as a key achievement.

In turn, our model can actually say something on the supranational DGS, the component policy makers currently seem to refuse to implement. Our model states that if there is sufficient interdependence between countries and a high degree of heterogeneity in the countries' fiscal soundness, a banking union as well as a fiscal union may eliminate the self-fulfilling crisis equilibrium. Observe that in fact banks are highly interconnected within the euro area. Moreover, observe that there are countries that may be considered fiscally sound (e.g., Germany and France), and others that may be considered fiscally weak (e.g., Spain and Italy).[9] If one is willing to believe in the self-fulfilling nature of financial crises, a deposit insurance scheme, potentially complemented by a fiscal union, may implement financial stability at no costs. This also implies that the refusal to implement a deposit insurance scheme may lead to potentially costly contagion across countries, which could be avoided.

A deposit insurance scheme works best if it is credible and never tested and thus eliminates the possibility of self-fulfilling crises at no costs. In order to understand the importance of this insight in the context of the European situation, consider the following: Assume that there is a fiscally sound country that would never experience a self-fulfilling crisis if it was in autarky. However, its interdependence with another country implies its banks will realize losses once there is a crisis in the foreign country. Thus, ensuring domestic financial stability through, e.g., a deposit insurance scheme will become costly for the government once its banks have realized losses. A crisis abroad may therefore cause real costs at home once there is sufficient interdependence. Given the self-fulfilling nature of the crisis abroad, it may be optimal for the home country to participate in a mechanism that prevents the crisis abroad at low (or even at zero) costs. Preventing the crisis abroad eliminates contagion and thus ensures financial stability at home in this setting. Our model shows that this is possible by implementing a banking union (equivalent to a joint deposit insurance scheme in our model) which is complemented with a fiscal union if necessary.

In order to apply this insight to the European situation, one needs to appreciate the fact that a deposit insurance in fiscally weak countries may not be credible. It may therefore not be able to prevent a banking crisis in the respective country, a crisis that can be contagious and thus costly for fiscally sound countries as well. A banking union with a joint deposit insurance scheme may increase the credibility of the deposit insurance. In fact, the deposit insurance scheme may become fully credible once it is backed by fiscally sound governments, eliminating the crisis equilibrium altogether. In fact, in our very simple setup, such a mechanism can eliminate the crisis equilibrium at no cost.

One may hypothesize that politicians in fiscally sound countries currently

[9] We do not consider our model to fit the case of Greece. It is more than questionable whether Greece could have repaid its debt even if its debt had been a rolled over. The crisis in Greece does not appear to be only self-fulfilling, but rather due to fundamental problems. Greece and Portugal rather had “old-fashioned sovereign debt crises” (Hellwig, 2014).

seem to be scared of implementing a joint deposit insurance scheme. The rationale is that it could appeal to voters as another form of mutualization of national debt, with a clear disadvantage for taxpayers in fiscally sound countries. E.g., German politicians may fear to scare their voters as a banking union may imply that German tax payer can potentially be liable for losses of, e.g., Spanish banks. Our model indicates that this may turn out to be bitter irony: exactly the refusal of implementing a full-fledged banking union with a joint deposit insurance scheme may make future crises more costly for the respective tax payers.

2.5 Conclusion

Our model has two main contributions. First, we discuss how banking crises and sovereign defaults can be contagious across countries. The setup allows us to rationalize supranational policies that aim at preventing sovereign and financial crises. Our specific setup gives conditions under which a fiscal and a banking union are effective measures to eliminate an adverse run equilibrium. We use these results to comment on the policy debates on the making of a banking union in the euro area. Importantly, our model indicates that a banking union with a joint deposit insurance scheme may be a mechanism to prevent contagious self-fulfilling banking crises. It possibly has to be complemented by a fiscal union to be entirely effective. We argue that the current proposal for a banking union, consisting only of supranational supervision and resolution mechanisms, is insufficient to break the vicious cycle between sovereigns and banks.

Second, the model illustrates the risks associated with banks holding government bonds. In our model, fragility arises whenever the fate of the government and the financial sector are closely connected. This condition is likely met in developed and highly leveraged financial systems where banks hold government bonds and where economic activity depends on the performance of the financial sector. This chapter thus sheds light on the debate regarding the regulation of government bonds holding by intermediaries. More specifically, it gives a rationale for why exposure of banks to sovereign risk may be problematic.

The stylized nature of our model implies that our insights and policy implications have to be taken with a grain of salt and cannot be translated one-to-one for every institutional arrangement. In our model, we abstract from fundamental uncertainty (i.e., macroeconomic shocks) as a source of a crisis, and from potential moral hazard resulting from an established banking and fiscal union. Both elements may be of importance in reality. For the case of negative macroeconomic shocks in a foreign country, a supranational deposit insurance may moderate a crisis, but this might come with real costs for the home country. In addition, the presence of an international insurance may induce a country's institutions (government, supervision, and banks) to gamble. Both aspects might induce fiscally strong countries to refrain from a fiscal and an extensive banking union. This is not an argument against such unions, though. It rather calls for detailed contractual definitions of the union's scope, and for strict regulation and supervision that is located at level of the union. The SSM and SRM can mitigate such moral hazard on the country level, and thus build the foundation which is necessary for implementing a supranational DGS.

Appendix 2.A Proofs

Proof of Proposition 2.1. We first analyze proof the existence of the Type I equilibrium where $(\omega, \eta) = (\pi, \alpha B)$: By Equation (2.1), banks do not engage in liquidation, $z(\omega, \eta) = 0$, yielding a late consumption of $c_2 = c_2^*$ and a tax revenue of $T = E + \tau RI$. Assumptions 2.1 to 2.3, and Equations (2.3) and (2.5) imply that patient consumers do not withdraw early, $\omega_i^*(\omega, \eta) = 0 \ \forall i$, and outside investors roll over the debt $\eta_j^*(\omega, \eta) = 1 \ \forall j$. Therefore $(\omega, \eta) = (\pi, \pi c_1^*)$ constitutes a Nash Equilibrium.

We now proof the existence of the Type II equilibrium. We distinguish two cases.

$E < B$ **Sovereign default and bank run** $(\omega, \eta) = (1, 0)$**:**
The liquidation is given by $z(\omega, \eta) = 1$, yielding $c_2 = 0$ and $T = E$. We get $\omega_i^*(\omega, \eta) = 1 \ \forall i$ and $\eta_j^*(\omega, \eta) = 0 \ \forall j$. Therefore $(\omega, \eta) = (1, 0)$ constitutes a Nash Equilibrium.

$E \geq B$ **Bank run** $(\omega, \eta) = (1, \alpha B)$**:**
The liquidation is given by $z(\omega, \eta) = 1$, yielding $c_2 = 0$ and $T = E$. We get $\omega_i^*(\omega, \eta) = 1 \ \forall i$ and $\eta_j^*(\omega, \eta) = 1 \ \forall j$. Therefore $(\omega, \eta) = (1, \alpha B)$ constitutes a Nash Equilibrium. □

Proof of Proposition 2.3. A sovereign default in the foreign country implies that $\eta^F = 0$, implying that domestic banks make a loss of $(1 - h)\alpha B$. Assumption 2.5 implies that this loss induces a liquidation which necessarily triggers a bank run in the home country in the absence of the a deposit insurance. To prove this fact, we show that even if there was rollover of sovereign debt and no run, depositors would still prefer to run, i.e., $c_2^H < c_1^*$. In this case, the liquidation would be $z^H(\pi, \alpha B, 0) = (\pi c_1^* - (1-\lambda)\pi c_1^*)/\ell I = \lambda \pi c_1^* / \ell I$. By Assumption 2.5, it follows that $z^H > \frac{c_2^* - c_1^*}{c_2^*} = 1 - \frac{c_1^*}{c_2^*}$. Late consumption is given by $c_2^H(z^H, \pi) = (1 - z^H)c_2^*$. It follows that $c_2^H(z^H, \pi) < (c_1^*/c_2^*)c_2^* = c_1^*$. Therefore, the bank run is inevitable in the absence of a deposit insurance.

$E^H < B$**: Sovereign default and bank run,** $(\omega^H, \eta^H) = (1, 0)$
Because the government cannot provide a deposit insurance, a bank run is triggered. This leads to full liquidation and reduces the tax revenue to $T^H = E^H < B$, inducing a sovereign default and a rollover freeze.

$E^H \in [B, B + \widetilde{DI})$**: Bank run,** $(\omega^H, \eta^H) = (1, \alpha B)$
Because the government cannot provide a deposit insurance, a bank run is triggered. This leads to full liquidation and reduces the tax revenue to $T^H = E^H > B$. The sovereign can repay its debt, and rollover is ensured.

$E^H \geq B + \widetilde{DI}$**: No Bank run but costly deposit insurance,** $(\omega^H, \eta^H) = (\pi, \alpha B)$
The government can provide a deposit insurance scheme and thus prevent a bank run, and it can also repay its debt, ensuring the rollover of debt. However, the deposit insurance is costly. □

Proof of Proposition 2.4. Given that $E^F \geq B + DI$, it is immediately clear that both countries are stable under political autarky, therefore a union is not needed.

In the presence of a banking union, the foreign government cannot experience a sovereign debt crisis if $E^F \geq \widetilde{E}$. Because joint funds suffice to make a banking union credible, it prevents a run in the foreign country. If there was a rollover freeze in the foreign country, banks would have to liquidate $z^F(\pi, 0, \alpha B) = \min[1, (1-\lambda)\pi c_1^*/(\ell I)]$. This induces a tax revenue of $T^F(z^F) \geq \widetilde{E} + \frac{[\ell I - (1-\lambda)\pi c_1^*]^+}{\ell I} \tau R I = B$. The foreign government can thus repay its debt, and a rollover freeze cannot occur in equilibrium. Therefore, the banking union is sufficient to eliminate any crisis altogether if $E^F \geq \widetilde{E}$.

If however $E^F < \widetilde{E}$, a rollover freeze constitutes an equilibrium even in the presence of a banking union which prevents a bank run. In case of a rollover freeze, the tax revenue is given by $T^F(z^F) < \widetilde{E} + \frac{[\ell I - (1-\lambda)\pi c_1^*]^+}{\ell I} \tau R I = B$. Therefore, the joint implementation of the banking and the fiscal union is required. This policy measure is costless for the home country because by providing the deposit insurance and guaranteeing to repay all government debt, it rules out a bank run and ensures rollover of foreign government debt. The deposit insurance will not be tested, and because foreign banks do not engage in liquidation, the foreign government has sufficient tax revenue to repay its debt by itself.

Returning to the case of $E^F \in [\widetilde{E}, B - DI]$, we can distinguish two different scenarios. If $E^F \in [\widetilde{E}, B]$, both countries strictly benefit from the implementation of the banking union. The foreign country does not experience any crisis, and because the default of the foreign sovereign is ruled out, losses of home banks on foreign government bonds are eliminated. In contrast, if $E^F \in [B, B + DI)$, the foreign country will always be able to repay its debt. Therefore, the home country cannot be affected by a crisis at all. Even if there was a bank run in the foreign country, the home country would not suffer because the exposure is only through foreign debt which is unaffected. Thus, only the foreign country benefits from the banking union, but the home country does not suffer. This distinction is illustrated by the regions (ii) and (iii) in Figure 2.4. □

3

Regulatory Arbitrage and Systemic Liquidity Crises

3.1 Introduction

Regulatory arbitrage has been identified as one of the main ingredients to the 2007-09 financial crisis (FCIC, 2011; Acharya et al., 2013). Hence, post-crisis reforms have targeted the main channels through which the turmoil in the shadow banking sector has affected the commercial banking sector, particularly, explicit or implicit contractual linkages between the two sectors. A natural question that arises in this context is whether the implemented and proposed reforms are effective in shielding the commercial banking sector in future crises?

This chapter discusses a new theoretical channel for how regulatory arbitrage in banking may contribute to overall financial fragility. We show that regulatory arbitrage may set the stage for panic-based runs in the shadow banking sector, which may affect the regulated banking sector via a deterioration of wholesale funding conditions. The non-regulated banking activities may thus affect the commercial banking sector via channels beyond those that have been targeted in the post-crisis reforms. We argue that restrictions on wholesale funding may be justified if the objective is to ensure stability of depository institutions in the presence of regulatory arbitrage.

Previous to the 2007-09 financial crisis, many commercial banks had set up off-balance sheet conduits to finance long-term real investment by issuing short-term debt. In the summer of 2007, increased delinquency rates on subprime mortgages ultimately led to the collapse of the conduits' main source of funding: the market for asset-backed commercial papers (ABCP) (see, e.g., Kacperczyk and Schnabl, 2009; Covitz et al., 2013). Many commercial banks had explicitly or implicitly sponsored these conduits,[1] and the collapse forced

[1] Asset-backed commercial paper conduits were set up to finance mortgage-backed securities (MBS) and asset-backed securities (ABS) by issuing ABCP or medium-term notes

those banks to take the conduits' asset and liabilities on their balance-sheets, thus creating severe solvency issues.

From an ex-post perspective, it appears that off-balance sheet banking had to a large extent been conducted to circumvent existing capital regulation (see, e.g., Acharya et al., 2013).[2] In this context, the adverse implications of explicit as well as implicit contractual linkages between regulated and unregulated banks have been identified as a particularly important source of instability (Segura, 2014). Consequentially, the overwhelming regulatory response has been to close the obvious loopholes in regulation by outright prohibition of contractual links between depository institutions and other parts of the financial system (compare, e.g., Section 619 of the Dodd-Frank Act, referred to as the "Volcker Rule", Report of the Vickers Commission, Liikanen Report).

In this chapter, we show that if intermediation optimally relies on wholesale funding, regulatory arbitrage poses a real threat to overall financial stability even in the absence of any contractual linkages between commercial banks and shadow banking activities. The main mechanism is that panic-based runs in the shadow banking sector may induce fire sales. A binding cash-in-the-market constraint may then also lead to a deterioration of the wholesale funding conditions for regulated banks. Contagion via the deterioration of funding conditions creates real costs for commercial banks and potentially leads to their illiquidity and insolvency. Moreover, these costs grow with the size of the shadow banking sector, as a larger shadow banking sector induces stronger fire-sale effects.

We argue that understanding such mechanisms is particularly important in the light of the new regulations implemented in the aftermath of the 07-09 financial crisis. Many of the new regulatory measures have been implemented under the premise that a prohibition of explicit or implicit contractual linkages between commercial banking and other types of banking can shield the former from turmoil originating in the latter. In particular, regulation has focused on prohibiting sponsor support, as well as on the separation of traditional banking and other activities, such as proprietary trading and market making.[3] At the same time, the question has been raised whether deposi-

(MTN), and were granted explicit credit or liquidity guarantees, or implicit guarantees as in the case of structured investment vehicles (SIV).

[2]To some observers, this had already been clear prior to the crisis; see Jones (2000).

