EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 1726 / AUGUST 2014

BANKS, SHADOW BANKING, AND FRAGILITY

Stephan Luck and Paul Schempp

In 2014 all ECB publications feature a motif taken from the €20 banknote.

RECIPIENT OF THE YOUNG ECONOMIST PRIZE AT THE FIRST ECB FORUM ON CENTRAL BANKING IN SINTRA

Acknowledgements
We are very thankful for Martin Hellwig's extensive advice and support. We also would like to explicitly thank Jean-Edouard Colliard for very helpful comments. Moreover, we also thank Tobias Adrian, Brian Cooper, Christian Hellwig, Sebastian Pfeil, Jean-Charles Rochet, Eva Schliephake, and Ansgar Walther. Financial support by the Alexander von Humboldt Foundation and the Max Planck Society is gratefully acknowledged.

Stephan Luck
University of Bonn and Max Planck Institute for Research on Collective Goods; e-mail: luck@coll.mpg.de

Paul Schempp
University of Bonn and Max Planck Institute for Research on Collective Goods; e-mail: schempp@coll.mpg.de

Address	Kaiserstrasse 29, 60311 Frankfurt am Main, Germany
Postal address	Postfach 16 03 19, 60066 Frankfurt am Main, Germany
Telephone	+49 69 1344 0
Internet	http://www.ecb.europa.eu

ISSN	1725-2806 (online)
ISBN	978-92-899-1134-4 (online)
EU Catalogue No	QB-AR-14-100-EN-N (online)

Abstract

This paper studies a banking model of maturity transformation in which regulatory arbitrage induces the coexistence of regulated commercial banks and unregulated shadow banks. We derive three main results: First, the relative size of the shadow banking sector determines the stability of the financial system. If the shadow banking sector is small relative to the capacity of secondary markets for shadow banks' assets, shadow banking is stable. In turn, if the sector grows too large, it becomes fragile: an additional equilibrium emerges that is characterized by a panic-based run in the shadow banking sector. Second, if regulated commercial banks themselves operate shadow banks, a larger shadow banking sector is sustainable. However, once the threat of a crisis reappears, a crisis in the shadow banking sector spreads to the commercial banking sector. Third, in the presence of regulatory arbitrage, a safety net for banks may fail to prevent a banking crisis. Moreover, the safety net may be tested and may eventually become costly for the regulator.

JEL: G21, G23, G28
Keywords: shadow banking, regulatory arbitrage, financial crisis, bank runs, maturity transformation

Non-technical summary

Shadow banking activities are banking activities such as credit, maturity, and liquidity transformation that take place outside the regulatory perimeter without having direct access to public sources of liquidity. Shadow banking has expanded rapidly over the last decades and was at the heart of the 2007-2009 financial crisis. This paper contributes to the theoretical understanding of how shadow banking activities can set the stage for a financial crisis.

We discuss a simple banking model of maturity transformation in order to illustrate how sharp contractions in short-term funding become possible in the shadow banking sector, and how such crises may ultimately spread to the commercial banking sector. In our model, commercial banks are covered by a safety net. They are therefore also subject to regulation, which induces regulatory costs for the banks. The shadow banking sector competes with commercial banks by also offering maturity transformation services. In contrast to commercial banks, shadow banking activities are neither covered by the safety net nor subject to regulatory costs.

In the analysis of our theoretical model, we derive three key results. Our first key result is that the relative size of the shadow banking sector may determine its stability. If the shadow banking sector is small relative to arbitrage capital, it appears to be stable. However, if it grows too large, fragility may arise. Fragility in our context is defined as the possibility that panic-based runs may occur. The underlying mechanism is as follows: If the short-term financing of shadow banks breaks down, they are forced to sell their securitized assets on a secondary market. If the size of the shadow banking sector is small relative to the capacity of this secondary market, shadow banks can sell their assets at face value in case of a run. Because they can raise a sufficient amount of liquidity, a run does not constitute an equilibrium. However, if the shadow banking sector is too large, the arbitrageurs' budget does not suffice to buy all assets at face value. Instead, cash-in-the-market pricing leads to depressed fire sale prices in case of a run. Because shadow banks cannot raise a sufficient amount of liquidity, self-fulfilling runs constitute an equilibrium.

The second key finding is that if commercial banks themselves engage in shadow banking activities, a larger shadow banking sector is sustainable. In this case, shadow banking indirectly benefits from the safety net for commercial banks. Banks being covered by the safety net implies that bank depositors never panic and banks thus have additional liquid funds to support their shadow banks. However, if this sustainable level is exceeded, the threat of a crisis may reappear. Moreover, a crisis in the shadow banking sector now also harms the sector of regulated commercial banking.

Finally, the third result is that a safety net for banks may not only be unable to prevent a banking crisis in the presence of regulatory arbitrage, but it may also be costly for the regulator (or taxpayer). If banks and shadow banking are separated, runs can only occur in the shadow banking sector and the regulated commercial banking sector is unaffected. If they are intertwined, a crisis in the shadow banking sector translates into a system-wide crisis, and ultimately the safety net becomes costly for the regulator. Regulatory arbitrage thus undermines the efficacy of the safety net while making it costly for the regulator.

The main result of our theoretical analysis is that the size of the shadow banking sector may play a crucial role for the stability of the financial system. However, the actual quantities of shadow banking activities are not completely clear to academics, policymakers, and regulators. Therefore, we suggest that the size of the shadow banking and the magnitude of maturity mismatch in the shadow banking sector as well as the interconnectedness with the regulated banking sector should be variables that authorities should track closely. More information on the volume of shadow banking activities should be collected and data issues should be tackled.

Moreover, our theoretical analysis suggests that shadow banking that exists due to regulatory arbitrage may pose a severe risk for financial stability. However, we argue that it would be wrong to conclude that regulation should thus be reduced. One needs to keep in mind that – under the presumption that regulation is in place for a good reason – it is not regulation in itself that poses a problem, but the circumvention of regulation. We therefore argue that if regulatory arbitrage can be prevented or reduced, it should be prevented or reduced. Given that regulatory arbitrage can be very costly in terms of creating systemic risk, we argue that it should be made very costly to those who are conducting it.

1. Introduction

A key ingredient to the 2007-2009 financial crisis was the maturity mismatch in the shadow banking sector (see, e.g., Brunnermeier, 2009; Financial Crisis Inquiry Commission, 2011). The shadow banking sector financed long-term real investments via short-term borrowing on a large scale. E.g., asset-backed securities (ABS) were financed through asset-backed commercial papers (ABCP). The increase in delinquency rates of subprime mortgages induced uncertainty about the performance of ABS, leading to the collapse of the market for ABCP, the central short-term financing instrument for off-balance-sheet banking activities (see, e.g., Kacperczyk and Schnabl, 2009; Covitz et al., 2013; Krishnamurthy et al., 2013). The collapse of shadow banking translated into broader financial-sector turmoil in which several commercial banks were on the brink of failure, and the insolvency of Lehman Brothers triggered a run on money market mutual funds (MMFs). Ultimately, governments and central banks had to intervene on a large scale.

This paper contributes to the theoretical understanding of how shadow banking activities can set the stage for a financial crisis. We develop a model in which shadow banking emerges to circumvent financial regulation.[1] We show that, if the shadow banking sector grows too large, fragility arises in the sense that panic-based runs may occur. The size of the shadow banking sector is crucial because it determines the volume of assets being sold on the secondary market in case of a run. We assume that arbitrage capital in this market is limited. Therefore, if the shadow banking sector is too large relative to available arbitrage capital, fire-sale prices are depressed due to cash-in-the-market pricing, and self-fulfilling runs become possible. Moreover, if shadow banking activities are intertwined with activities of commercial banks, a crisis in the shadow banking sector may also trigger a crisis in the regulated banking sector. Eventually, the efficacy of existing safety nets for regulated banks may be undermined.

Shadow banks are financial institutions that operate outside the regulatory perimeter and conduct "credit, maturity, and liquidity transformation without direct and explicit access to public sources of liquidity or credit backstops" (see the definition of Pozsar et al., 2013).[2] Shadow banking activities expanded rapidly over the decades prior to the

[1] There are several other rationales for why shadow banking exists: securitization can be an effective instrument to share macroeconomic interest rate risk (Hellwig, 1994) or to cater to the demand for safe debt (Gennaioli et al., 2013); it can make assets marketable by overcoming adverse selection problems (Gorton and Pennacchi, 1990, 1995; Dang et al., 2013); and it can increase the efficiency of bankruptcy processes (Gorton and Souleles, 2006). In contrast, we focus on the regulatory arbitrage hypothesis which has received considerable support by the empirical findings of Acharya et al. (2013).