[3]E.g., the Financial Services Act of 2013, which was based in the Vickers Commissions Report, suggests a limit on the exposure of depository institutions to other financial institution within the same bank holding company. Likewise, the Liikanen Report distinguishes between the "deposit bank" and the "trading entity" within the same bank holding company. The "Volcker Rule" limits proprietary trading activities in

tory institutions should be allowed to fund some of their activities by using wholesale funding. The Basel III liquidity regulation proposes to restrict the holding of illiquid assets and the reliance on unstable funding sources by introducing the liquidity coverage ratio (LCR) and the net stable funding ratio (NSFR).

This chapter sheds new light on how these measures affect financial stability in the presence of regulatory arbitrage. We argue that prohibiting contractual linkages is not sufficient to shield the regulated banking sector from financial fragility. It may be desirable to implement further restrictions on commercial banks. In particular, we show that, in the presence of regulatory arbitrage, restricting wholesale funding in depository institutions allows shielding the regulated banking sector from the consequences of turmoil in the unregulated shadow banking sector. Such macroprudential reasoning calls for liquidity regulation of depository institutions. However, we indicate that such regulation may achieve financial stability at the expense of allocative inefficiencies and lead to further growth of fragile shadow banking. Moreover, we argue that, in the context of the Basel III liquidity regulation, an explicitly macroprudential approach to regulation would be crucial for its efficacy. In particular, we use our model to argue that one should be more restrictive when declaring which assets are considered as *liquid* and which sources of funding are considered as *stable* in case of a systemic liquidity crisis.

Our model builds on – and at the same time nests – the banking model of maturity and liquidity transformation by Diamond and Dybvig (1983). We enrich the setup along two dimensions: On the one hand, we introduce a new type of agents, called "investors". These agents are only present from the interim period onwards and can provide funds to banks in exchange for claims on future cash-flows, as in Chapter 1. This makes it optimal for intermediation to use what we refer to as "wholesale funding" instead of using storage. On the other hand, we introduce a shirking technology that allows a disciplining role of short-term debt, reminiscent of Calomiris and Kahn (1991) and Diamond and Rajan (2001). First, we show that the disciplining effect of short-term debt allows intermediaries to implement the first-best allocation in our setup. Intermediaries refrain from shirking as they cannot enjoy private benefits if their depositors collectively withdraw. We also show that while short-term debt may be disciplining, it is also necessarily associated with the possibility of panic-based runs, as discussed by Admati and Hellwig (2013). A regulator that decides to provide a deposit insurance to eliminate panic-based runs would thus undermine the disciplining effect of short-term debt.

depository institutions.

This makes it necessary to complement a safety net with equity regulation, which induces diligent behavior via a textbook skin-in-the-game mechanism à la Tirole (2010).

Assuming that such equity regulation is costly, incentives to circumvent regulation may arise, and intermediaries may place themselves outside the regulatory perimeter in the so-called shadow banking sector. Institutions in this sector are not covered by the deposit insurance and have no access to the central bank's discount lending, implying that they can be subject to panic-based runs. We emphasize a new theoretical channel through which these runs may be contagious, affecting the regulated banking sector: A systemic run on shadow banks induces fire sales with cash-in-the-market pricing à la Allen and Gale (1994). A binding cash-in-the-market constraint implies that wholesale funding conditions also deteriorate in a fire sale. As regulated commercial banks optimally rely on wholesale funding, a fire sale creates real costs for them. Depending on the size of the shadow banking sector, regulated banks may ultimately become insolvent, and the provision of a deposit insurance may become costly. This contagion channel can be shut down by prohibiting wholesale funding or regulated banks. These restrictions do, however, lead to allocative inefficiencies in the regulated sector and increase the size of the shadow banking sector.

Throughout our main analysis, we treat runs on shadow banks as zero-probability events. In order to show that our results are robust to anticipated runs, we analyze the effects of sunspot runs that occur with positive probabilities, as in Cooper and Ross (1998). Naturally, the shadow banking sector becomes less attractive when runs are anticipated. However, this does not lead to a complete breakdown of shadow banking. Moreover, runs in the shadow banking sector also affect the regulated banking sectors' expected refinancing conditions and thus also make the intermediation of regulated banks less attractive. This mitigates the effect which anticipated runs in the shadow banking sector have on the relative attractiveness of the regulated banking sector. In this context, we also briefly analyze how government interventions in the form of a Lender of Last Resort (LoLR) and a Market Maker of Last Resort (MMLR) may affect the stability of the financial system. We find that while these interventions can shield the regulated banks, they necessarily also benefit the shadow banks.

Finally, we also discuss the role of direct contractual linkages between the two sectors in the form of liquidity guarantees. We show that liquidity guarantees are optimal from the perspective of a single institution, but they exacerbate the adverse consequences of runs in the shadow banking sector. This shows that the prohibition of contractual linkages may indeed be desirable. We argue, however, that prohibiting links alone is not enough to shield regulated banks from financial instability.

Literature

This chapter connects to the recent literature on theoretical aspects of shadow banking. The paper by Bolton et al. (2011) presents a model in which intermediaries originate assets, and can sell them on a secondary market to raise "outside liquidity". Similar to their paper, we discuss how the presence of such outside liquidity may affect the business model of banks. However, we abstract from adverse selection and emphasize the role of coordination.

Our modeling approach is also related to the papers by Martin et al. (2014a), Martin et al. (2014b), and Luck and Schempp (2014a). All three rely on the setup of Qi (1994). Martin et al. (2014a) investigate the differences between bilateral and tri-party repo in determining the stability of single financial institutions. Martin et al. (2014b) focus on the difference between runs on single institutions and systemic runs in secured funding markets. While our model has a similar notion of cash-in-the-market pricing, our focus lies on the effects of regulation and in particular on the adverse effects of shadow banking on regulated banks. In this respect, this chapter is close to Luck and Schempp (2014a), who analyze the effects of shadow banking on financial stability in the presence of contractual linkages between the two sectors. The crucial difference to Luck and Schempp (2014a) is that this chatper focuses on contagion in the absence of contractual linkages.

Segura (2014) explicitly discusses the role of liquidity guarantees. In particular, he gives a theoretical answer to the question posed by Acharya et al. (2013) on why banks supported their SIVs despite the absence of direct contractual linkages. While Segura elegantly shows how regulated banks may have incentives to support their shadow bank operation in order to preserve their reputation, we argue that there may be reasons to believe that regulatory arbitrage may affect stability of regulated banks beyond such contractual linkages.

Other important contributions that deal with shadow banking are Ordoñez (2013), Gennaioli et al. (2013) and Plantin (2014). Gennaioli et al. (2013) provide a model in which the demand for safe debt drives securitization, and fragility in the shadow banking sector arises when tail-risk is neglected. Ordoñez focuses on potential moral hazard on the part of banks. In his model, shadow banking is potentially welfare-enhancing as it allows one to circumvent imperfect regulation. However, it is only stable if shadow banks value their reputation and thus behave diligently; it becomes fragile otherwise. Plantin studies the optimal prudential capital regulation when regulatory arbitrage is possible. In his paper, relaxed capital requirements lead to a decline of the shadow banking sector, potentially improving welfare. In contrast to all three, we focus on the destabilizing effects of shadow banking in the sense that it gives rise to run equilibria.

3.2 Setup

Consider an economy that goes through a sequence of three dates, $t \in \{0, 1, 2\}$. There is a single good that can be used for consumption as well as for investment. The economy is populated by three types of agents: consumers, intermediaries, and investors.

Technologies
Altogether, there are three technologies available for investment (see a summary of the payoff structure in Figure 3.1). There is a short technology ("storage") available in $t = 0, 1$, transforming one unit invested in t into one unit in $t+1$. Moreover, there are two illiquid technologies available for investment in $t = 0$: a "productive technology" and an unproductive "shirking technology". Both technologies are technologically illiquid, i.e., for one unit invested they produce a return in $t = 1$ only if they are physically liquidated, and the physical liquidation rate of the technologies is assumed to be $\ell \to 0$. Note that the technologies (or claims on the technologies' future returns) may nonetheless be sold at higher values at a secondary market that will be specified below.

The return properties of the illiquid technologies in $t = 2$ are as follows: One unit invested in the productive technology yields a safe return of R units in $t = 2$. One unit invested in the shirking technology yields a safe return of $R_{shirk} < 1$ in $t = 2$. However, this technology yields a private benefit $B > 0$ in $t = 2$ which is available only to the agent who owns investment at this point in time, i.e., it is non-transferable and non-contractible. Moreover, it only accrues if the technology is not physically liquidated in the interim period.

	$t = 0$	$t = 1$	$t = 2$
Storage in $t = 0$	-1	1	0
Storage in $t = 1$	0	-1	1
Productive technology	-1	$\ell \to 0$	R
Shirking technology	-1	$\ell \to 0$	$R_{shirk} + B$

Figure 3.1: Payoff structure of technologies

We assume that $R_{shirk} + B \leq 1$. This implies that the shirking technology is inefficient, although it generates a private benefit. As will become clear later, the possibility of investing in this technology and financing the investment by short-term debt will give rise to moral hazard. This moral hazard will lead to the necessity of capital regulation once a deposit insurance undermines the disciplining effect of short-term debt.

Consumers

There is a continuum of consumers with mass one. Initially, consumers face idiosyncratic uncertainty with regard to their preferred date of consumption, and they may lend their endowment to intermediaries to invest on their behalf.[4]

Each consumer is endowed with 1 unit of the good in $t = 0$. There are two types of consumers, *patient* and *impatient* consumers: a fraction π is impatient and derives utility only from consumption in $t = 1$, $u(c_1)$, and a fraction $1 - \pi$ is patient and derives utility only from consumption in $t = 2$, $u(c_2)$. We restrict attention to CRRA utility, i.e., the period-utility function has the form $u(c_t) = \frac{1}{1-\eta} c_t^{1-\eta}$, with $\eta > 1$.

Initially, consumers do not know their type; their probability of being impatient is identical and independent, so all consumers have the same prior π initially. In period one, each consumer privately learns his type, this can be considered as a liquidity shock.

A consumption profile (c_1, c_2) denotes an allocation where an impatient consumer receives c_1 and a patient consumer receives c_2. As of period 0, such a consumption profile induces an expected utility of

$$U(c_1, c_2) = \pi u(c_1) + (1 - \pi) u(c_2) = \frac{1}{1 - \eta} \left[\pi c_1^{1-\eta} + (1 - \pi) c_2^{1-\eta} \right]. \qquad (3.1)$$

Notice that the attributes *patient* and *impatient* characterize the consumer's exogenous type, which determines his preference. In contrast, the attributes *late* and *early* will characterize the timing of actual consumption, and in the case of demand-deposit contracts, it denotes the withdrawals, which are endogenous: An "early consumer" withdraws in $t = 1$, while a "late consumer" withdraws in $t = 2$.

[4] We assume that consumers cannot invest in technologies directly in the initial stage and trade technologies in the interim period. They can only lend their funds to intermediaries. In a later section, we will argue briefly why we can focus on a banking solution directly, i.e., why a banking solution dominates a financial markets solution.

Intermediaries

There is a mass m of intermediaries.[5] While consumers cannot invest in the technologies directly, intermediaries face no investment restrictions. Intermediaries have no market power, they compete for the consumers' funds which they collect in exchange for a demand-deposit contract, and they invest the funds in the technologies. Moreover, they may choose to invest some of their own funds in the intermediation business.

Intermediaries only care about $t = 2$ consumption.[6] Each intermediary is endowed with E units of the good. We assume that E is large, implying that no result will be driven by the aggregate intermediaries' endowment mE becoming a binding resource constraint. Importantly, intermediaries are assumed to have an outside option, resulting in a required return of $\rho > R$ in $t = 2$ for each unit invested in $t = 0$. Because the required return is larger than the technologies' returns, it is costly for the consumers if the intermediaries invest their own endowment for investment. As we will see later, this assumption makes it costly to use a skin-in-the-game mechanism in order to provide intermediaries with incentives to invest in the productive technology instead of in the shirking technology in the presence of a deposit insurance.

On the liability side, intermediaries initially offer the deposit contract (c_1, c_2) to consumers in exchange for one unit of initial deposits. Moreover, intermediaries choose to invest e_0 units of their endowment in the intermediation business in $t = 0$, in exchange for receiving e_2 units in $t = 2$. While we do not initially impose restrictions on how intermediaries finance intermediation, equity financing will turn out to be optimal.

On the asset side, intermediaries make the following investment decision: We denote by I the investment in the productive technology, by I_{shirk} the investment in the shirking technology, and $1 + e_0 - I - I_{shirk}$ denotes the investment in storage. We assume that an intermediary's investment decision is unobservable in $t = 0$, but becomes public information in $t = 1$.

Investors

There is a continuum of investors of mass n. Investors only become active in the interim period and can provide liquidity to intermediaries: Investors can

[5]It is assumed that m is small compared to the mass of depositors such that each bank has a very large number of depositors, and thus does not face aggregate liquidity risk by a law of large numbers argument.

[6]As the model has no aggregate uncertainty, the shape of intermediaries' utility is not important. They may be risk-neutral or risk-averse. Only for the case of sunspot runs with positive probability we will assume that intermediaries are risk-neutral in order to keep the analysis tractable.

transfer some of their endowment to intermediaries in exchange for a claim on some of the future returns of the intermediaries' technologies. We refer to this activity as "wholesale funding". Later, we will show that the presence of these investors makes investment in storage inefficient, i.e., it is optimal for intermediaries to rely on wholesale funding from outsiders instead of storing real goods. However, this will also give rise to the main contagion channel between regulated and unregulated banking: When a run on shadow banks induces a fire sale, a cash-in-the-market constraint can become binding, and wholesale funding conditions for regulated banks deteriorate as well.

Investors are born in $t = 1$ and receive an endowment of A/n, so the investors' aggregate endowment is given by A. The endowment A will be one of the crucial parameters of the model: while it may be a sufficient source of liquidity in normal times, it may lead to a binding cash-in-the-market constraint in case of systemic runs. Given that investors are born in $t = 1$, it is not possible to contract with them in $t = 0$. Investors care about consumption[7] in period 2, and they are assumed to have an outside option, which induces a required return of γ, where $\gamma \in [1, R]$. That is, for each unit they transfer to intermediaries in $t = 1$, they need to receive at least γ units in $t = 2$.

Investors have no market power and thus are price takers as long as their liquidity A is not scarce, i.e., they take the conditions of wholesale funding as given. The required return γ implies that they are willing to provide liquidity as long as the return r they receive satisfies $r \geq \gamma$.

There are two different contractual specifications of wholesale funding (i.e., of how investors provide liquidity to intermediaries) which are economically equivalent: asset sales and collateralized lending. If investors purchase assets with a face value of R in period 2 at a price p in period 1, they get a return of $r = R/p$. On the other hand, they can also lend one unit to banks at the interest rate $r = R/p$ while receiving $r/R = 1/p$ units of asset as collateral.