[2] On shadow banking, see also Stein (2010), Gorton and Metrick (2010), Claessens et al. (2012), and

crisis (see, e.g., Financial Crisis Inquiry Commission, 2011; Financial Stability Board, 2013; Claessens et al., 2012). In August 2007, however, there was a sharp contraction of short-term funding in the shadow banking sector. The spreads on ABCP rapidly increased after BNP Paribas suspended convertibility of three of its funds that were exposed to risk of subprime mortgages bundled in ABS (Kacperczyk and Schnabl, 2009). The empirical evidence suggests that this contraction resembled the essential features of a run-like event or a rollover freeze (see Gorton and Metrick, 2012; Covitz et al., 2013). In the direct aftermath of the crisis, the academic debate had – due to the availability of data – largely focused on the run on repo (Gorton and Metrick, 2012). However, it is now clear that the market for asset-backed commercial papers has been quantitatively much more important as a source of funding for the shadow banking sector. Moreover, the breakdown of the ABCP market in summer 2007 was quantitatively also more pronounced (Krishnamurthy et al., 2013).[3] Our model is an attempt to illustrate how this sharp contraction in short-term funding such as ABCP became possible and how it ultimately spread to the commercial banking sector.

We discuss a simple banking model of maturity transformation in the tradition of Diamond and Dybvig (1983) and Qi (1994) in order to illustrate how shadow banking can sow the seeds of a financial crisis. In our model, commercial banks' liabilities are covered by a deposit insurance. Because this might induce moral hazard on the part of the banks, they are subject to regulation, which induces regulatory costs for the banks. The shadow banking sector competes with commercial banks in offering maturity transformation services to investors. In contrast to commercial banks, shadow banking activities are neither covered by the safety net nor burdened with regulatory costs.

Our first key result is that the relative size of the shadow banking sector determines its stability. If the short-term financing of shadow banks breaks down, they are forced to sell their securitized assets on a secondary market. The liquidity in this market is limited by the budget of arbitrageurs. If the size of the shadow banking sector is small relative to the capacity of this secondary market, shadow banks can sell their assets at face value in case of a run. Because they can raise a sufficient amount of liquidity in this way, a run does not constitute an equilibrium. However, if the shadow banking

---

Adrian and Ashcraft (2012); on securitization, see, e.g., Gorton and Souleles (2006) and Gorton and Metrick (2013); and on structured finance, see, e.g., Coval et al. (2009).

[3] Gorton and Metrick (2012) focus on repurchase agreements (repos) and hypothesize that there was a run on repo. However, Krishnamurthy et al. (2013) have shown that the run on ABCP was more important (from a quantitative perspective) for the collapse of the shadow banking than the contraction in repo. However, they also emphasize that the contraction in repo selectively affected important investment banks.

sector is too large, the arbitrageurs' budget does not suffice to buy all assets at face value. Instead, cash-in-the-market pricing à la Allen and Gale (1994) leads to depressed fire sale prices in case of a run. Because shadow banks cannot raise a sufficient amount of liquidity, self-fulfilling runs constitute an equilibrium. Depressed fire-sale prices are reminiscent of theories on the limits to arbitrage (see, e.g., Shleifer and Vishny, 1997, 2011) and give rise to multiple equilibria in our model.

As a second key result we find that if commercial banks themselves operate shadow banks, a larger size of the shadow banking sector is sustainable. In this case, the shadow banking sector indirectly benefits from the safety net for commercial banks. Because of this safety net, bank depositors never panic and banks thus have additional liquid funds to support their shadow banks. This enlarges the parameter space for which shadow banking is stable. However, once the threat of a crisis reappears, a crisis in the shadow banking sector also harms the sector of regulated commercial banking.

Finally, the third important result is that a safety net for banks may not only be unable to prevent a banking crisis in the presence of regulatory arbitrage. In fact, it may become tested and costly for the regulator (or taxpayer). If banks and shadow banking are separated, runs only occur in the shadow banking sector, while the regulated commercial banking sector is unaffected. If they are intertwined, a crisis in the shadow banking sector translates into a system-wide crisis and ultimately the safety net becomes tested, and eventually costly, for its provider. This is at odds with the view that safety nets such as a deposit insurance are an effective measure to prevent panic-based banking crises. In traditional banking models of maturity transformation, such as Diamond and Dybvig (1983) and Qi (1994), credible deposit insurance can break the strategic complementarity of investors and eliminate adverse run equilibria at no costs, as it is never tested. The efficacy of such safety nets was widely agreed upon until recently; see, e.g., Gorton (2012) on "creating the quiet period". We show that this may not be the case when regulatory arbitrage is possible. Regulatory arbitrage may undermine the efficacy of safety nets.

The main contribution of our paper is to show how regulatory-arbitrage-induced shadow banking can contribute to the evolution of financial crises. We illustrate how shadow banking activities undermine the effectiveness of a safety net that is installed to prevent self-fulfilling bank runs. Moreover, we show how shadow banking may make the safety net costly for the regulator in case of a crisis. We argue that the understanding of how shadow banking activities contribute to the evolution of systemic risk is not only key to understanding the recent financial crisis. Our results indicate that circumvention of regulation can generally have severe adverse consequences on financial stability. We argue that it is an essential part of any analysis of the efficacy of regulatory interventions to consider the extent of possible regulatory arbitrage.

While the simple nature of our model keeps the analysis tractable, we exclude certain features that might be considered relevant. In our view, the most important ones are the following two: First, in our model, a financial crisis is a purely self-fulfilling phenomenon. We do not claim that the turmoils in summer 2007 were a pure liquidity problem. Clearly, ABCP conduits had severe solvency problems as a consequence of increased delinquency rates. However, our paper is an attempt to demonstrate how the structure of the financial system can set the stage for a severe fragility: because of maturity mismatch in a large shadow banking sector without an explicit safety net, small shocks can lead to large repercussions. Second, by focusing on regulatory arbitrage as the sole reason for the existence of shadow banking, we ignore potential positive welfare effects of shadow banking and securitization. Whenever we speak of shadow banking and its consequences for financial stability, we mainly address shadow banking that originates from regulatory arbitrage. However, the fragility that we find in our model may arguably also exists in a different context.

There is a fast-growing literature on theoretical aspects of shadow banking. Our modeling approach is related to the paper by Martin et al. (2014). However, their focus lies the run on repo and on the differences between bilateral and tri-party repo in determining the stability of single financial institutions. In turn, we focus on ABCP and system-wide crises. The paper by Bolton et al. (2011) is the first contribution to provide an origination and distribution model of banking with multiple equilibria in which adverse selection is contagious over time. Gennaioli et al. (2013) provide a model in which the demand for safe debt drives securitization. In their framework, fragility in the shadow banking sector arises when tail-risk is neglected.

Other contributions that deal with shadow banking are Ordoñez (2013), Goodhart et al. (2012, 2013), and Plantin (2014). Ordonez focuses on potential moral hazard on the part of banks. In his model, shadow banking is potentially welfare-enhancing as it allows to circumvent imperfect regulation. However, it is only stable if shadow banks value their reputation and thus behave diligently; it becomes fragile otherwise. The emphasis of Goodhart et al. lies on incorporating shadow banking into a general equilibrium model. Plantin studies the optimal prudential capital regulation when regulatory arbitrage is possible. In contrast to all three, we focus on the destabilizing effects of shadow banking in the sense that it gives rise to run equilibria.

The paper proceeds as follows: In Section 2, we illustrate the baseline model of maturity transformation. In Section 3, we extend the model by a shadow banking sector and analyze under which conditions fragility arises. In Section 4, we show how the results change when commercial banks themselves operate shadow banks. Finally, we analyze different types of runs in the shadow banking sector in Section 5 and conclude in Section 6.

2. Model Setup

Our baseline model is an overlapping-generation version of the model of maturity transformation by Diamond and Dybvig (1983) which was first introduced by Qi (1994).

There is an economy that goes through an infinite number of time periods $t \in \mathrm{Z}$. There exists a single good that can be used for consumption as well as investment. In each period t, a new generation of investors is born, consisting of a unit mass of agents. Each investor is born with an endowment of one unit of the good, and her lifetime is three periods: $(t, t+1, t+2)$. Upon birth, all investors are identical, but in period $t+1$, their type is privately revealed: With a probability of π, an investor is impatient and her utility is given by $u(c_{t+1})$. With a probability of $1-\pi$, the investor is patient and her utility is given by $u(c_{t+2})$. Assume that the function $u(\cdot)$ is strictly increasing, strictly concave, twice continuously differentiable, and satisfies the following Inada conditions: $u'(0) = \infty$, and $u'(\infty) = 0$.

In each period t, there are two different assets (investment technologies): 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 $t+1$, thus effectively storing the good. The long asset is represented by a continuum of investment projects. An investment project is a metaphor for an agent who is endowed with a project (e.g., an entrepreneur with a production technology or a consumer who desires to finance a house), but has no funds she can invest.

There is no aggregate, but only idiosyncratic return risk: each investment project requires one unit of investment in t and yields a stochastic return of R_i units in $t+2$. 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 some interval $[\underline{R}, \overline{R}] \subset \mathrm{R}^+$, 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 whoever finances the project.

The idiosyncratic return risk of the long asset implies that financial intermediaries dominate a financial markets solution in terms of welfare because of adverse selection in the financial market.[4] In turn, unlike participants of a financial market, a financial intermediary will not be subject to these problems as he is able to diversify and create assets that are not subject to asymmetric information.[5]

Finally, an investment project may be physically liquidated prematurely in $t+1$, yielding a liquidation return of $\ell R_i/R$, where $\ell \in (0, 1/R)$. The liquidation return of a project thus depends on the project's stochastic long-term return. The average liquidation return of a project is equal to ℓ.

Intergenerational Banking

In the following, we describe the mechanics of intergenerational banking and derive steady state equilibria, closely following Qi (1994). We assume that there is a banking sector operating in the economy, consisting of identical infinitely lived banks that take deposits and make investments. It is assumed that the law of large numbers applies at the bank level, i.e., a bank neither faces uncertainty regarding the fraction of impatient investors nor regarding the aggregate return of the long asset.

In each period $t \in \mathrm{Z}$, banks receive new deposits D_t. They sign a demand-deposit contract with investors which specifies a short and a long interest rate. Per unit of deposit, an investor is allowed either to withdraw $r_{t,1}$ units after one period, or $r_{t,2}$ units after two periods. In period t, banks yield the returns from the last period's investment in storage, S_{t-1}, and the returns from investment in the production technology in the second but last period, I_{t-2}. They can use these funds to pay out withdrawing investors and to make new investment in the production and in the storage technology.

We are interested in steady states of this intergenerational banking. A steady state is given by a collection of payoffs, i.e., a short and a long interest rate, (r_1, r_2), a deposit decision D, and an investment decisions I and S. We are only interested in those steady states in which investors deposit all their funds in the banks, $D = 1$, and the total investment in the storage and production technology does not exceed new deposits, i.e., $S + I \leq D$.[6] This yields the investment constraint

$$S + I \leq 1. \tag{1}$$

[4] Because asset quality is not observable, there is only one market price. Impatient consumers with high-return assets have an incentive to liquidate them instead of selling them, and patient consumers with low-return assets have an incentive to sell. This drives the market price below average return and inhibits the implementation of the first-best.

[5] Critiques on the coexistence of financial markets and intermediaries, as by Jacklin (1987) and Thadden (1998), therefore do not apply to our model.

[6] There also exist steady states with $S + I > D$, but this implies that banks have some wealth which is kept constant over time and the net returns of which are paid out to investors each period. This scenario does not appear particularly plausible or interesting.

Moreover, we restrict attention to those steady states in which only impatient consumers withdraw early. We will show later that these withdrawal decisions as well as the deposit decision are actually optimal choices in a steady state equilibrium. In such a steady state, banks have to pay πr_1 units to impatient investors and $(1-\pi)r_2$ units to patient consumers in every period. Since payoffs and investments are limited by returns and new deposits, the following resource constraint must hold:

$$\pi r_1 + (1-\pi)r_2 + S + I \leq RI + S + 1. \tag{2}$$

This constraint can be simplified to obtain a simple feasibility condition for steady-state payoffs:

Definition 1 (Steady-state Payoff)**.** *A steady-state payoff* (r_1, r_2) *is budget feasible if*

$$\pi r_1 + (1-\pi)r_2 \leq (R-1)I + 1. \tag{3}$$

In a next step, we want to select the optimal steady state among the set of budget feasible steady states. Our objective is to choose the steady state that maximizes the welfare of a representative generation of investors, or equivalently, the expected utility of one representative investor. We can partition this analysis by deriving the optimal investment behavior of banks in a first step, and then addressing the optimal interest rates. We see that the budget constraint (3) is not influenced by S. Thus, the banks' optimal investment behavior follows directly:

Lemma 1 (Optimal Investment)**.** *The optimal investment behavior of banks is given by* $I = 1$ *and* $S = 0$*, i.e., there is no investment in storage. The budget constraint reduces to*

$$\pi r_1 + (1-\pi)r_2 \leq R. \tag{4}$$

The intergenerational feature of banking implies that storage is not needed for the optimal provision of liquidity. Any investment in storage would be inefficient and would hence imply a deterioration.

We can now derive the optimal steady-state payoffs (r_1, r_2), i.e., the optimal division between long and short interest rate. It is straightforward to see that the first-best steady-state payoff is given by perfect consumption smoothing, $(r_1^{FB}, r_2^{FB}) = (R, R)$. However, the first-best cannot be implemented as it is not incentive compatible. The incentive-compatibility and participation constraints are given by

$$r_1 \leq r_2, \tag{5}$$

$$r_1^2 \leq r_2, \tag{6}$$

$$\text{and} \quad r_2 \geq R. \tag{7}$$

Constraint (5) ensures that patient investors wait until the last period of their lifetime instead of withdrawing early and storing their funds. Constraint (6) ensures that patient investors do not withdraw early and re-deposit their funds. By this type of re-investment, investors can earn the short interest rate twice. As long as net returns are positive, the latter condition is stronger, implying that the yield curve must not be decreasing. Finally, constraint (7) ensures that investors do not engage in private investment and side-trading. In fact, this condition is the upper bound to the side-trading constraint. The adverse selection problem induced by the idiosyncratic return risk relaxed this constraint, but the constraint will turn out not to be binding anyhow.

Obviously, constraint (6) is violated in the first-best, inducing patient investors to withdraw early and to deposit their funds in the banks a second time. In the second-best, constraints (4) and (6) are binding, resulting in a flat yield curve, $r_2 = r_1^2$. Following Equation (4), the interest rate is such that

$$\pi r_1 + (1 - \pi) r_1^2 = R. \tag{8}$$

Proposition 1 (Qi 1994)**.** *In the second-best steady state, the intergenerational banking sector collects the complete endowment, $D = 1$, and exclusively invests in the long-asset, $I = 1$. In exchange, banks offer demand-deposit contracts with a one-period interest rate given by*

$$r_1^* = \frac{\sqrt{\pi^2 + 4(1 - \pi)R} - \pi}{2(1 - \pi)}, \tag{9}$$

and a two-period interest rate given by

$$r_2^* = r_1^{*2}. \tag{10}$$

It holds that $r_2^* > R > r_1^* > 1$. Unlike in the Diamond and Dybvig model, the first-best and the second-best do not coincide. The intergenerational structure introduces the new IC constraint that the long interest rate must be sufficiently larger than the short one in order to keep patient investors from withdrawal and reinvestment.[7]

Steady-State Equilibrium

Until now, we have not formally specified the game in a game-theoretic sense. Consider the infinite game where in each period $t \in \mathrm{Z}$, investors born in period t decide whether to

[7] However, the intergenerational structure also relaxes the feasibility constraint. Although the yield curve is allowed to be decreasing in the model of Diamond and Dybvig, the second-best of intergenerational banking dominates the first-best of Diamond and Dybvig for a large set of utility functions because banks do not have to rely on inefficient storage.

deposit, and investors born in $t-1$ decide whether to withdraw or to wait for one more period. We do not engage in a full game-theoretic analysis. In particular, we do not characterize all equilibria of this game, but only focus on the equilibrium characterized by the above steady state, and analyze potential deviations. Banks are assumed to behave mechanically according to this steady state.