As long as liquidity is not scarce, competition among investors will induce $r = \gamma$. However, if liquidity becomes scarce, we assume that the asset price and the interest charged in collateralized lending is determined by a cash-in-the-market constraint.

[7] Again, the shape of their utility function is not important as long as it is compatible with the specified outside option.

3.3 Optimal Intermediation and Runs

3.3.1 First-Best

We will now derive the allocation that maximizes the expected utility of consumers, subject to the participation constraints of intermediaries and investors, and subject to the resource constraints. Since our objective is to maximize consumers' welfare, we treat these participation constraints as resource constraints. We refer to the resulting allocation as the first-best allocation and denote it by (c_1^*, c_2^*).

In the first-best, the shirking technology is not used because the productive technology strictly dominates the shirking technology, i.e., $I_{shirk} = 0$. We denote by L the units of the productive technology that get transferred from intermediaries to investors, in exchange for Lp units of the good ("liquidity") from investors to intermediaries in the interim period. This transaction can be interpreted as asset sales, but is mainly referred to as "wholesale funding" in the following.

The first-best maximization program is given by

$$\max_{(c_1,c_2,e_0,e_2,I,L,p)\in\mathbb{R}_+^7} \quad \pi u(c_1) + (1-\pi)u(c_2), \tag{3.2}$$

$$\text{subject to} \quad \pi c_1 \leq (1 + e_0 - I) + Lp, \tag{3.3}$$

$$(1-\pi)c_2 \leq (I-L)R - e_2, \tag{3.4}$$

$$e_2 \geq \rho e_0 \geq 0, \tag{3.5}$$

$$R \geq \gamma p, \tag{3.6}$$

$$pL \leq A, \tag{3.7}$$

$$I \leq 1 + e_0, \tag{3.8}$$

$$L \leq I. \tag{3.9}$$

The budget constraints for periods one and two are given by (3.3) and (3.4). Investors may transfer Lp to consumers in $t = 1$ in exchange for L units in $t = 2$. As indicated above, we refer to this as wholesale funding. (3.5) represents the participation constraint of the intermediary and non-negativity constraint, and (3.6) represents the participation constraint of investors. The resource constraint on investors' capital A ("interim liquidity") in the interim period is given by (3.7). Finally, (3.8) and (3.9) denote the constraint on initial investment as well as the constraint on the units of assets that can be sold in the interim period.

The constraint (3.9) cannot be binding because the Inada conditions require that $c_2 > 0$ and thus $L < I$. The two budget constraints are always binding. Furthermore, the participation constraint of intermediaries must also be binding. Moreover, $e_0 > 0$ cannot be optimal. Because $\rho > R$, we can reduce e_0 and thus relax the second-period constraint. It therefore follows that $e_0^* = e_2^* = 0$. Because the intermediaries' required return is higher than the asset return, intermediaries' funds are not used for intermediation in the first-best. As we will see below, moral hazard may make it necessary to force the intermediary to invest some of his endowment.

Let us now turn towards the use of interim liquidity, i.e., wholesale funding in period 1. In the first-best, it also has to hold that the participation constraint of investors is binding. Whenever $p < R/\gamma$ and $pL < A$, we can increase p and thereby relax the period 1 constraint. Whenever $p < R/\gamma$ and $pL = A$, we can increase p and decrease L as much as necessary, thereby relaxing the period 2 constraint. Therefore, it holds that $p = R/\gamma$.

We are now left with a maximization problem with two weak inequalities.

$$\max_{(c_1,c_2,I,L)\in\mathbb{R}_+^4} \quad \pi u(c_1) + (1-\pi)u(c_2), \tag{3.10}$$

$$\text{subject to} \quad \pi c_1 = (1-I) + LR/\gamma, \tag{3.11}$$

$$(1-\pi)c_2 = (I-L)R, \tag{3.12}$$

$$LR/\gamma \leq A, \tag{3.13}$$

$$I \leq 1. \tag{3.14}$$

Depending on the model parameters A, R, γ, and π, as well as on the shape of the utility function, the first-best program now has three solution candidates. As discussed in detail in the appendix, investment in storage is only optimal if A is small, and becomes unnecessary when A is sufficiently large. For the remaining part of this chapter, we will assume that we are in the case in which the endowment of the investors A is large enough such that the investors' budget constraint (3.13) is not binding. In this case, storage is not used, and there is only investment in the productive technology, i.e., $I^* = 1$. This translates into the following assumption:

Assumption 3.1. $A \geq \xi \equiv \pi\gamma^{-\frac{1}{\eta}} \frac{R}{(1-\pi)+\pi\gamma^{1-\frac{1}{\eta}}}$.

For a detailed discussion of the implications of Assumption 3.1, see Appendix 3.A, in which we also characterize the the first-best for the case that investors' capital is scarce. Assumption 3.1 allows us to focus on a setup where intermediation optimally relies exclusively on investors providing interim liquidity through wholesale funding and refrains from the use of storage.

Lemma 3.1 (First-best allocation)**.** *The first-best allocation is characterized by*

$$I^* = 1, \quad L^* = \xi\gamma/R, \ \textit{and}\ e_0 = e_2 = 0,$$

and the optimal consumption profile is given by

$$c_1^* = \gamma^{-\frac{1}{\eta}} \frac{R}{(1-\pi) + \pi\gamma^{1-\frac{1}{\eta}}} \quad \textit{and} \quad c_2^* = \frac{R}{(1-\pi) + \pi\gamma^{1-\frac{1}{\eta}}}. \tag{3.15}$$

The risk-sharing between early and late consumers is described by the FOC $u'(c_1) = \gamma u'(c_2)$ because under wholesale funding, the technological rate of substitution between period 1 and 2 is given by the investors' required return γ. Diamond and Dybvig (1983) restrict attention to utility functions with a relative risk aversion larger one. In their setup, risk-sharing between patient to impatient consumers is optimal, implying that $c_1^{DD} > 1$, where 1 is the technological rate of return between periods 0 and 1 (storage). However, this condition also enables self-fulfilling runs. In contrast, we focus on the special case of constant relative risk aversion. The parameter of relative risk aversion is constant and given by $\eta > 1$. We get a similar result with respect to risk-sharing: It holds that $c_1^* > R/\gamma$, where R/γ is the rate of return between periods 0 and 1 under wholesale funding. But as we shall see in next subsection, this condition also has similar and important implications for fragility and self-fulfilling runs.

Lemma 3.5 in the Appendix describes the first-best if Assumption 3.1 does not hold. For $A < \xi$, the investors' endowment constraint (3.13) becomes binding. Furthermore, there exists some threshold $\xi_0 < \xi$ below which partial investment in storage becomes optimal. In the extreme case of $\gamma = R$ or $A = 0$, the optimal consumption profile is identical to that in the Diamond and Dybvig model with CRRA utility, which is thus nested in our model.

3.3.2 Intermediary Implementation

In the following, we will show that the first-best allocation can be implemented by demand-deposit contracts offered by the intermediaries. We first show that consumers are willing to lend to intermediaries, although intermediaries have the option of investing in the shirking technology. We will show that the demand-deposit contracts allow depositors to discipline the intermediary. In a second step, we show that the disciplining element of the demand-deposit contract is associated with financial fragility in the sense that panic-based runs may take place in the interim period.

Disciplining Demand-Deposit Contracts

We assume that consumers cannot invest in the technologies directly, but only via intermediaries. Let us first consider the agency problems on the part of the intermediary. To this end, let us first devote more attention to the timing and the action space of consumers and intermediaries.[8]

A consumer can choose whether and where to deposit her endowment in period 0, and an intermediary can then choose how to invest this endowment on her behalf. In period 1, consumers learn their type and observe the intermediary's investment choice from the initial period, and they can decide whether to withdraw based on this information.

Let us assume that competition among intermediaries forces them offer to the first-best demand-deposit contract (c_1^*, c_2^*) in exchange for the consumers' endowment. In period 1, consumers have the possibility to withdraw the promised amount of c_1^*, or to wait until period 2. We have assumed that an intermediaries investment decision I_{shirk} is not observable in $t = 0$, but becomes publicly observable before consumers make their withdrawal decision in $t = 1$.

Proposition 3.1 (Implementation of the first-best: demand-deposit contracts)**.** *There exists a subgame-perfect Nash equilibrium in which the first-best consumption profile (c_1^*, c_2^*) is implemented by the intermediaries offering demand-deposit contracts.*

Consider the following strategy of a consumer for the period-1 subgame: She withdraws if she turns out to be impatient or if the intermediary has chosen $I_{shirk} > 0$, and she does not withdraw if she turns out to be patient and the intermediary has chosen $I_{shirk} = 0$. We will now show that if all consumers use this strategy, this strategy profile constitutes a Nash equilibrium in the period-1 subgame for any investment decision of the intermediary, and the optimal strategy of the intermediary is to choose $I_{shirk} = 0$.

Assume that the intermediary has chosen $I_{shirk} > 0$. Because all other consumers withdraw, it is a best response to do so as well because the intermediary is illiquid and insolvent already in $t = 1$. Notice further that if I_{shirk}

[8]Notice that in case of unsecured wholesale funding, one would have to worry about the behavior of investors as well; compare Chapter 1. However, for the case of asset sales or collateralized lending, we do not have to worry about the investors' behavior as long as they cannot collude in order to extract rents from consumers.

is large enough, withdrawing actually becomes a dominant strategy because the intermediary will be illiquid and insolvent in $t = 2$ even without a run.

Now assume the intermediary has only invested in the productive technology, i.e., $I = 1$ and $I_{shirk} = 0$. Given that only impatient consumers withdraw, the intermediary will be able to serve all early consumers by selling L^* units of her investment to investors. Because $A \geq \xi = \pi c_1^*$ by assumption, the investors' funds are sufficient to serve all early depositors. As $c_2^* > c_1^*$, it is a best response for patient consumers to wait.

This withdrawal strategy is a credible punishment strategy, and it uses the threat of a bank run as a disciplining device: Because the intermediary anticipates that all consumers will withdraw in $t = 1$ whenever she invests in the shirking technology, she knows that she will not be able to enjoy the private benefit B. Therefore, she does not invest in the shirking technology in the first place. This disciplining effect of short-term debt is reminiscent of the findings of Calomiris and Kahn (1991), and Diamond and Rajan (2001), and allows intermediaries to implement the first-best allocation via demand-deposit contracts.

Note that there also exists a continuum of subgame-perfect Nash equilibria in which the bank chooses to invest a positive fraction in the shirking technology, but is not disciplined by the depositors up to this fraction. We discuss such equilibria in the Appendix 3.B. In the following, we restrict attention to the equilibrium proposed above. This is equivalent to assuming that an intermediary can only exclusively invest in either the productive or the shirking technology.

3.3.3 Fragility

While short-term debt is disciplining in our model, it is also a source of fragility. In fact, the model exhibits multiple equilibria in the period-1 subgame. Depending on the amount of investors' funds A, qualitatively different run equilibria emerge. As long as the amount of funds A is sufficiently large, potential runs on some intermediaries do not affect other intermediaries. However, if the endowment of investors A is relatively small, liquidity can become scarce in case of a run on many intermediaries. This puts the market for liquidity under stress and deteriorates the funding conditions of other intermediaries.

The price p of assets sold in period 1 depends on the aggregate amount L of assets sold if and only if the investors' resource constraint becomes binding. As long as the resource constraint is not binding, competition among investors ensures that the price is equal to the investors' willingness to pay. Thus, if

A is so large such that L units of the asset can purchased by investors at price $p = R/\gamma$, this is the market-clearing price, i.e., the price is equal to the assets' rate of return R divided by the rate of the investors' outside option γ. If, however, A is scarce relative to the amount L of assets sold (i.e., if A is not sufficient to purchase L units at price R/γ), the market clears via cash-in-the-market pricing, i.e., it must hold that $pL = A$.

Given L, the amount of assets sold, the price of the assets in period 1 is given by

$$p(L) = \begin{cases} \frac{R}{\gamma} & \text{if } A\frac{\gamma}{R} \geq L \\ \frac{A}{L} & \text{if } A\frac{\gamma}{R} < L. \end{cases} \tag{3.16}$$

Recall that by Assumption 3.1, we restrict our attention to the case in which intermediaries exclusively invest in the productive technology, i.e., $I^* = 1$, so the amount of assets sold is at most one. The price for assets in period 1 is depicted in Figure 3.2 for the case that $A < R/\gamma$.

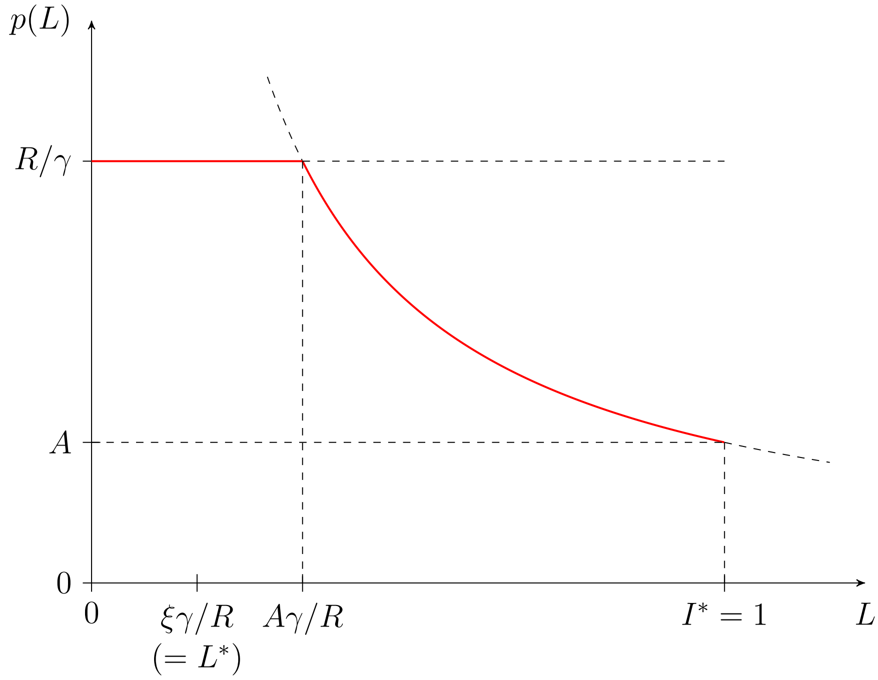

Figure 3.2: This graph depicts the potential fire-sale price for the case $R/\gamma > A$. In this case, a run may lead to depressed fire-sale prices via the binding cash-in-the-market constraint.

As long as $A \geq R/\gamma$, liquidity cannot become scarce, and the price is always given by R/γ. However, if $A < R/\gamma$, runs on some intermediaries can have negative external effects on others. If sufficiently many intermediaries experience a run, the price p gets depressed, thus deteriorating the refinancing condition of other intermediaries.

Micro-Fragility: Runs on Single Institutions
Let us start by considering the stability of a single intermediary. Notice that, on the one hand, the price on the secondary market is limited by the investors' willingness to pay, i.e., $p \leq R/\gamma$, but, on the other hand, the optimal demand-deposit contract promises an early consumption level that is strictly larger than this amount, $c_1^* > R/\gamma$. Because $p < c_1^*$, it holds true that if all depositors of one specific intermediary i run, this intermediary has to sell all assets, but still cannot fulfill all her obligations to her depositors. This particular intermediary becomes illiquid and insolvent already in period 1, and in particular could not serve any late consumer. Thus, a run on intermediary i constitutes an equilibrium.

Lemma 3.2 (Single-institution runs)**.** *Assume that intermediaries choose the first-best investment level and demand-deposit contract. There exists a Nash equilibrium in the period-1 subgame in which there is a run on some intermediary i, inducing a complete asset sale and immediate illiquidity and insolvency of this intermediary. In particular, there exists an equilibrium in which there is a run on all intermediaries.*

Notice that the run on a mass j of intermediaries does not necessarily affect the remaining mass $1-j$ of other intermediaries. If there is sufficient investor capital A, the price on the market remains high enough to make it possible that there exists an equilibrium where some mass j of intermediaries face a run, but the rest does not face a run. The reason is that if A is large enough and if (conditional on A) the mass j of intermediaries who face a run is sufficiently low, the price in the secondary market is high enough to make "prudent" behavior at the intermediaries $1-j$ compatible in equilibrium with runs elsewhere. Nonetheless, it may be true that all intermediaries are experiencing a run at the same time.

Macro-Fragility: Systemic Runs and Cash-in-the-market-pricing
Notice first that, if $A > R/\gamma$, it holds that $p(L) = R/\gamma$ for all L. This means that even in case of an economy-wide run, the price on the secondary market is unaffected and there is no binding cash-in-the-market constraint. This also implies, that if all intermediaries except for i had a run, this run would not affect i at all, because it can sell the designated amount L^* at the expected price $p = R/\gamma$, so it could refinance at the ex-ante expected conditions.

Now consider the case of $A < R/\gamma$. This implies that there is cash-in-the-market pricing in case of an economy-wide run, implying that $p(1) = A$. Thus, if all intermediaries but one are experiencing a run, the intermediary who is not experiencing a run will yet face deteriorated funding conditions. We refer to runs as "systemic runs" if they induce cash-in-the-market pricing and thus affect the overall funding conditions.

Proposition 3.2 (Systemic runs). *Assume that $A < \frac{R}{\gamma}$, and assume that intermediaries choose the first-best investment level and demand-deposit contract. Then there exist "systemic runs", i.e., an economy-wide run in the period-1 subgame leads to cash-in-the-market pricing and thus a deterioration of overall funding conditions.*

Proposition 3.1 shows that the ability to withdraw early induces diligent behavior of the intermediary, so short-term debt has a disciplining effect in our model. However, there always exist multiple equilibria. In one class of equilibria, only some single institutions experience runs while others do not, and the latter ones remain completely unaffected. From Lemma 3.2 we learn that runs are always possible (on single institutions, but also economy-wide runs). However, the runs on single institutions occur independently. From Proposition 3.2 we learn that runs are contagious via deteriorated funding conditions only if $A < R/\gamma$, i.e., if investor capital is scarce. Whenever $A < R/\gamma$, there exists a second class of equilibria, in which runs also become contagious in the sense that they affect funding conditions of other institutions. The second type of run will be particularly important when we analyze later how runs in the shadow banking sector may affect funding conditions for the regulated banking sector.

Finally, it is important to notice that there is an implicit assumption underlying the existence of systemic runs: Our model does not allow that funds withdrawn at one shadow bank immediately re-enter the system as deposits at another intermediary or via the secondary market. Without this restriction, further frictions would be needed to explain systemic runs, such as frictions in interbank trade as pointed out by Skeie (2008).

3.3.4 Financial Markets Implementation

Until this point, we have by assumption ignored the possibility of implementing an allocation via a financial market instead via intermediaries. The allocation that can be attained via a financial market in which consumers invest in the technologies directly and trade with investors in $t = 1$, is $(c_1^{fm}, c_2^{fm}) = (R/\gamma, R)$. This allocation, however, only coincides with the first-best if $\eta \to 1$, i.e., if $u(c) = \ln(c)$. This is reminiscent of the result of Diamond and Dybvig (1983).

Nonetheless, we need to make the investment restriction: If we allowed for the coexistence of financial markets and intermediaries, the incentive to conduct side trading would destroy the ability to implement the first-best via intermediaries, due to the same reasoning as in Jacklin (1987) and Farhi et al. (2009). If intermediaries offered the first-best demand-deposit contract, a consumer has an incentive to invest his endowment in the productive technology and consume the returns $R > c_2^*$ if he turns out to be patient, and to trade with a patient depositor otherwise, thereby consuming c_1^* units.

3.4 Deposit Insurance and Optimal Bank Regulation

As we have seen in the previous section, the first-best is implementable through non-regulated intermediaries, but this implementation is fragile in the sense that there always exist run equilibria in the period-1 subgame. To eliminate such panic-based bank runs, assume that the regulator provides a credible deposit insurance:[9] The regulator guarantees each bank depositor that she will receive at least the amount in $t = 2$ that she was promised in $t = 1$.

In a setup without aggregate uncertainty and with multiple equilibria, introducing a deposit insurance that is credible may eliminate the adverse run equilibrium at no cost, e.g., as discussed by Diamond and Dybvig (1983). By guaranteeing patient consumers that they will get at least as much in the final period than in the interim period, the strategic complementarity is eliminated. Thus, the deposit insurance is never tested in equilibrium and is costless.[10]

In our setup, however, a deposit insurance – if implemented without further regulatory policy measures – can give rise to opportunistic behavior on the part of intermediaries, which imposes costs on the provider of such deposit

[9] Alternatively, one could consider a lender of last resort, who grants access to the discount window of the central bank in order to prevent panics. The role of the lender of last resort is discussed in a later section.

[10] An alternative measure often discussed in the literature is to allow intermediaries to suspend convertibility. One can easily see that the discussion below would be equivalent under suspension of convertibility: suspension of convertibility may successfully prevent panic-based runs, but also undermines the disciplining effect of demand deposit contracts. If banks are able to suspend convertibility, regulation will also be necessary to ensure diligent behavior of intermediaries.

insurance. The reason is that in the presence of deposit insurance, consumers do not care about the investment behavior of the intermediary, thus eliminating the disciplining effect of short-term debt. Even if they know that the intermediary will be insolvent in the second period, they do not run because they know that the deposit insurance will pay them at least the amount that the demand deposit contract entitles them to withdraw in the interim period. Therefore, an intermediary may have incentives to invest in the shirking technology.

Given the moral hazard problem arising from the deposit insurance, there exists an optimal regulatory response. In the first-best, the intermediary does not invest any of his funds in the intermediation business. The intermediary has no skin in the game, and the participation constraint $e_2 \geq \rho e_0$ is trivially satisfied by $e_0 = e_2 = 0$. This is efficient because there is no need to provide the intermediary with incentives, and given $\rho > R$, it would be costly for consumers to use the intermediary's funds.

If a regulator wants to rule out moral hazard, she can do so by requiring the intermediaries to hold junior claims on their intermediation business. Optimal regulation thus calls for a minimal equity requirement via a classic skin-in-the-game argument. To insure diligence, the incentive compatibility constraint of the intermediary has to be satisfied. It is given by

$$e_2 \geq (1 + e_0)B. \tag{3.17}$$

At the same time, the intermediary's participation constraint, $e_2 \geq \rho e_0$, still needs to be fulfilled.

In the second-best, both constraints are binding, i.e., $e_2 = (1 + e_0)B$ and $e_2 = \rho e_0$, yielding the second-best equity stakes

$$e_0^{**} = \frac{B}{\rho - B}, \quad \text{and} \tag{3.18}$$

$$e_2^{**} = \frac{\rho B}{\rho - B}. \tag{3.19}$$

Because it is costly to use intermediary's funds, it holds that in an optimal regulatory regime that tries to prevent the intermediary from investing in the shirking technology, as little as possible intermediary capital is used, but enough to ensure diligent behavior. Given this necessary equity level, we can derive the second-best demand-deposit contract. The FOC is again given by $u'(c_1) = \gamma u'(c_2)$.

Proposition 3.3 (Second-best contract)**.** *Assume that demand deposits are* *protected by a credible deposit insurance. Optimal bank regulation requires intermediaries to satisfy an equity-to-debt ratio of $B/(\rho - B)$, and intermediaries will hold exactly $e_0^* = B/(\rho - B)$. There exists no run equilibrium in the period-1 subgame. Given that $\xi \leq A$, investment and sales are given by*

$$I^{**} = 1 + \frac{B}{\rho - B} \quad \text{and} \quad L^{**} = \frac{\pi\gamma^{1-\frac{1}{\eta}}}{R} \frac{R - \frac{B}{\rho - B}(\rho - R)}{(1-\pi) + \pi\gamma^{1-\frac{1}{\eta}}}, \tag{3.20}$$

and the optimal consumption is given by

$$c_1^{**} = \gamma^{-\frac{1}{\eta}} \frac{R - \frac{B}{\rho - B}(\rho - R)}{(1-\pi) + \pi\gamma^{1-\frac{1}{\eta}}} \quad \text{and} \quad c_2^{**} = \frac{R - \frac{B}{\rho - B}(\rho - R)}{(1-\pi) + \pi\gamma^{1-\frac{1}{\eta}}}. \tag{3.21}$$

In the regime with a deposit insurance, the consumption levels are decreasing in the private benefit B as well as in the required return of intermediaries ρ. Obviously, first-best (Lemma 3.1) and second-best coincide if $B = 0$ or $\rho = R$. For any other $B > 0$ and $\rho > R$, the second-best consumption levels are strictly lower. In fact, the case of $B > 0$, but $\rho = R$, is very interesting. In this case, using intermediary capital is not costly, and the first-best can always be implemented by using intermediary capital and investing it in the production technology until incentives are provided.

Importantly, there are no run equilibria in the interim period. The allocative inefficiency comes with the benefit of financial stability. However, as we will emphasize in the next section, this overall stability can only be attained if we exclude the possibility of regulatory arbitrage.

3.5 Regulatory Arbitrage and Fragility

In the previous section, we have abstracted from the possibility of regulatory arbitrage. In the following, we assume that the regulator provides a deposit insurance and regulates those intermediaries who are covered by the deposit insurance, hereafter referred to "commercial banks" or "regulated banks". However, we assume that it is also possible for intermediaries to place themselves outside of the regulatory perimeter of banking. Intermediaries who engage in this kind of regulatory arbitrage are referred to as "shadow banks" in the following. In this case, they will neither be regulated nor covered by the deposit insurance. However, shadow banks are disciplined in their investment behavior by demand-deposit contracts.[11]

In the following, we want to analyze a situation in which both regulated commercial banks and shadow banks exist, and analyze how systemic risk that emerges in the shadow banking sector can spread to the commercial banking sector. In our model, the coexistence of regulated banking and shadow banking results from an additional assumption concerning the heterogeneity of the depositors with regard to their "taste for commercial banking", which we model as a cost of moving to a shadow bank as in Luck and Schempp (2014a).

In period 0, consumers can choose whether to deposit their endowment in a regulated bank or in a shadow bank. We assume that consumers do not care about potential self-fulfilling runs on shadow banks, as runs are considered to be zero-probability events. In a later section, we will relax this assumption. Nonetheless, depositing at a shadow bank is assumed to come at some opportunity cost. We assume that investors are initially located at a regulated bank. For some consumer i, switching to a shadow bank comes at a cost of $s_i \in \mathbb{R}^+$, where s_i is independently and identically distributed according to the distribution function G. We assume that G is a continuous function that is strictly increasing on its support $\mathbb{R}^+$, and that $G(0) = 0$. The switching cost is assumed to enter into the investors' utility additively separable from the consumption utility.

This switching cost should not be taken literally. We have two preferred interpretations in mind: First, one can think of s_i as a screening a cost or as the cost of monitoring a shadow bank's investment decision and thus ensuring its diligent behavior. Alternatively, one can interpret s_i as the forgone service benefits that depositors lose when leaving commercial banks, such as payment services and ATMs. Both types of costs are likely to be heterogeneous across the population of depositors. It may be relatively more costly to monitor for a consumer that deposits only a small amount with the intermediary than for a firm that deposits larger sums. Likewise, consumers may care more about the payment services that banks provide.

As above, it is important to notice that the subsequent analysis also relies on the assumption that funds withdrawn do not re-enter the system immediately. Funds withdrawn at a shadow bank in $t = 1$ cannot be redeposited at the regulated banking sector, nor can they be invested in the secondary market. Thus, we assume that funds do not fully re-enter the banking system

[11] While by legal standards shadow banks have historically not offered demand deposits in reality, they do issue claims that are essentially equivalent to demand deposits, such as equity shares with a stable net assets value (stable NAV), or other instruments such as asset-backed commercial papers or repurchase agreements. For tractability, we will assume that shadow banks are literally taking demand deposits.

immediately. This is equivalent to assuming that no frictionless transfer from the shadow banking sector to the banking sector or to the secondary market is possible.

Coexistence of Banks and Shadow Banks

First of all, consider the contracts that can be offered. A regulated bank will offer a contract as discussed in Proposition 3.3, so the consumption levels promised are given by $(c_1^b, c_2^b) = (c_1^{**}, c_2^{**})$, where the superscript b stands for *bank.* The expected utility of a bank customer is decreasing in B and ρ. A shadow bank will offer the unconstrained contract described in Lemma 3.1, so the consumption profile is given by $(c_1^{sb}, c_2^{sb}) = (c_1^*, c_2^*)$, where the superscript sb stands for *shadow bank.*

When born, a consumer can decide whether he wants to remain at a bank, or pay the switching costs and get the consumption profile of the shadow banking contract. Thus, a consumer decides to become a customer of the shadow bank if

$$U(c_1^{sb}, c_2^{sb}) - s_i \geq U(c_1^b, c_2^b).$$

Let us define $s^*(B, \rho)$ such that a consumer with switching cost $s_i = s^*$ is indifferent between the two sectors, i.e.,

$$s^* = U(c_1^{sb}, c_2^{sb}) - U(c_1^b, c_2^b). \tag{3.22}$$

Lemma 3.3 (Size of shadow banking)**.** *The size of the shadow banking sector is given by $G(s^*)$. The cost threshold s^* that makes consumers indifferent, and thus also sector size G^*, is increasing in B and ρ. It holds that $G^* = 0$ if $B = 0$ or $\rho = R$.*

The relative size of the two sectors is determined by the distribution G as well as the two parameters B and ρ. The size of the shadow banking sector increases with the threat of moral hazard (i.e., in the intermediaries' private benefits from shirking B) and in the cost of intermediary equity ρ.