Lemma 2. *The second-best steady state constitutes an equilibrium of the infinite game.*

If all investors deposit their funds in the banks, and if only impatient consumers withdraw early, it is in fact individually optimal for each investor to do the same. The second-best problem already incorporates the incentive compatibility constraints as well as the participation constraint. Patient investors have no incentive to withdraw early, given that all other patient investors behave in the same way and given that new investors deposit in the bank. Nor do investors have an incentive to invest privately in the production or storage technology, as the bank offers a weakly higher long-run return that R.

Fragility

We will now study the stability of intergenerational banking in the absence of a deposit insurance. Models of maturity transformation such as Diamond and Dybvig (1983) and Qi (1994) may exhibit multiple equilibria in their subgames. Strategic complementarity between the investors may give rise to equilibria in which all investors withdraw early, i.e., bank run equilibria.

In the following, we analyze the subgame starting in period t under the assumption that behavior until date $t-1$ is as in the second-best steady-state equilibrium. We derive the condition under which banks might experience a run by investors, i.e., the condition for the existence of a run equilibrium in the period-t subgame. In the case of intergenerational banking, we consider a “run” in period t to be an event in which all investors born in $t-1$ withdraw their funds, and none of the newly-born investors deposit their endowment. In case of such a run, the bank has to liquidate funds in order to serve withdrawing investors. In addition to the expected withdrawal of impatient consumers, the bank now also has to serve one additional generation of patient investors withdrawing early. Thus, it needs an additional amount of liquid funds $(1-\pi)r_1^* = 1/2\left[\sqrt{\pi^2+4(1-\pi)R}-\pi\right]$. Let us assume that the liquidation rate is sufficiently small relative to the potential liabilities of banks in case of a run:

Assumption 1. $\ell < 1/2\left[\sqrt{\pi^2+4(1-\pi)R}-\pi\right]$.

Assumption 1 implies that, if in some period t all depositors withdraw their funds and newborn investors do not deposit their endowment, the liquidation return that the bank can realize does not suffice to serve all withdrawing consumers. Therefore, the bank is illiquid and insolvent.

Proposition 2. *Assume that the economy is in the second-best steady state. In the subgame starting in period t, a run of investors on banks constitutes an equilibrium.*

This proposition states that the steady state is fragile in the sense that there is scope for a run. Assumption 1 implies that it is optimal for a patient investor to withdraw early if all other patient investors do so and if new investors do not deposit. Note that Proposition 2 only states that a run is an equilibrium of a subgame, but does not say anything about equilibria of the whole game. However, our emphasis lies on the stability/fragility of the steady-state equilibrium.

An important insight from Diamond and Dybvig (1983) and Qi (1994) is that a credible deposit insurance may actually eliminate the adverse equilibrium at no cost. If the insurance is credible, it eliminates the strategic complementarity and is thus never tested. In fact, this is also true in the setup described above. Assume that there is a regulator that can cover the liquidity shortfall in any contingency, including a full-blown bank run. In the context of our model, this amounts to assuming that the regulator has funds of $(1-\pi)r_1^* - \ell$ at its disposal in any period. Whenever patient investors are guaranteed an amount r_1^* by the regulator, they do not have an incentive to withdraw early.[8] In contrast, this does not hold in the presence of regulatory arbitrage, as we will show in the following sections.

3. A Model of Banks and Shadow Banking

We now extend the model described above by three elements: First, we make the assumption that commercial banks are covered by a safety net, but are also subject to regulation and therefore have to bear regulatory costs. Second, there is an unregulated shadow banking sector that competes with banks by also offering maturity transformation services. Investors can choose whether to deposit their funds in a bank or in the shadow banking sector. Depositing in the shadow banking sector is associated with some opportunity cost that varies across investors. Third, there is a secondary market in which securitized assets can be sold to arbitrageurs. The amount of liquidity in this market is assumed to be exogenous.

[8] We ignore the possibility for suspension of convertibility. Diamond and Dybvig (1983) already indicate that suspension of convertibility is critical if there is uncertainty about the fraction of early and late consumers. Moreover, as Qi (1994) shows, suspension of convertibility is also ineffective if withdrawing depositors are paid out by new depositors.

In the following, we describe the extended setup in detail and derive the steady-state equilibrium, before analyzing whether the economy is stable or whether it features multiple equilibria and panic-based runs may occur.

Commercial Banking and Regulatory Costs

From now on, we assume that commercial banks are covered by a safety net that is provided by some unspecified regulator, ruling out runs in the commercial banking sector.[9] Because of this safety net, banks are not disciplined by their depositors, such that – in a richer model – moral hazard could arise. We therefore assume that banks are regulated (e.g., they are subject to a minimum capital requirement). This is assumed to be costly for the bank. In what follows, we will not model the moral hazard explicitly and assume that regulatory costs are exogenous. However, in Appendix A we provide an extension of our model in which we illustrate how moral hazard may arise from the existence of the safety net, and why costly regulation is necessary to prevent moral hazard.

We assume that banks have to pay a regulatory cost γ per unit invested in the long asset, resulting in a gross return of $R - \gamma$. We assume that regulatory costs are not too high, i.e., even after subtracting the regulatory costs, the long asset is still more attractive than storage.

Assumption 2. *$R > 1 + \gamma$.*

Because of the lower gross return, banks can now only offer a per-period interest rate r_b such that

$$\pi r_b + (1 - \pi) r_b^2 = R - \gamma.$$

Under this regulation, the interest rate on bank deposits is explicitly given by

$$r_b = \frac{\sqrt{\pi^2 + 4(1 - \pi)(R - \gamma)} - \pi}{2(1 - \pi)}. \tag{11}$$

The banking sector thus functions like the banking sector in the previous section. The only difference is that banks cannot transfer the gross return R to investors, but only the return net of regulatory cost, $R - \gamma$.

Shadow Banking Sector

We now introduce a shadow banking sector that also offers credit, liquidity, and maturity transformation to investors. The structure of the shadow banking sector (compare Figure 1) is exogenous in our model. We selectively follow and simplify the descriptions by

[9] The regulator is assumed to have sufficient funds to provide a safety net. Moreover, he can commit to actually applying the safety net in case it is necessary, i.e., in case of a run.