Systemic Runs on Shadow Banks and Contagion

We are now equipped to derive conditions under which runs in the shadow banking sector affect the consumption allocations in the regulated banking. We show that if the shadow banking sector is large enough, it will induce cash-in-the-market pricing and thus be contagious via a deterioration in funding conditions for regulated banks.

Recall that $c_1^{sb} \geq R/\gamma$. This implies that a run on one individual shadow bank constitutes an equilibrium independent of the amount of investors' capital A. That is, because the claims of all consumers withdrawing early cannot be satisfied via complete liquidation there can always be runs on single institution, compare Lemma 3.2. Moreover, a run on all shadow banks also constitutes an equilibrium for the same reason. Such a run on all shadow banks is not necessarily contagious. However, if the shadow banking sector is sufficiently large, or the amount of investors' capital too small, then the run on shadow banks soaks up so much liquidity that the funding conditions for commercial banks deteriorate.

We now analyze under which conditions a run on shadow banks leads to cash-in-the-market pricing. Observe that, in case of a run on shadow banks, the fire sale price of assets is given by

$$p(s^*) = \begin{cases} R/\gamma & if \quad G(s^*) \leq \bar{g} \\ \frac{A-(1-G(s^*))\pi c_1^b}{G(s^*)} & if \quad G(s^*) > \bar{g}. \end{cases} \tag{3.23}$$

In case of a run on the shadow banking sector, all shadow banks try to serve all withdrawing depositors. In order to fulfill their obligations, they sell all their assets, i.e., the shadow banking sector sells a total amount of $G(s^*)$ units. As long as there is no cash-in-the-market pricing, shadow banks thus absorb an amount of liquidity $G(s^*)R/\gamma$. However, this does not suffice to serve all customers because $c_1^{sb} > p \geq R/\gamma$. In addition, commercial banks also need an amount $[1 - G(s^*)]\pi c_1^b$ of liquidity to satisfy their withdrawing impatient consumers. A run on shadow banks is not compatible with a price $p = R/\gamma$ (i.e., it leads to cash-in-the-market pricing) and thus has a negative effect on regulated banks if the sum of these two terms exceeds the available funds A.

Proposition 3.4 (Coexistence and runs on shadow banks)**.** *When regulated banks and shadow banks coexist, a run on all shadow banks leads to cash-in-the-market pricing whenever the shadow banking sector is large relative to the budget of investors, specifically, if*

$$G(s^*) > \frac{A - \pi c_1^b}{R/\gamma - \pi c_1^b} = \bar{g}(A) > 0. \tag{3.24}$$

In this case, the run also affects commercial banks because their funding conditions deteriorate. There exists some threshold, $\tilde{g}(A) > \bar{g}(A)$, such that in case of a run, regulated banks become insolvent if $G(s^*) > \tilde{g}(A)$.

If $A < R/\gamma$, it holds that $\bar{g}(A) < 1$, i.e., the price in a run decreases whenever the shadow banking sector is large enough. Although regulated banks cannot be subject to runs because they are covered by the deposit insurance, they are affected via a deterioration of their funding conditions as the cash-in-the-market constraint becomes binding due to the runs on shadow banks. That is, in case of a run on all shadow banks, the asset price in the secondary market will be lower than initially expected. Thus, commercial banks will have to refinance themselves at worse conditions than initially expected. However, the banks' equity cushion may be able to absorb these losses. Given a relatively large shadow banking sector, the fire-sale price will be so low such that banks become insolvent in $t = 2$, or already illiquid and insolvent in $t = 1$. Importantly, this kind of contagion does not stem from direct contractual linkages between the two sectors such as explicit or implicit liquidity guarantees.

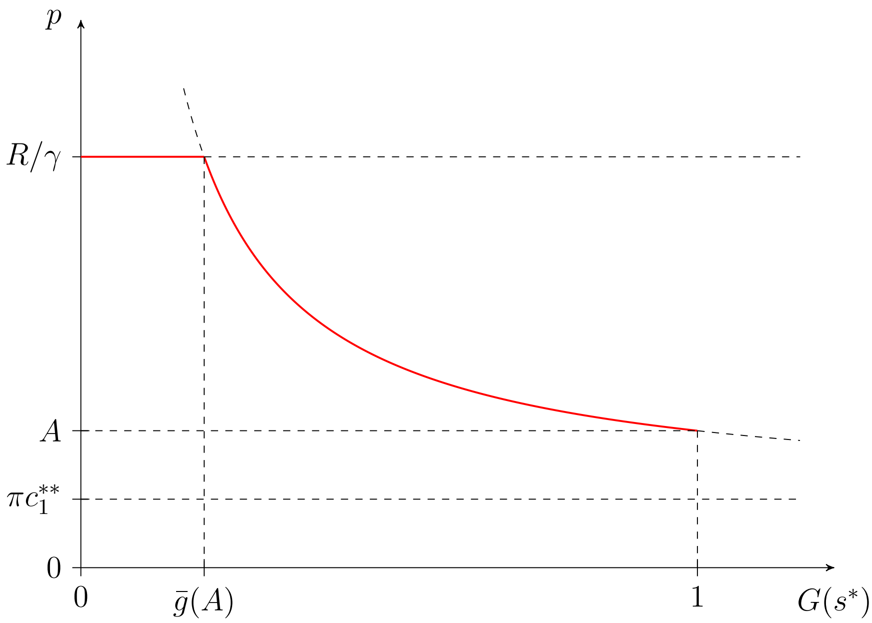

Figure 3.3: This graph depicts the fire-sale price in case of a run on all shadow banks. A run leads to cash-in-the-market pricing and thus affects regulated banks if $G(s^*) > \bar{g}(A)$.

In the previous section, we have shown that if there is deposit insurance and regulation, but no regulatory arbitrage, it holds that the economy attains an allocation which is inefficient compared to the first-best, but exhibits no fragility. If we assume that regulatory arbitrage is possible, this may no longer be true. Moreover, fragility is not contained in the shadow banking sector only. Whenever the shadow banking sector is large relatively to the available capital of investors, runs on shadow banks become possible and can affect regulated banks via a deterioration of funding conditions. Importantly, the contagion operates only via the asset market and thus there can be contagion even in the absence of any contractual linkages.

3.6 Wholesale Funding Restrictions

A natural question that arises in this context is whether limiting wholesale funding of commercial banks can improve financial stability. In the following, we analyze the most simple and the most extreme case: a complete prohibition of wholesale funding for regulated banks, i.e., a restriction that $L = 0$ for regulated banks. We show that this shuts down the contagion channel, but the associated allocation is less efficient and implies a further growth of the shadow banking sector.

The optimal consumption profile a bank can offer when wholesale funding is prohibited is identical to the second best for the case of $\gamma = R$. In this latter case, the bank could sell assets in the interim period, however, the return of doing so is identical to using storage. The FOC is given by $u'(c_1) = Ru(c_2)$.

Lemma 3.4. *The constrained optimal allocation under the prohibition of wholesale funding requires an investment of*

$$I^r = 1 + \frac{B}{\rho - B} - \pi c_1^r,$$

and the consumption profile is given by

$$c_1^r = R^{-\frac{1}{\eta}} \frac{R - \frac{B}{\rho - B}(\rho - R)}{\pi R^{1-\frac{1}{\eta}} + (1-\pi)}, \quad \textit{and} \quad c_2^r = \frac{R - \frac{B}{\rho - B}(\rho - R)}{\pi R^{1-\frac{1}{\eta}} + (1-\pi)}.$$

It holds that

$$U(c_1^b, c_2^b) > U(c_1^r, c_2^r).$$

Note that this optimal contract under the exclusion of wholesale funding (c_1^r, c_2^r) coincides with the first-best allocation of the Diamond and Dybvig model for $B = 0$ or $\rho = R$. However, this allocation is ont optimal in the presence of investors that can provide liquid funds in the interim period.

Shadow banks still offer the same contract as above, i.e., $(c_1^{sb}, c_2^{sb}) = (c_1^*, c_2^*)$. Given the switching cost s_i, let us define $s^r(B, \rho, L = 0)$ such that a consumer with switching costs $s_i = s^r$ is indifferent between the two sectors, i.e.,

$$s^r = U(c_1^{sb}, c_2^{sb}) - U(c_1^r, c_2^r). \tag{3.25}$$

One can directly see that the restrictions will shield regulated banks from the adverse consequence of a run in the shadow banking sector, but that at the same time they lead to a larger shadow banking sector in equilibrium. That is, we will find that $s^r > s^*$.

Proposition 3.5. *Wholesale funding restrictions successfully shield regulated banks from the adverse consequences of runs in the shadow banking sector. However, the shadow banking sector is larger than without these restrictions,* $G(s^r) > G(s^*)$.

Wholesale funding restrictions, together with a deposit insurance, eliminate the fragility altogether. They do, however, induce a further allocative inefficiency by deteriorating the consumption profile of bank customers and by pushing more depositors into the shadow banking sector.

3.7 Sunspot Runs in Shadow Banking

So far, we have assumed that runs on shadow banks are zero-probability events. Clearly, if agents anticipated that such runs occur with a positive probability, the expected utility from holding the same claim on the shadow banking sector would be lower. However, we argue that while it does change the incentive to switch to a shadow bank, it does not lead to a breakdown of the shadow banking sector in general. As long as the probability of a run is small, there is a positive mass of consumers who still prefer shadow banks over regulated banks. Thus, the shadow banking sector shrinks, but continues to have a positive size. Moreover, there is a second effect of systemic runs that partially offsets the direct effect: Because systemic runs on shadow banks worsen the wholesale funding conditions of regulated banks, runs on shadow banks also worsen the expected consumption profile that regulated banks can offer.

In order to illustrate the robustness of our results with regard to positive-probability runs, we analyze a variant of the model in which a run on the shadow banking sector occurs with an exogenous sunspot probability q, as in Cooper and Ross (1998). In order to keep the analysis tractable, we assume that the intermediaries are risk-neutral and that intermediaries' equity is sufficient to be completely loss-absorbing.

Observe that if q was high and the endowment of investors A was small, shadow banks would at some point change their investment behavior by also investing in storage.[12] This would shield them against low fire-sale prices in

[12] Cooper and Ross (1998) show that if the physical liquidation rate is sufficiently high and

case of a run. However, to keep things simple, we make the assumption that $A > 1$ in this section. In this case, storage is still dominated by wholesale funding for any q because $A > 1$ implies that even if we have cash-in-the-market pricing, the price can never fall below 1. This in turn implies that for any probability of a run on shadow banks q, shadow banks will not invest in storage. It follows that the contract that is offered by shadow banks does not depend on q, so it is still given by the contract (c_1^{sb}, c_2^{sb}), as specified in Section 3.5. If there is no run, the utility of a shadow bank customer is given by $U(c_1^{sb}, c_2^{sb})$. However, in case of a run, shadow banks will not be able to satisfy this contract. In case of a run, shadow banks sell all assets at price p^r, and the consumers' utility is given by $u(p^r)$. The utility of shadow bank customers decreases in q because the probability of the unfavorable event decreases directly, and because the price p^r also decreases if shadow banks are systemic (as we will see later).

Let us now consider the effect of sunspot runs on regulated banks. For a sunspot probability q, let (c_1^q, c_2^q) denote the contract offered by regulated banks, and let s^q denote the level of switching cost that makes a consumer indifferent between shadow banks and regulated banks. It is given by

$$s^q = (1-q)U(c_1^{sb}, c_2^{sb}) + qu(p^r) - U(c_1^q, c_2^q). \qquad (3.26)$$

If the parameters are such that runs in the shadow banking sector are not systemic for the case of $q = 0$ (i.e., if $G(s^*) \leq \bar{g}$, which is equivalent to the situation of runs as zero-probability events, as discussed before), then it holds that a positive bank-run probability does not affect regulated banks at all either. If runs do not induce cash-in-the-market pricing in this case, they will not induce cash-in-the-market pricing for positive q either. Regulated banks can thus always receive wholesale funding at $p = R/\gamma$, so the contract offered by regulated banks and the utility of bank customers is not changed. However, if the run on shadow banks realizes, shadow bank customers only get the proceeds from selling the assets, i.e., $p = R/\gamma < c_1 sb$. Thus, shadow banking becomes unambiguously less attractive, and since the utility from staying at a regulated bank is unchanged, and the size of the shadow banking sector decreases.

We saw that the expected consumption utility of a shadow bank customer (given by the first two terms) is decreasing in q. For $G(s^*) < \bar{g}$, we saw that the contract of commercial banks does not depend on q, so s^q and thus

the probability for the sunspot run is also sufficiently high, banks will prefer to exclude the possibility of sunspot runs by investing in storage.

the size of the shadow banking sector, $G(s^q)$, shrinks in q. Now assume that $G(s^*) > \bar{g}$. If we now introduce the sunspot probability q, the expected utility of a deposit contract offered by a regulated bank will decrease as well.

Assuming $G(s^*) > \bar{g}$, it still holds true that the utility of shadow bank customers decreases in q. However, because a run on shadow banks can be systemic and induce cash-in-the-market pricing, the run also negatively affects commercial banks. Because banks anticipate this, their optimal contract changes and commercial banking also becomes less attractive. This reduces the limiting effect of the positive probability of bank runs on the size of the shadow banking sector.

Let us first consider the contract that banks are offering, given that a run will hit the shadow banks with probability q. We have the two states, “run” r and “no run” n. If no run occurs, the price in the secondary market is $p^n = R/\gamma$. In case of a run, the price is denoted by p^r.

We assume that equity is completely loss-absorbing, i.e., the equity is sufficient to buffer the losses in case of a run, and the bank can pay out the amounts promised in the demand-deposit contract without having to use the deposit insurance. Let $e_2 = qe_2^r + (1-q)e_2^n$ denote the expected payoff of the intermediary in period 2. Binding IC and IR condition of intermediaries still imply that $e_2 = \frac{\rho B}{\rho - B}$, and $e_0 = \frac{B}{\rho - B}$.

For each state $i = n, r$, the budget constraints are given by

$$\pi c_1^q = L^i p^i, \quad \text{and} \tag{3.27}$$

$$(1-\pi)c_2^q = (1+e_0)R - e_2^i - L^i R. \tag{3.28}$$

The FOC is given by

$$u'(c_1^q) = u'(c_2^q) R \left(\frac{1-q}{p^n} + \frac{q}{p^r} \right). \tag{3.29}$$

Note that, for $q = 0$, we get the well-known FOC $u'(c_1^q) = \gamma u'(c_2^q)$.