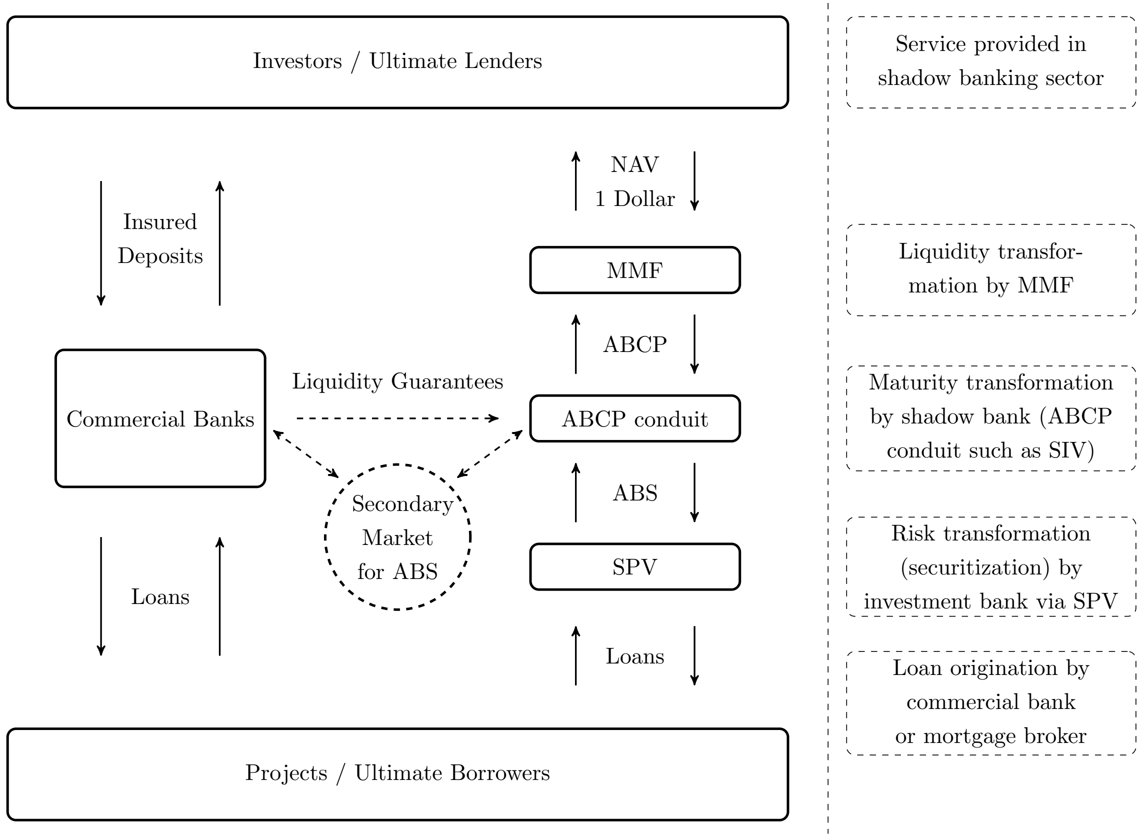

Figure 1: Structure of the financial system: The structure of the shadow banking sector is mostly exogenous in our model. We selectively follow and simplify the descriptions by Poszar et al. (2013). In our setup, shadow banking consists of investment banks, shadow banks (ABCP conduits such as structured investment vehicles (SIVs)), and money market mutual funds (MMFs). Investment banks securitize assets via special purpose vehicle (SPVs) in order to make them tradable, i.e., they conduct risk and liquidity transformation. Once the projects are securitized, they are purchased by shadow banks that finance their long-term assets by borrowing short-term from money market mutual funds (MMFs) via, e.g., ABCP, i.e., they conduct maturity transformation. MMMFs are the door to the shadow banking sector by offering deposit-like claims to investors such as shares with a stable net assets value (NAV), conducting another form of liquidity transformation. Finally, there is a secondary market in which ABS can be sold to arbitrageurs.

Pozsar et al. (2013). Altogether, the actors of the shadow banking system invest in long assets and transform these investments into short-term claims. However, we distinguish between different actors in the shadow banking sector. This structure is exogenous and empirically motivated.

In our setup, shadow banking consists of investment banks, shadow banks, and money market mutual funds (MMFs). Investment banks securitize assets such as loans (i.e., the long assets in our model) via special purpose vehicles (SPVs), thereby transforming them into asset-backed securities (ABS). Through diversified investment, they eliminate the idiosyncratic risk of loans and conduct risk transformation. Note that SPVs typically do not lend to firms or consumers directly, but rather purchase loans from loan originators such as mortgage agencies or commercial banks.

Shadow banks purchase securitized long assets and finance their business by issuing short-term claims that they sell to MMFs. To put it more technically, ABCP conduits such as structured investment vehicles (SIVs) purchase ABS and finance themselves through ABCPs which they sell to MMFs.[10] Shadow banks hence conduct maturity transformation. Maturity transformation is the central element and the key service of banking in our model, and it is the main source of fragility.

For investors, MMFs are the door to the shadow banking sector as they transform short-term debt (such as ABCP) into claims that are essentially equivalent to demand deposits, such as equity shares with a stable net assets value (stable NAV). MMFs thus conduct liquidity transformation. For tractability, we will assume that MMFs are literally taking demand deposits.

Investment banks use their SPVs to invest in a continuum of long assets with idiosyncratic returns R_i. As the law of large numbers is assumed to apply, the return of their portfolio is R. Securitization is assumed to come with a per-unit cost of ρ. Therefore, the per-unit return of securitized loans is $R - \rho$. Investment banks sell these securitized loans to shadow banks. Similar to the regulatory cost γ, we also assume that the securitization cost is not too high, i.e., even after subtracting the securitization cost, the long asset is still more attractive than storage:

Assumption 3. $R > 1 + \rho$.

The empirically motivated narrative is that investment banks purchase loans from loan originators such as mortgage brokers or commercial banks. They bundle the claims into securitized loans (ABS) via SPVs, successfully diversifying the idiosyncratic return risk. Securitization costs accrue and can be thought of as the costs of creating an ABS

[10]Note that we ignore other securities that shadow banks also use to finance their activities, such as medium term notes (MTNs).

out of many small loans, e.g., the costs of hiring a rating agency and a lawyer and setting up the information technology to process payments.[11] Securitization ultimately makes the long assets tradable by eliminating the adverse selection problem that is associated with idiosyncratic return risk. Ultimately, securitized loans (ABS) are sold to shadow banks.

At the heart of the shadow banking sector is the maturity transformation by shadow banks (ABCP conduits). Shadow banks purchase securitized assets (ABS) from investment banks' SPVs. As described above, these assets have a return of $R - \rho$ and a maturity of two periods. Shadow banks can finance themselves by borrowing from MMFs via ABCPs. Moreover, they can also sell ABS to arbitrageurs in the secondary market which is specified below.

Shadow banks offer a per-period interest rate r_{abcp} such that

$$\pi r_{abcp} + (1 - \pi) r_{abcp}^2 = R - \rho,$$

implying a return of

$$r_{abcp} = \frac{\sqrt{\pi^2 + 4(1 - \pi)(R - \rho)} - \pi}{2(1 - \pi)}.$$

We assume that there exists a secondary market for securitized assets (ABS). There are arbitrageurs who are willing to buy ABS at a price that equals expected revenue. Arbitrageurs can be thought of as experts (pension funds, hedge funds) that do not necessarily hold ABS in normal times, but purchase them if they are available at small discounts and promise gains from arbitrage.

The secondary market is assumed to be such that there is no market power on any side of the market. Moreover, there is a fixed amount of cash in this market. We assume that arbitrageurs have a total budget of A which they are willing to spend for buying ABS which can lead to cash-in-the-market pricing. The equilibrium supply and price of ABS on the secondary market will be derived below.

The idea behind this assumption is that not every individual or institutions has the expertise to purchase financial products such as ABS. Moreover, the equity and collateral of these arbitrageurs is limited, so they cannot borrow and invest infinite amounts.[12]

Investors can access the services of the shadow banking sector vian MMFs, which assumed to intermediate between investors and shadow banks.[13] MMFs offer demand-

[11]Securitization costs could also be understood as the regulatory costs that accrue in shadow banking. While shadow banking activities are outside the regulatory perimeter of banking regulation, there are nonetheless existing regulations.

[12]See theories on the limits to arbitrage Shleifer and Vishny (1997).

[13]This is like assuming that investors face large transaction costs or do not have the expertise to deal with shadow banks directly.

deposit contracts to investors while purchasing short-term claims on shadow banks.[14] MMFs offer a per-period interest rate r_{mmf} to investors and purchase ABCP (short-term debt) with a per-period return r_{abcp} from shadow banks. Competition among MMFs implies that $r_{mmf} = r_{abcp}$.

Upon birth, investors can choose whether to deposit their endowment in a regulated bank or in an MMF. Depositing at MMFs comes at some opportunity cost. We assume that investors are initially located at a regulated bank. Switching to an MMF comes at a cost of s_i, 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 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. One can think of these costs as monitoring or screening costs for investors that become necessary when choosing an MMF as these are not protected by a deposit insurance (see Appendix A for more details). For simplicity, we have assumed that all depositors have the same size. However, we could alternatively write down a model where investors have different endowments (see Appendix B). It is very plausible that the ratio of switching costs to the endowment is lower for larger investors (e.g., for corporations that need to store liquid funds of several millions for a few days). Another interpretation is the forgone service benefits that depositors lose when leaving commercial banks, such as payment services and ATMs.

Investors' Behavior

Given the interest rates of commercial banks, r_b, of MMFs, r_{abcp}, and given the switching cost distribution G, we can pin down the size of the shadow banking sector.

Lemma 3. *Assume that banks offer an interest rate r_b and MMFs offer an interest rate of r_{abcp}, as specified above. Then there exists a unique threshold s^* such that an investor switches to an MMF if and only if $s_i \leq s^*$. The mass of investors depositing in the shadow banking sector is given by $G(s^*)$. It holds that $s^* = f(\gamma, \rho)$, where $f_\gamma > 0$ and $f_\rho < 0$.*

Proof. Take r_b and r_{abcp} as described above. We know r_b decreases in γ, and r_{abcp} decreases ρ. Staying at a commercial bank provides an investor with an expected consumption utility of $EU_b = \pi u(r_b) + (1-\pi)u(r_b^2)$. Switching to an MMF is associated

[14] an MMF typically sells shares to investors, and the fund's sponsor guarantees a stable NAV, i.e., it guarantees to buy back shares at a price of one at any time. As mentioned above, the stable NAV implies that an MMF share is a claim that is equivalent to a demand-deposit contract. For simplicity, we will assume that MMFs are literally taking demand deposits.

with an expected consumption utility of $EU_{sb} = \pi u(r_{abcp}) + (1-\pi)u(r^2_{abcp})$. Observe that EU_b decreases in γ and EU_{sb} decreases in ρ.