The positive probability of a bank run has two negative effects on bank customers’ utility: First, it tightens the budget constraint because banks’ funding conditions deteriorate in a run, but intermediaries’ participation constraint still needs to be satisfied. For the same reason, the optimal marginal rate of substitution between late and early consumption increases (see FOC), which means that there will be less consumption smoothing. Thus, it holds that for any sunspot probability $q > 0$, the bank contract (c_1^q, c_2^q) yields a lower utility than for the case of $q = 0$, which is given by (c_1^b, c_2^b):

$$U(c_1^q, c_2^q) < U(c_1^b, c_2^b). \tag{3.30}$$

Thus, if we compare the utility that consumers get from commercial banks and shadow banks, respectively, the direct effect of shadow bank runs on the attractiveness of shadow banks is partly offset by an indirect effect on shadow banks via the wholesale funding channel.

3.8 Liquidity Guarantees

Until this point, we have restricted attention to intermediaries becoming either regulated banks or unregulated shadow banks. An interesting question is how our results change when banks and shadow banks are interdependent not only via effects on secondary markets, but if they are operated by the same intermediary. This had been practice prior to the recent financial crisis, as documented by Acharya et al. (2013), and has been targeted by post-crisis reforms.

To this end, we analyze a version of our model in which intermediaries operate a bank and a shadow bank at the same time. We investigate under which conditions intermediaries may have incentives to use funds from their regulated banking branch to support their shadow-banking activities in case of distress. We will thus analyze the effect of explicit contractual linkages.

Private Optimality of Liquidity Guarantees

In the previous section, we assumed that an economy-wide run is a sunspot phenomenon that occurs with a probability q. We will now again assume that economy-wide runs occur with zero probability.

In contrast, we assume that shadow banks experience idiosyncratic sunspot runs: With a probability q_i, each individual shadow bank experiences a run. However, we will assume that this probability is very small, i.e., $q_i \to 0$. This allows us to keep the optimal contracts fixed. A commercial bank optimally offers (c_1^b, c_2^b) as before, and a shadow bank (c_1^{sb}, c_2^{sb}).

Given that there is a run on the shadow bank with probability q_i, the intermediary may now have an incentive to guarantee the liquidity of her shadow bank to protect her from idiosyncratic runs. Observe that the possibility of an idiosyncratic sunspot run can be eliminated if a regulated branch of an intermediary provides a credible liquidity guarantee for its unregulated operations. Moreover, observe that it is optimal to provide this support guarantee for each institution as it makes the offered contract more attractive and will thus attract more consumers.

The liquidity guarantee is credible if the bank can serve all its impatient bank customers as well as all those who own a shadow bank contract by selling all its assets. The regulated bank has conducted an initial investment of $(1-G(s^*))(1+e_0^*)$ and the shadow banking an investment of $G(s^*)$. In total, the bank can thus raise $R/\gamma(1+(1-G(s^*))e_0^*)$. The funds raised by selling all assets need to be sufficient to serve all impatient consumers, $(1-G(s^*))c_1^b$, of the regulated bank and all consumers of the shadow bank, $G(s^*)c_1^{sb}$. A liquidity guarantee is thus credible whenever

$$G(s^*) \leq \frac{R/\gamma(1+(1-G(s^*))e_0^*)-\pi c_1^b}{c_1^{sb}-\pi c_1^b}$$

A guarantee can thus only be credible if the shadow banking operations are not too large compared to the regulated banking activities.

Systemic Runs

While it is optimal from the perspective of a single institution to provide a liquidity guarantee, it leads to an increased parameter space for runs on the aggregate level. In case of a run, commercial banks have to provide an amount of $G(s^*)c_1^{sb}$ to shadow banks. In addition, they require an amount of $(1-G(s^*))\pi c_1^b$ to satisfy their own impatient customers. Therefore, a run is systemic whenever $G(s^*)c_1^{sb}+(1-G(s^*))\pi c_1^b > A$.

Proposition 3.6. *Assume that intermediaries can own a regulated bank and a shadow bank at the same time and that liquidity guarantees are credible. It is privately optimal for each intermediary to guarantee the liquidity for her shadow bank branch by using funds from her regulated bank. In turn, this decreases the threshold size above which the shadow banking becomes systemic: A systemic run can now already occur and affect regulated banks if*

$$G(s^*) \geq \frac{A-\pi c_1^b}{c_1^{sb}-\pi c_1^b} < \bar{g}(A).$$

Without liquidity guarantees, systemic runs are only possible if the shadow banking sectors size exceeds $\bar{g}$. With liquidity guarantees, this is already true for a sector size of $G(s^*) \geq \frac{A-\pi c_1^b}{c_1^{sb}-\pi c_1^b}$. The underlying mechanism is as follows: while liquidity guarantees are optimal from the individual intermediary's perspective, they increase the number of assets sold in case of an unanticipated economy-wide run. This shows that there is a clear benefit of preventing direct contractual linkages via regulation, as it reduces the parameter space in which systemic runs may take place. However, as shown in the earlier section, it may not be sufficient to rule out adverse effect for regulated banks entirely.

3.9 Lender of Last Resort and Market Maker of Last Resort

So far, we have only considered a deposit insurance and capital regulation as policy measures, and we have ignored other government interventions. In practice, regulated banks also have access to the discount window of a central bank (Lender of Last Resort, LoLR). Alternatively, the central bank may decide to intervene in the secondary market in order to prevent cash-in-the-market pricing (Market Maker of Last Resort, MMLR). We now want to ask whether the possibility of such interventions may help to shield commercial banks from turmoil in the shadow banking sector.

Assume that there is an institution called central bank that has unlimited funds at its disposal in the interim period and can commit to being a LoLR or a MMLR. In a richer setup that distinguishes between nominal and real values, one could also consider potential costs of central bank interventions. We abstract from such trade-offs in our analysis and briefly discuss them at the end of this section.

Let us consider a central bank that acts as a LoLR or a MMLR, and analyze the policy it would choose in order to protect regulated banks from systemic crises. In our setup, a LoLR would intervene in case of a systemic crisis (i.e., if $p < R/\gamma$). Regulated banks would be allowed to use the discount window to borrow funds, so they do not have to sell assets at fire-sale prices. If we assume that the discount window lending has the same terms as the wholesale funding in normal times, then LoLR effectively shields regulated banks from turmoil in the shadow banking sector. Thus, banks can offer the second-best contract even in the case of crises occurring with a positive probability. However, shadow banks also benefit from such intervention, although they do not have access to the discount window. Because regulated banks do not have to sell their assets, fewer assets are on the market compared to a situation without central bank intervention, implying that fire-sale prices are higher.

A MMLR in our setup could ensure that the price of assets is always equal to R/γ by buying assets at this price in case of a systemic crisis. In contrast to the LoLR's discount window lending, the MMLR conducts an open market operation, and it appears plausible to assume that the MMLR cannot selectively buy assets from regulated banks only. As in the case of the LoLR, regulated banks are again unaffected by runs in the shadow banking sector and can offer the second-best. Shadow bank customers again benefit from the intervention, and in this case even more than in case of a LoLR intervention because the MMLR eliminates cash-in-the-market pricing altogether.

It is worth noticing that both the LoLR and the MMLR as discussed, are effective in shielding regulated banks, but they do not prevent runs in the shadow banking sector as the strategic complementary among consumers prevails independent of the wholesale funding conditions.

Therefore, we conclude that, for both types of interventions, it holds true that they make the shadow banking sector relatively more attractive in the presence of positive probabilities for runs and thus lead to an increased shadow banking sector; and this effect is more pronounced for the MMLR than for the LoLR intervention. Still, a LoLR or a MMLR that shields regulated banks from the adverse consequences of runs on shadow banks could be considered as an alternative to restrictions on wholesale funding. However, as mentioned above, our model is not suited to draw such conclusions in general, as we do not consider the costs and potential distortions resulting from interventions. While interventions have to be financed via taxation or inflation, there would be a trade-off in a richer setup that allows to distinguish between nominal and real values. Moreover, an ex-post intervention by a LoLR or a MMLR may not always be desirable in richer setups because of other potential frictions, e.g., if the government cannot distinguish between insolvent and illiquid banks or if interventions give rise to moral hazard.

3.10 Liquidity Regulation in Basel III

In the following, we relate our findings to the regulatory treatment of liquidity risk in the context of prudential regulation. Our model can shed light on the economic consequences of liquidity regulation in the Third Basel Accord (Basel III). Basel III introduces a new assessment and regulation of liquidity risk by defining two minimum standards of funding liquidity, first described in Basel Committee (2010). The two central measures are the Liquidity Coverage Ratio (LCR) and the Net Stable Funding Ratio (NSFR). The LCR requirement aims to ensure that a bank can withstand a “significantly severe liquidity stress scenario” with a horizon of 30 days and is described in detail in Basel Committee (2013). It aims to ensure that a bank has a sufficient stock of liquid assets in order to cover its liquidity needs during the next month. The objective of the NSFR requirement, elaborated in Basel Committee (2014), is to ensure stable funding over a one-year horizon. It requires a bank to have an amount of equity, long-term debt, and other "stable" funding that is sufficient to finance its stock of illiquid assets during the next year.

Our model can be interpreted as an argument in favor of liquidity regulation: it indicates that, in the presence of regulatory arbitrage, regulated banks may be affected by systemic liquidity crises originating in the shadow banking sector. This in turn may call for a macroprudential regulation of liquidity management whenever the regulator wants to shield regulated banks from the adverse consequences of runs in the shadow banking sector. However, we will now argue that it is important to use a macroprudential approach and in the context of the proposed liquidity regulation in Basel III this would imply a necessity for being more restrictive when defining which assets are considered as *liquid* and which sources of funding are considered as *stable*.

In the following, we briefly summarize the Basel committee's proposal for liquidity regulation and analyze its implications in the context of our model.

Liquidity Coverage Ratio (LCR)

The LCR is defined as the ratio of High Quality Liquid Assets (HQLA) and the (hypothetical) total net cash outflows over the next 30 calendar days, and Basel III requires this measure to be above 1.[13] Thus, the LCR sets a lower bound for the stock of liquid assets, conditional on a bank's (expected) cash flows. "Total net cash outflow" is defined as the maximum of "total expected net cash outflow" and "25% of total expected cash outflow". In this context, "expected" denotes a scenario of a "combined idiosyncratic and market-wide shock" that entails (among others) a partial run-off of retail deposits and a partial reduction in unsecured wholesale funding and secured short-term financing.

HQLA consist of two categories: Level 1 assets are cash, central bank reserves, and government bonds with 0% risk weight. Level 2 assets can again be divided in two sub-categories, Level 2A and Level 2B assets. A minimum haircut of 15% has to be applied to all Level 2 assets, which is supposed to capture their devaluation in a crisis scenario. After applying this haircut, Level 2 assets must not make up more than 40% of the whole stock of HQLA. Level 2A assets include government bonds with risk weights below 20%, as well as corporate debt securities (including commercial papers) and covered bonds with a rating of at least AA-. Level 2B assets also include Residential Mortgage-Backed Securities (RMBS) with ratings of at least AA, corporate debt securities with ratings of at least BBB-, and common equity shares which are constituent of a major stock index. These assets are subject to a haircut

[13]For details on the LCR, see Basel Committee (2013).

between 25% and 50% and must not make up more than 15% of the stock of HQLA.[14]

In addition to these requirements, the Basel Committee specifies liquidity requirements of eligible assets in the following way: Level 2 assets must be

> "traded in large, deep and active repo or cash markets characterised by a low level of concentration [and] have a proven record as a reliable source of liquidity in the markets (repo or sale) even during stressed market conditions (ie maximum decline of price not exceeding 10% or increase in haircut not exceeding 10 percentage points over a 30-day period during a relevant period of significant liquidity stress)."(Basel Committee, 2013).

The requirements for Level 2 assets are thus defined in terms of their past and present liquidity. The underlying notion seems to be that an asset's past and present liquidity predicts its future liquidity. This is particularly evident in the condition that an asset's value must have been stable in a "period of significant liquidity stress".

Net Stable Funding Ratio (NSFR)

As a second measure of liquidity regulation, Basel III requires the NSFR to be above 1. The NSFR is defined as the ratio of available stable funding (AFS) and required stable funding (RSF), both with a horizon of one year.[15] Thus, the NSFR sets a lower bound for the amount of stable funding, conditional on a bank's portfolio of illiquid assets and off-balance sheet exposures. The ASF is defined in order to capture the "capital and liabilities expected to be reliable over the time horizon [of] one year". ASF comprises regulatory capital, preferred stock, and liabilities with maturities of at least one year, but also

Liabilities of the latter categories have to be multiplied by an ASF factor of less than one. ASF aims to exclude unstable short-term funding, i.e., funding that might quickly be withdrawn or not rolled over. It excludes short-term wholesale funding, such as interbank lending, but includes retail deposits, because deposit insurance is supposed to make deposits a source of stable of funding.

RSF is a measure of a bank's illiquid asset portfolio. It is defined as the sum of the value of a bank's assets, multiplied by a specific RSF factor that should reflect an asset's liquidity risk, plus a similarly weighted sum of the bank's off-balance sheet activities or potential liquidity exposures. An asset's

[14]Note that they must also be included in the 40% cap of all Level 2 assets.

[15]For details on the NSFR, see Basel Committee (2014).

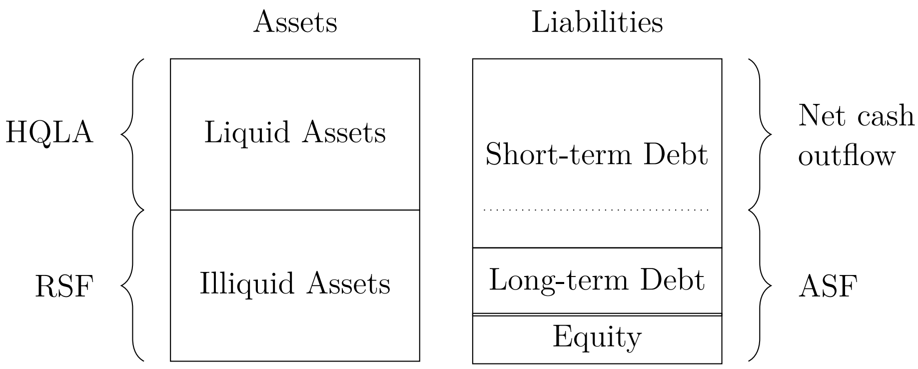

Figure 3.4: Schematic representation of the bank balance sheet under the assumption that LCR and NSFR have been defined for the same time horizon.

RSF factor is lower the more liquid this asset is. Cash and securities with a maturity of less than one year have an RSF factor of 0%; other securities and corporate bonds with good ratings have low, but positive RSF factors; other bonds, mortgages and loans have higher RSF factors, and other assets (particularly encumbered assets) have RSF factors of 100%.

The notion behind the NSFR requirement is that the ASF serves a bank to finance its illiquid asset contained in the RSF in times of a liquidity crisis. It is assume that those assets not contained in the RSF are liquid and can thus be sold even in times of a crisis in order to compensate the "unstable" funding that might disappear in a crisis.

Comparison and Discussion

Although the definitions of the LCR and NSFR appear quite different at first sight, a closer look reveals that their time horizon and the rhetoric are the only distinct differences. To illustrate this point, let us consider a stylized bank balance sheet; see Figure 3.4. The bank's assets can be divided into a portfolio of liquid and a portfolio of illiquid assets, and the liability side consists of short-term debt, long-term debt, and equity.

If we abstract from the different time horizons of the two liquidity measures, we see that the illiquid assets count as RSF, and the liquid assets count as HQLA. Which asset is considered to be liquid or illiquid is determined by the scenario of stress that is specified by the regulator. On the liability side, the stress scenario specifies which kind of funding is expected to disappear and which is expected to stay during a crisis. The expected net cash outflow measures the expected change in the bank's short-term liabilities. The part of short-term funding which (in the relevant scenario) is assumed not to disappear, together with long-term debt and equity, forms the ASF. Because total liabilities equal total assets, the LCR and the NSFR requirements are equivalent: HQLA exceed expected net cash outflow if and only if ASF exceeds RSF. It follows that the two measures only vary in their time horizon.

Both the LCR and the NSFR requirement are part of a micro-prudential regulatory approach. They aim to guarantee the liquidity of a single financial institution, but while they may assure that banks survive in normal times or in times of mild or idiosyncratic stress, it is questionable whether the underlying stress scenario is severe enough to capture potential systemic crises of a self-fulfilling nature.

Our model suggests that liquidity regulation may in general be considered as a useful tool. According to our model, liquidity regulation can be particularly useful if it has a macroprudential focus that allows shielding regulated banks from financial distress events that lie outside the regulated banking sector (compare Proposition 3.5). However, our model also indicates that we should use a macroprudential approach when it comes to defining what assets are considered to be *liquid* and which sources of funding are considered as *stable*.

With regard to the definition of HQLA, we argue that Level 2 assets in particular may not be very helpful in case of systemic liquidity crises. In our model, assets that are based on the productive technology can be traded in the interim period. In fact, these assets are traded at $p = R/\gamma$ in normal times. Moreover, even if some part of the financial system is in turmoil, the price may remain unaffected (as argued in Lemma 3.2 and in Proposition 3.4). This suggests that a real-life counterpart of the long asset in our setup could qualify as a Level 2 asset. However, in a systemic liquidity crisis, fire sales may depress the market value of these assets. This results from the fact that, in a setup with multiple equilibria, assets may be seemingly safe and the value might drop substantially and in a fashion that is not predictable by their past performance.

We further argue that the ASF may be too broad and, in particular, that short-term wholesale funding cannot be considered as a stable source of funding. While insured demand-deposits can be considered as stable funding, this does not apply to short-term wholesale funding. As our model indicates, the conditions of wholesale funding may be stable in normal times and even in case of stress events of parts of the financial system. However, they are not stable in case of a systemic liquidity crisis.

We thus argue that liquidity regulation as in Basel III may in principle be a useful tool to shield commercial banks from adverse consequences of turmoil that takes place outside the regulated banking sector. However, we argue that its focus is rather microprudential and it is thus not suited to deal with systemic crises. Consequentially, if we take our model seriously, the definition of HQLA and ASF should be tightened such that they are suited to deal with systemic liquidity crises.

3.11 Conclusion

This chapter provides a banking model in which intermediation optimally relies on wholesale funding next to funding via demand deposits. We show that in an unregulated banking environment, short-term debt has a disciplining role, but panic-based runs are also possible. If a safety net is in place to eliminate the possibility of panic-based runs, regulation becomes necessary. This in turn leads to the emergence of an unregulated and fragile shadow banking sector. We then show that because regulated banks optimally rely on wholesale funding, runs on shadow banks are contagious via a deterioration of the wholesale funding conditions. Therefore, turmoil in the shadow banking sector may affect the commercial banking sector even in the absence of any contractual linkages between the two sectors.

We emphasize two main findings of our analysis: The ideal policy measure would be to prevent the circumvention of regulation altogether. In the context of our model, this would allow a constrained-efficient allocation to be attained for in which financial stability is ensured by deposit insurance, and diligence is ensured through capital requirements. However, if regulatory arbitrage cannot be prevented altogether, attaining stability becomes difficult. One can consider ring-fencing the part of the banking sector that is covered by deposit insurance. However, besides prohibiting direct contractual linkages, one would then have to restrict or prohibit wholesale funding altogether. Yet, for such restrictions to be effective in fulfilling their macro-prudential objective, they need to be sufficiently strict. Moreover, achieving financial stability in the regulated banking sector by restrictions on wholesale funding can only be achieved at the cost of allocative inefficiency and a larger shadow banking sector.

Appendix 3.A First-Best

As discussed in the text, we are left with a maximization problem with two weak inequalities.

$$\max_{(c_1,c_2,I,L)\in\mathbb{R}^4_+} \quad \pi u(c_1) + (1-\pi)u(c_2), \tag{3.31}$$

$$\text{subject to} \quad \pi c_1 = (1-I) + LR/\gamma, \tag{3.32}$$

$$(1-\pi)c_2 = (I-L)R, \tag{3.33}$$

$$LR/\gamma \leq A, \tag{3.34}$$

$$I \leq 1. \tag{3.35}$$

Let us define the following thresholds:

$$\xi \equiv \gamma^{-\frac{1}{\eta}} \frac{\pi R}{(1-\pi) + \pi\gamma^{1-\frac{1}{\eta}}} < R/\gamma, \tag{3.36}$$

and

$$\xi_0 \equiv \frac{\pi R^{-\frac{1}{\eta}}}{(1-\pi) + \pi R^{-\frac{1}{\eta}}\gamma} < \xi, \tag{3.37}$$

Depending on the model parameters A, R, γ, and π, as well as on the shape of the utility function, the first-best program now has three solution candidates.

Lemma 3.5 (First-best)**.** *If $A \geq \xi$, then the first-best allocation is characterized by*

$$I^* = 1 \text{ and } L^* = \xi\gamma/R,$$

and optimal consumption is given by

$$c_1^* = \gamma^{-\frac{1}{\eta}} \frac{R}{(1-\pi) + \pi\gamma^{1-\frac{1}{\eta}}} \text{ and } c_2^* = \frac{R}{(1-\pi) + \pi\gamma^{1-\frac{1}{\eta}}}. \tag{3.38}$$

For $A \in (\xi_0, \xi)$, we have that

$$I^* = 1 \text{ and } L^* = A\gamma/R, \tag{3.39}$$

and optimal consumption is given by

$$c_1^* = \frac{A}{\pi} \text{ and } c_2^* = \frac{R - A\gamma}{(1-\pi)}. \tag{3.40}$$

Finally, if $A \leq \xi_0$, then the first-best allocation is characterized by

$$I^* = \frac{(1-\pi)(1+A) + \pi R^{-\frac{1}{\eta}} A\gamma}{(1-\pi) + \pi R^{-\frac{1}{\eta}}} \text{ and } L^* = A\gamma/R, \tag{3.41}$$

and optimal consumption is given by

$$c_1^* = R^{-\frac{1}{\eta}} \frac{R + (R-\gamma)A}{(1-\pi) + \pi R^{-\frac{1}{\eta}}} \text{ and } c_2^* = \frac{R + (R-\gamma)A}{(1-\pi) + \pi R^{-\frac{1}{\eta}}}. \tag{3.42}$$

In the first case ($A \geq \xi$), it holds that $I^* = 1$ and $L^* < 1$, and the optimal allocation is characterized by

$$u'(c_1) = \gamma u'(c_2) \tag{3.43}$$

$$\pi c_1 \gamma + (1-\pi)c_2 = R. \tag{3.44}$$

In the third case ($A < \xi_0$), we have $L^* = A\gamma/R$, and $I^* < 1$, and the optimal allocation is characterized by

$$u'(c_1) = Ru'(c_2) \tag{3.45}$$

$$\pi c_1 R + (1-\pi)c_2 = R + (R-\gamma)A. \tag{3.46}$$

Appendix 3.B Shirking Equilibria

We have stated above that in a subgame where the intermediary has chosen $I_{shirk} = 1$, it is the dominant strategy of all consumers to withdraw their funds early, so this constitutes the unique Nash equilibrium of this subgame. This illustrates the potential for short-term debt to be disciplining. Because the intermediary enjoys the benefit only when he does not experience a bank run in the interim period, there exists a run strategies in which consumers credibly threaten to punish any shirking. Thus, there exists a subgame-perfect Nash equilibrium in which the intermediary does not shirk, but implements the first-best.

However, as we pointed out in the main text, there also exist subgame-perfect Nash equilibria in which that short-term debt is only partially disciplining. In fact, some investment in the shirking technology may be tolerated by depositors before they run. The reason is that if the intermediary has chosen some $I_{shirk} \geq 0$ that is small enough, "running" is not the dominant strategy.

In order to illustrate this, assume that $R_{shirk} = 0$. Assume the contract (c_1^*, c_2^*) has been promised. We can ask what level of I_{shirk} the intermediary can choose before consumers will run, irrespectively of the behavior of other in the interim period. We have that the following two budget constraints hold:

$$\pi c_1^* = LR/\gamma$$

$$(1-\pi)c_2 = (1 - I_{shirk})R - LR$$

Consumers will run after observing I_{shirk} whenever $c_2 < c_1$. That is, whenever

$$c_2 = \frac{1}{1-\pi}\left((1 - I_{shirk})R - \frac{\pi c_1^* \gamma}{R}\right) < c_1^*$$

$$\Leftrightarrow I_{shirk} > 1 - \frac{c_1^*(1-\pi) + \pi c_1^* \gamma / R}{R} \equiv I_x$$

Therefore, there exist multiple subgame-perfect Nash equilibria in pure strategies which differ in the extent of the discipline that they ensure. In particular, for any $\Psi \in [0, I_x]$, there exist an equilibrium in which each consumer runs if and only if $I_{shirk} > \Psi$, and intermediaries choose $I_{shirk} = \Psi$. In the main part of our analysis, we ignore such equilibria, which is equivalent to assuming that the intermediaries can only invest in either the productive or the shirking technology, i.e., $I_{Shirk} \in \{0, 1\}$.

Appendix 3.C Second-Best

In the presence of equity regulation, the second-best is the solution of the maximization problem of the first-best subject to the equity requirement. If the investors' capital A is not scarce (i.e., if $pL < A$), it also holds that $L < I$, and the following constraints are binding:

$$\begin{aligned}
\pi c_1 &= (1 + e_0 - I) + Lp, \\
(1-\pi)c_2 &= (I - L)R - e_2, \\
e_0 &= \frac{B}{\rho - B}, \\
e_2 &= \rho e_0, \\
p &= R/\gamma, \\
I &= 1 + e_0,
\end{aligned}$$

The maximization problem is reduced to

$$\begin{aligned}
\max_{(c_1,c_2,L)} \quad & \pi u(c_1) + (1-\pi)u(c_2), \\
\text{subject to} \quad & \pi c_1 = LR/\gamma, \\
& (1-\pi)c_2 = (1+e_0-L)R - \rho e_0, \\
& e_0 = \frac{B}{\rho - B}.
\end{aligned}$$

As the utility function is characterized by CRRA with the RA-parameter η, the FOC can be rewritten as

$$c_1 = \gamma^{-\frac{1}{\eta}} c_2,$$

and the period budget equations yield the following consumption levels:

$$\begin{aligned}
c_1 &= \frac{RL}{\pi\gamma}, \quad \text{and} \\
c_2 &= \frac{1}{1-\pi}\left[(1+e_0-L)R - \rho e_0\right],
\end{aligned}$$

and Proposition 3.3 follows.

Appendix 3.D Wholesale Funding Restrictions

In the case that wholesale funding is prohibited, commercial banks can only offer an allocation where liquidity is provided by investing in storage as in Diamond and Dybvig (1983). The optimal allocation is the solution of the following problem:

$$\max_{(c_1,c_2,I)} \quad \pi u(c_1) + (1-\pi)u(c_2), \tag{3.47}$$

$$\text{subject to} \quad \pi c_1 = (1+e_0-I), \tag{3.48}$$

$$(1-\pi)c_2 = IR - \rho e_0, \tag{3.49}$$

$$e_0 = \frac{B}{\rho - B}. \tag{3.50}$$

We see that if we set $\rho = R$ in the second-best problems, the two optimization problems are identical in terms of the consumption allocation they imply, and it holds that

$$L^{**} = (1 + e_0 - I^r), \tag{3.51}$$

i.e., the amount of assets sold in the second-best L^{**} and the investment in storage under wholesale funding restrictions are equal.

Bibliography

Acharya, Viral V., and Raghuram G. Rajan (2013). “Sovereign Debt, Government Myopia, and the Financial Sector”. *Review of Financial Studies* 26.6, pp. 1526–1560.

Acharya, Viral V., Philipp Schnabl, and Gustavo A. Suarez (2013). “Securitization without Risk Transfer”. *Journal of Financial Economics* 107.3, pp. 515–536.

Acharya, Viral V., Itmar Drechsler, and Philipp Schnabl (2014). “A Pyrrhic Victory? Bank Bailouts and Sovereign Credit Risk”. *The Journal of Finance* LXIX.6, pp. 2689–2739.

Admati, Anat R., and Martin F. Hellwig (2013). “Does Debt Discipline Bankers? An Academic Myth about Bank Indebtedness”. *Rock Center for Corporate Governance at Stanford University Working Paper* 132.

Alesina, Alberto, Alessandro Prati, and Guido Tabellini (1990). “Public confidence and debt management: A model and a case study of Italy”. In: *Debt Management and Capital Markets*. Ed. by R. Dornbusch, and M. Draghi. June. Cambridge University Press, pp. 94–118.

Allen, Franklin, and Douglas Gale (1994). “Limited Market Participation and Volatility of Asset Prices”. *American Economic Review* 84.4, pp. 933–955.

——— (1998). “Optimal Financial Crises”. *The Journal of Finance* LIII.4, pp. 1245–1284.

——— (2000). “Financial Contagion”. *Journal of Political Economy* 108.1, pp. 1–33.

Bagehot, Walter (1873). *Lombard Street: A Description of the Money Market*. London, UK: H.S. King.

Barro, Robert J. (1974). “Are Government Bonds Net Wealth?” *Journal of Political Economy* 82.6, p. 1095.

Basel Committee – Basel Committee on Banking Supervision (2010). *Basel III: International Framework for Liquidity Risk Measurement, Standards and Monitoring*. December. Basel: Bank for International Settlement.