An investor with switching cost s_i switches to the shadow banking sector if $EU_b < EU_{sb} - s_i$. This implies that all investors with $s_i \leq EU_{sb} - EU_b$ switch to MMFs. We define $s^* \equiv f(\gamma,\rho) = EU_{sb}(\rho) - EU_b(\gamma)$. A mass $G(s^*)$ of each generation's investors switches to MMFs, and a mass $1-G(s^*)$ stays at commercial banks. Because u is twice continuously differentiable, it holds that $\partial EU_b/\partial\gamma < 0$ and $\partial EU_{sb}/\partial\rho > 0$. Thus, f is a continuously differentiable function with $f_\gamma > 0$ and $f_\rho < 0$. □

An investor with $s_i = s^*$ is indifferent between depositing at a bank or an MMF. All investors with lower switching costs choose an MMF; their mass is given by $G(s^*)$. The size of the shadow banking sector increases in the regulatory cost γ and decreases in the cost of securitization ρ. For example, if investors had linear consumption utility, it would hold that $s^* = \gamma - \rho$.

We are now equipped to characterize the economy's steady state equilibrium:

Proposition 3. *In the second-best steady-state equilibrium, the intergenerational banking sector collects an amount of deposits $D_b = 1 - G(s^*)$ in each period, and invests all funds in the long-asset, $I_b = 1 - G(s^*)$. They offer demand-deposit contracts with a per-period interest rate of*

$$r_b = \frac{\sqrt{\pi^2 + 4(1-\pi)(R-\gamma)} - \pi}{2(1-\pi)}. \tag{12}$$

MMFs collect an amount of deposits $D_{sb} = G(s^*)$, offering demand-deposit contracts with a per-period interest rate of $r_{mmf} = r_{abcp}$. MMFs lend all funds to shadow banks which exclusively invest in ABS, $I_{sb} = 1 - G(s^*)$. They offer a per-period interest of

$$r_{abcp} = \frac{\sqrt{\pi^2 + 4(1-\pi)(R-\rho)} - \pi}{2(1-\pi)}. \tag{13}$$

It holds that $s^* = f(\gamma,\rho)$, where $f_\gamma > 0$ and $f_\rho < 0$. There are no assets traded in the secondary market.

Proposition 3 described the steady state in which regulated commercial banks and shadow banking coexist. The interest rates are given by r_b and r_{abcp} and depend on γ and ρ, which determines the size of the shadow banking sector as described by Lemma 3. It is important to notice that, in this steady-state equilibrium, no assets are being sold to arbitrageurs on the secondary market, as there are no gains from trade.

Fragility

As in the previous section, we now study the stability of shadow banks. We analyze the subgame starting in period t under the assumption that behavior until date $t-1$ is as in the steady-state equilibrium specified in Proposition 3. We derive the condition under which shadow banks might experience a run by MMFs, i.e., the condition for the existence of a run equilibrium in the period-t subgame.

Because short-term liabilities in the shadow banking sector are not insured, a run in the shadow banking sector is not excluded per se. However, as will become clear below, runs are only possible if the shadow banking sector is too large. Generally, there are two types of runs that could potentially take place in the adverse equilibrium of the $t = 1$ subgame. First, all investors withdraw their funds from the MMFs, and no new funds are deposited. Second, MMFs withdraw all their funds from shadow banks, i.e., they stop rolling over ABCPs. Throughout this section, we will assume that MMFs have sponsors that are able credibly to guarantee the stable NAV of MMFs in any contingency. That is, an MMF sponsor credibly guarantees stable NAV for the MMF shares, i.e., it buys back shares at face value in case of liquidity problems. This allows us to focus on the second case where MMFs run on shadow banks. In a later section, we will analyze under which conditions runs by investors on MMFs can accompany runs of MMFs on shadow banks.

Whether a run of MMFs on shadow banks constitutes an equilibrium depends on whether shadow banks can raise enough liquidity in the secondary market to serve all their obligations. A run of MMFs implies that MMFs withdraw all their funds from shadow banks and deposit no new funds (i.e., they stop rolling over ABCP). In case of such a run in period t, shadow banks have to repay what the MMFs have invested on behalf of the mass of $(1-\pi)$ patient investors in $t-2$ who have claims worth r^2_{abcp}. Moreover, they have to pay all funds that were invested on behalf of those investors from $t-1$ who have claims worth r_{abcp}. Given that only a fraction of $G(s^*)$ investors deposit their funds in the shadow banking sector each period, shadow banks have to serve MMFs with a total amount of $G(s^*)[(1-\pi)r^2_{abcp} + r_{abcp}]$.

However, shadow banks only have an amount $G(s^*)(R-\rho)$ of liquid funds available in t from the investment in ABS they made in $t-2$. The liquidity shortfall of shadow banks in case of a run by MMFs is given by

$$G(s^*)[(1-\pi)r^2_{abcp} + r_{abcp} - (R-\rho)].$$

In order to cover this shortfall, shadow banks can either sell the ABS that they bought in $t-1$ to the arbitrageurs, or they can liquidate these assets.[15] We assume that

[15] Liquidating ABS might not be straightforward, as all tranches would have to be collected and the un-

liquidation of ABS will never be enough to cover the shortfall. Similar to Assumption 1, this is equivalent to making the following assumption:

Assumption 4. $\ell < 1/2\left[\sqrt{\pi^2 + 4(1-\pi)(R-\rho)} - \pi\right]$

Observe that in case of a run, the supply of shadow banks is partially inelastic: they have to cover their complete liquidity shortfall. There are two cases to be considered: In the first case, the arbitrageurs' funds are sufficient to purchase all funds the shadow banks sell at face value, while in the second case, the arbitrageurs' budget is not sufficient and the price is determined by cash-in-the-market pricing. Runs of MMFs on shadow banks become possible in this second case.

Proposition 4. *Assume that the economy is in the second-best steady state. A run of MMFs on shadow banks (ABCP conduits) constitutes an equilibrium of the period-t subgame if and only if*

$$G(s^*) > \frac{A}{1/2\left[\sqrt{\pi^2 + 4(1-\pi)(R-\rho)} - \pi\right]} \equiv \xi,$$

where $s^* = f(\gamma, \rho)$*, with* $f_\gamma > 0$ *and* $f_\rho < 0$.

Proof of Proposition 4. Assume that MMFs collectively withdraw funds from shadow banks and deposit no new funds in period t. It will be optimal for a single MMF to also withdraw if the shadow banks become illiquid and insolvent in t.

We calculate the liquidity shortfall if shadow banks in case of a run as

$$G(s^*)[(1-\pi)r_{abcp}^2 + r_{abcp} - (R-\rho)].$$

Recall from Proposition 3 that $\pi r_{abcp} + (1-\pi)r_{abcp}^2 = R - \rho$. Making use of this by substituting for $(R-\rho)$, we know that the shortfall is given by $G(s^*)[(1-\pi)r_{abcp}]$. Recalling Equation (13), the liquidity shortfall is given by

$$1/2\left[\sqrt{\pi^2 + 4(1-\pi)(R-\rho)} - \pi\right] G(s^*).$$

Liquidating the ABS portfolio would yield $\ell G(s^*)$. According to Assumption 4, this will always be less than the liquidity shortfall. We therefore know that shadow banks will never be able to cover the liquidity shortfall by liquidating their ABS portfolio.

The relevant question is whether shadow banks can raise sufficient funds by selling the ABS portfolio to arbitrageurs. There are two cases to be considered: in the first case,

derlying assets would have to be liquidated. However, our model also goes through in case liquidation is not possible; we can just set $\ell = 0$.

the arbitrageurs' funds are sufficient to purchase all funds the shadow banks sell at face value, while in the second case the arbitrageurs' budget is not sufficent and the price is determined by cash-in-the-market pricing.

We are in the first case if $A \geq 1/2\left[\sqrt{\pi^2+4(1-\pi)(R-\rho)}-\pi\right]G(s^*)$. All ABS held by shadow banks can be sold at face value, i.e., the price of ABS in the secondary market is equal to the expected return, $p = R - \rho$. The value of shadow banks' ABS as well as the amount of cash in the market exceeds the shadow banks' potential liquidity needs. Therefore, in case of a run, all old MMFs can be served. Therefore, it is weakly dominant strategy for each MMF to rollover and to deposit new funds. A run therefore does not constitute an equilibrium.