Basel Committee – ——— (2011). *Basel III: A global regulatory framework for more resilient banks and banking systems*. June. Basel: Bank for International Settlement.

Basel Committee – ——— (2013). *Basel III: The Liquidity Coverage Ratio and Liquidity Risk Monitoring Tools*. January. Basel: Bank for International Settlement.

Basel Committee – Basel Committee on Banking Supervision (2014). *Basel III: The Net Stable Funding Ratio.* October. Basel: Bank for International Settlement.

Bhattacharya, Sudipto, and Douglas Gale (1987). "Preference Shocks, Liquidity and Central Bank Policy". In: *New Approaches to Monetary Economics, Cambridge University Press,* ed. by W. Barnett, and K. Singleton. Cambridge University Press, pp. 69–89.

Bolton, Patrick, and Olivier Jeanne (2011). "Sovereign Default Risk and Bank Fragility in Financially Integrated Economies". *IMF Economic Review* 59.2, pp. 162–194.

Bolton, Patrick, Tano Santos, and Jose A. Scheinkman (2011). "Outside and Inside Liquidity". *Quarterly Journal of Economics* 126.1, pp. 259–321.

Broner, Fernando, Alberto Martin, and Jaume Ventura (2010). "Sovereign Risk and Secondary Markets". *American Economic Review* 100.4, pp. 1523–1555.

Brunnermeier, Markus K. (2009). "Deciphering the Liquidity and Credit Crunch 2007-2008". *Journal of Economic Perspectives* 23.1, pp. 77–100.

Brunnermeier, Markus K., Arvind Krishnamurthy, and Gary Gorton (2013). "Liquidity Mismatch Measurement". *Working Paper.*

Bryant, John (1980). "A Model of Reserves, Bank Runs, and Deposit Insurance". *Journal of Banking & Finance* 4.4, pp. 335–344.

Bulow, Jeremy, and Kenneth Rogoff (1989). "Sovereign Debt: Is to Forgive to Forget?" *American Economic Review* 79.1, pp. 43–50.

Caballero, Ricardo J. (2010). "Sudden Financial Arrest". *IMF Economic Review* 58, pp. 6–36.

Calomiris, Charles, and Charles Kahn (1991). "The Role of Demandable Debt in Structuring Optimal Banking Arrangements". *American Economic Review* 81.3, pp. 497–513.

Calvo, Guillermo (1988). "Servicing the Public Debt: The Role of Expectations". *American Economic Review* 78.4, pp. 647–661.

Chari, VV, and Ravi Jagannathan (1988). "Banking Panics, Information, and Rational Expectations Equilibrium". *The Journal of Finance* XVIII.3, pp. 749–761.

Cole, Harold L., and Timothy Kehoe (2000). "Self-Fulfilling Debt Crises". *Review of Economic Studies* 67.1, pp. 91–116.

Cooper, Russell, and Kalin Nikolov (2013). "Government Debt and Banking Fragility: The Spreading of Strategic Uncertainty". *NBER Working Paper No. 19278.*

Cooper, Russell, and Thomas W. Ross (1998). "Bank runs: Liquidity Costs and Investment Distortions". *Journal of Monetary Economics* 41.1, pp. 27–38.

Council of the European Union (2013). "COUNCIL REGULATION (EU) No 1024/2013 of 15 October 2013 conferring specific tasks on the European Central Bank concerning policies relating to the prudential supervision of credit institutions". *Official Journal of the European Union* 287.1024, pp. 63–89.

Covitz, Daniel, Nellie Liang, and Gustavo A. Suarez (2013). "The Evolution of a Financial Crisis: Panic in the Asset-Backed Commercial Paper Market". *The Journal of Finance* LXVII.3, pp. 815–847.

Dang, Tri Vi, Gary B. Gorton, and Bengt Holmström (2013a). "Ignorance, Debt and Financial Crises". *Working Paper*.

——— (2013b). "The Information Sensitivity of a Security". *Working Paper*.

Dasgupta, Amil (2004). "Financial Contagion through Capital Connections: A Model of the Origin and Spread of Bank Panics". *Journal of the European Economic Association* 2.6, pp. 1049–1084.

Diamond, Douglas W. (1997). "Liquidity, Banks, and Markets". *Journal of Political Economy* 105.5, pp. 928–956.

Diamond, Douglas W., and Philip H. Dybvig (1983). "Bank Runs, Deposit Insurance, and Liquidity". *Journal of Political Economy* 91.3, pp. 401–419.

Diamond, Douglas W., and Raghuram G. Rajan (2001). "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking". *Journal of Political Economy* 109.2, pp. 287–327.

——— (2005). "Liquidity Shortages and Banking Crises". *The Journal of Finance* XI.2, pp. 615–647.

Eaton, Jonathan, and Mark Gersovitz (1981). "Debt with Potential Repudiation: Theoretical and Empirical Analysis". *Review of Economic Studies* 48.2, pp. 289–309.

European Commission (2013a). *A comprehensive EU response to the financial crisis: substantial progress towards a strong financial framework for Europe and a banking union for the eurozone*. December, MEMO.

——— (2013b). *Establishing uniform rules and a uniform procedure for the resolution of credit institutions and certain investment firms in the framework of a Single Resolution Mechanism and a Single Bank Resolution Fund and amending Regulation (EU) No 1093/2010*. Proposal for a Regulation of the European Parliament and of the Council.

Farhi, Emmanuel, and Jean Tirole (2012). “Collective Moral Hazard, Maturity Mismatch and Systemic Bailouts”. *American Economic Review* 102.1, pp. 60–93.

——— (2014). “Deadly Embrace: Sovereign and Financial Balance Sheets Doom Loops”. *Working Paper*.

Farhi, Emmanuel, Mikhail Golosov, and Aleh Tsyvinski (2009). “A Theory of Liquidity and Regulation of Financial Intermediation”. *Review of Economic Studies* 76.3, pp. 973–992.

Fecht, Falko (2004). “On the Stability of Different Financial Systems”. *Journal of the European Economic Association* 2.6, pp. 969–1014.

FCIC – Financial Crisis Inquiry Commission (2011). *The Financial Crisis Inquiry Report*. Washington D.C.: Printing Office of the US Government.

Flannery, Mark J. (1996). “Financial Crises, Payment System Problems, and Discount Window Lending”. *Journal of Money, Credit and Banking* 28.4, pp. 804–824.

Gale, Douglas (1990). “The Efficient Design of Public Debt”. In: *Public debt management: theory and history*. Ed. by Rudiger Dornbusch, and Mario Draghi. Cambridge University Press.

Gennaioli, Nicola, Andrei Shleifer, and Robert W. Vishny (2013). “A Model of Shadow Banking”. *The Journal of Finance* LXVIII.4, pp. 1331–1363.

Gennaioli, Nicola, Alberto Martin, and Stefano Rossi (2014). “Sovereign Default, Domestic Banks, and Financial Institutions”. *The Journal of Finance* LXIX.2, pp. 819–866.

Goldstein, Itay (2005). “Strategic Complementarities and the Twin Crises”. *Economic Journal* 115, pp. 368–390.

Goldstein, Itay, and Ady Pauzner (2005). “Demand-Deposit Contracts and the Probability of Bank Runs”. *The Journal of Finance* LX.3, pp. 1293–1327.

Goodhart, Charles (2011). *The Basel Committee on Banking Supervision: A History of the Early Years 1974-1997*. Cambridge University Press.

Gorton, Gary B., and G Ordoñez (2013). “The Supply and Demand for Safe Assets”. *NBER Working Paper No. 18732*.

Gorton, Gary B., and George Pennacchi (1990). “Financial Intermediaries and Liquidity Creation”. *The Journal of Finance* XLV.1, pp. 49–71.

Gourinchas, Pierre-Olivier, and Olivier Jeanne (2012). “Global Safe Assets”. *BIS Working Paper No. 399*.

Greenwood, Robin, Samuel G. Hanson, and Jeremy C. Stein (2012). “A Comparative-Advantage Approach to Government Debt Maturity”. *Working Paper*.

Guembel, Alexander, and Oren Sussman (2009). "Sovereign Debt without Default Penalties". *Review of Economic Studies* 76.4, pp. 1297–1320.

Hart, Oliver, and John Moore (1994). "A Theory of Debt Based on the Inalienability of Human Capital". *Quarterly Journal of Economics* 109.4, pp. 841–879.

He, Zhiguo, and Wei Xiong (2012). "Dynamic Debt Runs". *Review of Financial Studies* 25.6, pp. 1799–1843.

Hellwig, Martin F. (1994). "Liquidity Provision, Banking, and the Allocation of Interest Rate Risk". *European Economic Review* 38.7, pp. 1363–1389.

——— (2009). "Systemic Risk in the Financial Sector: An Analysis of the Subprime-Mortgage Financial Crisis". *De Economist* 157.2, pp. 129–207.

——— (2014). "Yes Virginia, There is a European Banking Union! But It May Not Make Your Wishes Come True". *Preprint of the Max Planck Institute for Research on Collective Goods Bonn No. 2014/12.*

Holmström, Bengt, and Jean Tirole (1997). "Financial Intermediation, Loanable Funds, and the Real Sector". *Quarterly Journal of Economics* 112.3, pp. 663–691.

——— (1998). "Private and Public Supply of Liquidity". *Journal of Political Economy* 106.1, pp. 1–40.

——— (2011). *Inside and Outside Liquidity.* Cambridge, MA: The MIT Press.

Jacklin, Charles J. (1987). "Demand Deposits, Trading Restrictions, and Risk Sharing". In: *Contractual Arrangements for Intertemporal Trade.* Ed. by Edward C. Prescott, and Neil Wallace. University of Minnesota Press, pp. 26–47.

Jacklin, Charles J., and Sudipto Bhattacharya (1988). "Distinguishing Panics and Information-based Bank Runs: Welfare and Policy Implications". *Journal of Political Economy* 96.3, pp. 568–592.

Jones, David (2000). "Emerging Problems with the Basel Capital Accord: Regulatory Capital Arbitrage and Related Issues". *Journal of Banking & Finance* 24.1-2, pp. 35–58.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor (2013). "Sovereigns Versus Banks: Credit, Crises, and Consequences". *NBER Working Paper No. 19506.*

Kacperczyk, Marcin, and Philipp Schnabl (2009). "When Safe Proved Risky: Commercial Paper during the Financial Crisis of 2007-2009". *Journal of Economic Perspectives* 24.1, pp. 29–50.

Kashyap, Anil K., Raghuram Rajan, and Jeremy C. Stein (2002). "Banks as Liquidity Providers: An Explanation for the Coexistence of Lending and Deposit taking". *The Journal of Finance* LVII.1, pp. 33–73.

Kiyotaki, Nobuhiro, and John Moore (1997). "Credit Cycles". *Journal of Political Economy* 105.2, pp. 211–248.

König, Philipp, Kartik Anand, and Frank Heinemann (2013). "The 'Celtic Crisis': Guarantees, Transparency, and Systemic Liquidity Risk". *SFB 649 Discussion Paper 2013-025.*

Krishnamurthy, Arvind (2010). "How Debt Markets Have Malfunctioned in the Crisis". *Journal of Economic Perspectives* 24.1, pp. 3–28.

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen (2012). "The Aggregate Demand for Treasury Debt". *Journal of Political Economy* 120.2, pp. 233–267.

——— (2013). "Short-term Debt and Financial Crises: What We Can Learn from US Treasury Supply". *Working Paper.*

Lagos, Ricardo (2006). "Inside and Outside Money Inside and Outside Money". *Federal Reserve Bank of Minneapolis Research Department Staff Report 374* May.

Leonello, Agnese (2013). "Government Guarantees and the Two-Way Feedback between Banking and Sovereign Debt Crises". *Working Paper.*

Longstaff, Francis A. (2004). "The Flight-to-Liquidity Premium in US Treasury Bond Prices". *Journal of Business* 77.3, pp. 511–526.

Luck, Stephan, and Paul Schempp (2014a). "Banks, Shadow Banking, and Fragility". *ECB Working Paper No 1726.*

——— (2014b). "Outside Liquidity, Rollover Risk, and Government Bonds". *Preprint of the Max Planck Institute for Research on Collective Goods Bonn No. 2014/14.*

——— (2014c). "Sovereign Default, Bank Runs, and Contagion". *Preprint of the Max Planck Institute for Research on Collective Goods Bonn No. 2014/15.*

——— (2015). "Regulatory Arbitrage and Systemic Liquidity Crises". *mimeo.*

Martin, Antoine, David Skeie, and Ernst Ludwig Von Thadden (2014a). "Repo Runs". *Review of Financial Studies* 27.4, pp. 957–989.

——— (2014b). "The Fragility of Short-Term Secured Funding Markets". *Journal of Economic Theory* 149, pp. 15–42.

Ordoñez, Guillermo (2013). "Sustainable Shadow Banking". *NBER Working Paper No. 19022.*

Plantin, Guillaume (2014). "Shadow Banking and Bank Capital Regulation". *Review of Financial Studies, forthcoming.*

Postlewaite, Andrew, and Xavier Vives (1987). "Bank Runs as an Equilibrium Phenomenon". *Journal of Political Economy* 95.3, p. 485.

Qi, Jianping (1994). "Bank Liquidity and Stability in an Overlapping Generations Model". *Review of Financial Studies* 7.2, pp. 389–417.

Reinhart, Carmen M., and Kenneth S. Rogoff (2009). "The Aftermath of Financial Crises". *American Economic Review* 99.2, pp. 466–472.

——— (2011). "From Financial Crash to Debt Crisis". *American Economic Review* 101.August, pp. 1676–1706.

Rochet, Jean-Charles, and Xavier Vives (2004). "Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?" *Journal of the European Economic Association* 2.6, pp. 1116–1147.

Saint-Paul, Gilles (2005). "Fiscal Policy and Economic Growth: The Role of Financial Intermediation". *Review of International Economics* 13.3, pp. 612–629.

Schularick, Moritz (2014). "Public and Private Debt: The Historical Record (1870-2010)". *German Economic Review* 15.1, pp. 191–207.

Segura, Anatoli (2014). "Why did Sponsor Banks Rescue their SIVs? A Signaling Model of Rescues". *Working Paper*.

Skeie, David R. (2008). "Banking with Nominal Deposits and Inside Money". *Journal of Financial Intermediation* 17.4, pp. 562–584.

Tirole, Jean. (2010). *The Theory of Corporate Finance*. Princeton University Press.

Uhlig, Harald (2010). "A Model of a Systemic Bank Run". *Journal of Monetary Economics* 57.1, pp. 78–96.

von Thadden, Ernst-Ludwig (1998). "Intermediated versus Direct Investment: Optimal Liquidity Provision and Dynamic Incentive Compatibility". *Journal of Financial Intermediation* 7.2, pp. 177–197.

Woodford, Michael (1990). "Public Debt as Private Liquidity". *American Economic Review* 80.2, pp. 382–388.