The second case is given by $A < 1/2\left[\sqrt{\pi^2+4(1-\pi)(R-\rho)}-\pi\right]G(s^*)$, where shadow banks cannot sell all their assets at face value. If all MMFs stop rolling over ABCP, shadow banks cannot raise the required funds to fulfill their obligations by selling their ABS because the amount of assets on the secondary market exceeds the budget of arbitrageurs. The price of ABS drops below face value and shadow banks are forced to sell their complete ABS portfolio. Still, shadow banks can only raise a total amount A of liquidity, which is insufficient to serve withdrawing MMFs. It follows that it is not optimal for an MMF to roll over ABCP if no other MMF does so. A self-fulfilling run thus constitutes an equilibrium whenever

$$G(s^*) > \frac{A}{1/2\left[\sqrt{\pi^2+4(1-\pi)(R-\rho)}-\pi\right]} \equiv \xi.$$

□

The key mechanism giving rise to multiple equilibria is cash-in-the-market pricing (see, e.g., Allen and Gale, 1994) in the secondary market for long-term securities that results from limited arbitrage capital and is related to the notion of limits to arbitrage (see, e.g., Shleifer and Vishny (1997)). The fact that there are not enough arbitrageurs (and that these arbitrageurs cannot raise enough funds) to purchase all assets of the shadow banking system possibly induces the price of ABS to fall short of their face value. This implies that shadow banks may in fact be unable to serve their obligations once they sell all their long-term securities prematurely. This in turn makes it optimal for an MMF to run on shadow banks once all other MMFs run.

In order to illustrate the role of limited availability of arbitrage capital we examine the hypothetical fire-sale price in the market for ABS. Cash-in-the-market pricing describes a situation where the buyers' budget constraint is binding and the supply is fixed. The price adjusts such that demand balances the fixed supply. In our case, the price p is

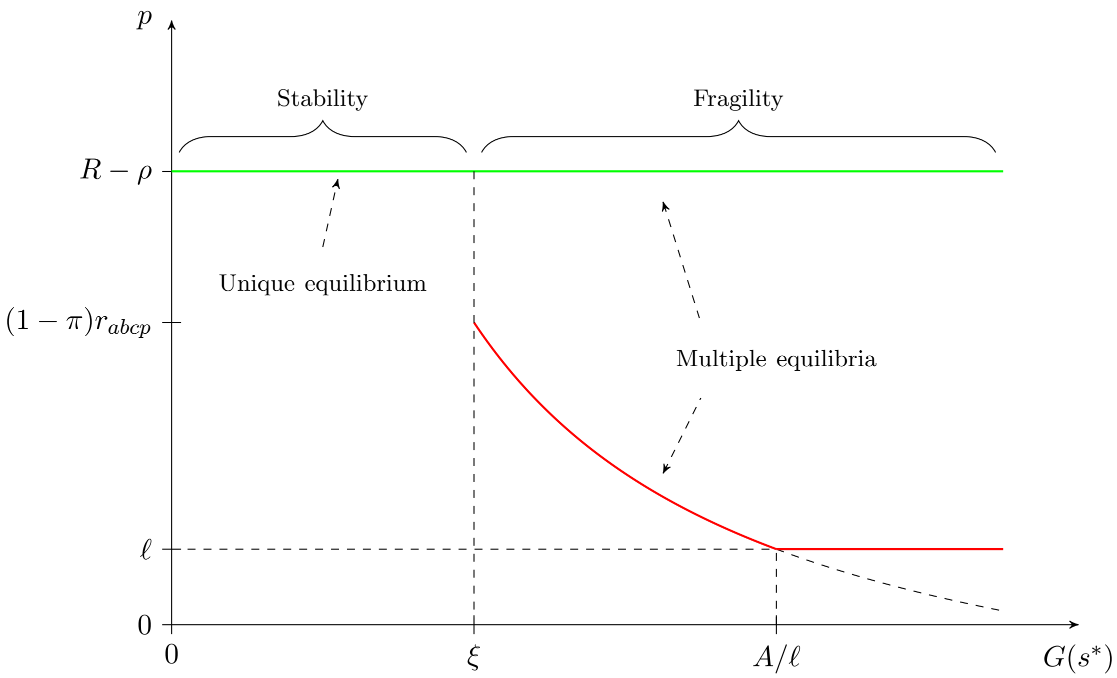

Figure 2: This graph depicts the potential fire-sale price of ABS. Whenever $G(s^*) \leq \xi = A/(1-\pi)r_{abcp}$, the funds of arbitrageurs are sufficient to purchase all assets of shadow banks at face value. There is a unique equilibrium of the period t subgame in which there are no panic-based withdrawals of MMFs. In turn, if $G(s^*) > \xi$, the funds of arbitrageurs are insufficient, and the period t subgame has multiple equilibria. If all MMFs withdraw from shadow banks, the price of ABS in the secondary market drops to the red line.

such that

$$pG(s^*) = A.$$

The fire-sale price is a function of the amount of assets that are on the market in case of a run on the shadow banking sector, which is given by the size of the shadow banking sector $G(s^*)$. The price is given by

$$p(s^*) = \begin{cases} R - \rho & \text{if } G(s^*) \leq \xi, \\ A/G(s^*) & \text{if } G(s^*) \in (\xi, A/\ell], \\ \ell & \text{if } G(s^*) > A/\ell. \end{cases}$$

The equilibrium fire-sale prices as a function of the size of the shadow banking sector is illustrated in Figure 2.

Whether the period t subgame has multiple equilibria ultimately depends on the parameters ρ and γ, as they determine the size of the shadow banking sector. This is depicted in Figure 3. Whenever the regulatory costs γ exceed the costs of securitizing assets ρ (i.e., if we are above the 45 degree line), the shadow banking sector has positive size in equilibrium, i.e. $G(s^*) > 0$. However, as long as the shadow banking sector is small relative to the capacitiy of arbitrageurs to purchase its assets at face value in a fire-sale, it is stable. Only when regulatory costs γ are sufficiently larger than securitization cost ρ, the shadow banking sector's size $G(s^*)$ exceeds the critical threshold ξ, and shadow banking becomes fragile.

4. Liquidity Guarantees

So far, there has been no connection between the regulated commercial banking sector and the shadow banking sector; both sector compete for the investors' funds. We now assume that commercial banks themselves actively engage in shadow banking, i.e., they operate shadow banks through off-balance-sheet subsidiaries (ABCP conduits).[16] In fact, we assume that commercial banks explicitly or implicitly provide their ABCP conduits with liquidity guarantees. They may have strong incentives to support their conduits in case of distress, e.g., in order to protect their reputation, see Segura (2014).

As above, we assume that the commercial banks' demand-deposit liabilities are covered by a credible safety net. This safety net being credible implies that commercial banks

[16] Note that the fact that banks operate shadow banks themselves does not result from optimal behavior in our setup. However, our idea is that it is profitable for banks to found their own shadow banks.

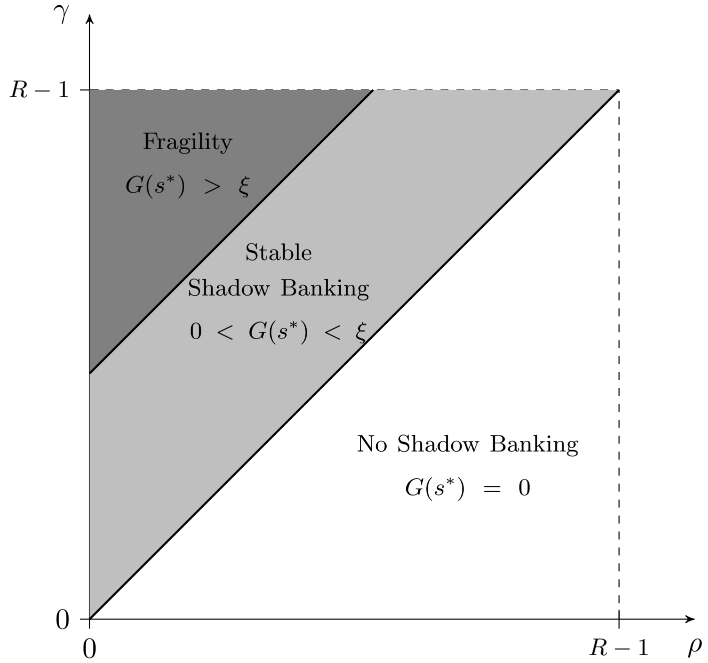

Figure 3: This figure visualizes the equilibrium characteristics of the financial system for different values of γ and ρ. For $\gamma < \rho$, shadow banking is not made use of in equilibrium, as it is dominated by commercial banking. If $\gamma > \rho$, the shadow banking sector has positive size. As long as the difference $\gamma - \rho$ is small, shadow banking is stable. If the difference increases, the size of the shadow banking sector also increases and finally introduces fragility into the financial system.

do not experience runs by investors. Patient investors who are located at a commercial bank will thus never withdraw their funds early.

Liquidity guarantees imply that in case of a run on shadow banks, commercial banks supply liquid funds to shadow banks. This increases the critical size up to which the shadow banking sector is stable. However, this comes with an unfavorable side effect: once this critical size is exceeded an shadow banks experience a run, the crisis spreads to the commercial banking sector and makes the safety net costly.

Proposition 5. *Assume that the economy is in the second-best steady state described in Proposition 3 and all shadow banks (ABCP conduits) are granted liquidity guarantees by commercial banks. A run of MMFs on shadow banks constitutes an equilibrium of the* *subgame starting in period t if and only if*

$$G(s^*) > \frac{\max[A,\ell]+1}{1/2\left[\sqrt{\pi^2+4(1-\pi)(R-\rho)}-\pi\right]+1} \equiv \vartheta,$$

where $s^* = f(\gamma,\rho)$. *It holds that* $\vartheta > \xi$.

Proof. In case of a run, the shadow banks' need for liquidity is given as above by

$$1/2\left[\sqrt{\pi^2+4(1-\pi)(R-\rho)}-\pi\right]G(s^*).$$

Banks can sell their loans on the same secondary market in case of a crisis. Still, the total endowment of arbitrageurs in this market is given by A. Therefore, either banks and shadow banks sell their assets in the secondary market, or both types of institutions liquidate their assets. They jointly still only raise an amount A from selling long-term securities on the secondary market or ℓ units from liquidating all long assets. The maximum amount they can raise is thus $\max[A,\ell]$. On top, commercial banks also have an additionally amount $1-G(s^*)$ of liquid funds available since new investors still deposit their endowment at commercial banks because of the safety net for commercial banks.

The liquidity guarantees by commercial banks can satisfy the shadow banks' liquidity needs in case of a run if

$$\max[A,\ell]+(1-G(s^*)) \geq 1/2\left[\sqrt{\pi^2+4(1-\pi)(R-\rho)}-\pi\right]G(s^*),$$

which is equivalent to

$$G(s^*) \leq \frac{\max[A,\ell]+1}{1/2\left[\sqrt{\pi^2+4(1-\pi)(R-\rho)}-\pi\right]+1} = \vartheta.$$

If $G(s^*) \leq \vartheta$, the liquidity guarantees suffice to satisfy the liquidity needs in case of a run, so a run does not constitute an equilibrium. If $G(s^*) > \vartheta$, the liquidity guarantees do suffice to satisfy the liquidity needs in case of a run, and a run equilibrium. □

If commercial banks themselves operate shadow banks and provide them with liquidity guarantees, the parameter space in which shadow banking is stable is enlarged compared to a situation without liquidity guarantees, i.e., the critical threshold for the size of the shadow banking sector ϑ is now larger than ξ, the threshold in the absence of liquidity guarantees. This shift is also depicted in Figure 4. The reason for this result is that banks have additional liquid funds, even in case of a crisis: because of the deposit insurance, they always receive funds from new depositors, and their patient depositors never have an incentive to withdraw early.

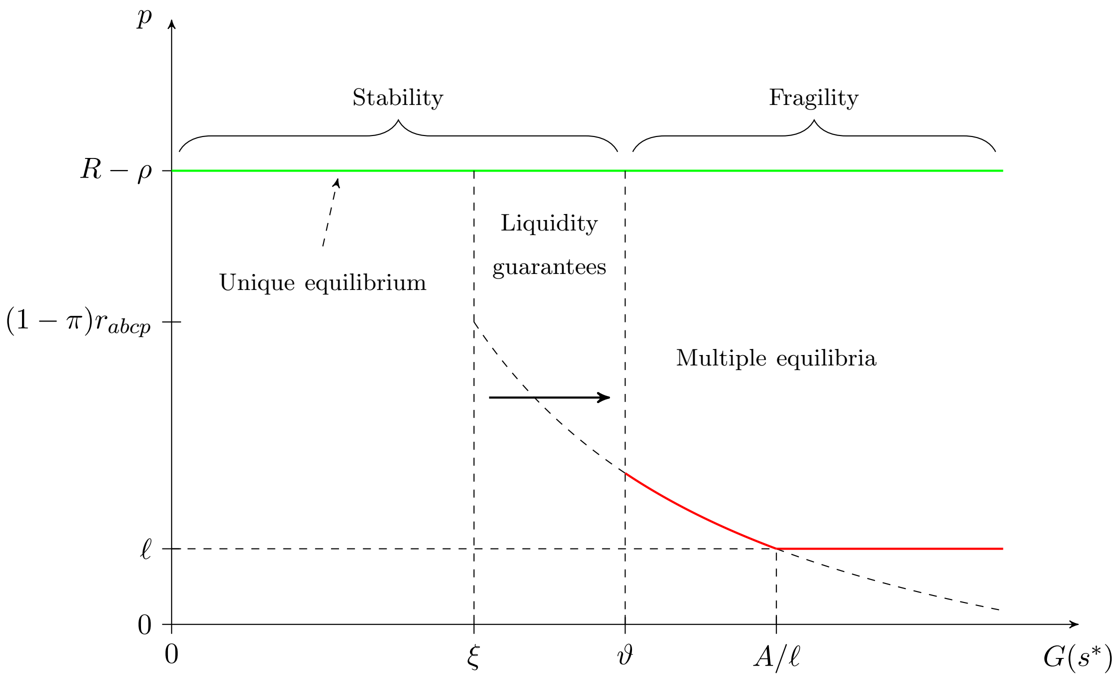

Figure 4: This graph depicts the potential fire-sale price of ABS for the case that regulated commercial banks provide liquidity guarantees to shadow banks. The critical size above which multiple equilibria exist moves from ξ to ϑ.

In traditional banking models, policy tools like a deposit insurance eliminate self-fulfilling adverse equilibria at no cost. This is not necessarily true in our model: once the shadow banking sector exceeds the size ϑ, a run in the shadow banking sector constitutes an equilibrium despite the safety net for commercial banks, and despite the liquidity guarantees of banks. Shadow banks – by circumventing the existing regulation – place themselves outside the safety net and are thus prone to runs. If the regulated commercial banks offer liquidity guarantees, a crisis in the shadow banking sector also spreads to the regulated banking sector. Ultimately, self-fulfilling adverse equilibria are not necessarily eliminated by the safety net and may become costly.

Corollary 1. *Assume that $G(s^*) > \vartheta$ and assume banks provide liquidity guarantees to shadow banks. In case of a run in the shadow banking sector, the safety net for regulated commercial banks is tested and the regulator must inject an amount*

$$1/2\left[\sqrt{\pi^2 + 4(1-\pi)(R-\rho)} - \pi\right] G(s^*) - \max[A, \ell] - (1 - G(s^*)) > 0.$$

If the regulated commercial banking and the shadow banking sector are intertwined, a crisis may not be limited to the shadow banking sector, but also spread to the commercial banks, thus testing the safety net. Ultimately, the regulator has to step in and cover the commercial banks' liabilities. Therefore, the model challenges the view that policy measures like a deposit insurance necessarily are an efficient mechanism for preventing self-fulfilling crises. Historically, safety nets such as a deposit insurance schemes were perceived as an effective measure to prevent panic-based banking crises. The view is supported by traditional banking models of maturity transformation such as Diamond and Dybvig (1983) and Qi (1994). In the classic models of self-fulfilling bank runs, a credible deposit insurance can break the strategic complementarity in the withdrawal decision of bank customers at no cost. We show that this may not be the case when regulatory arbitrage is possible and regulated and unregulated banking activities are intertwined.

5. Runs on MMFs

In the previous sections, we ruled out runs on MMFs by assuming that they have credible support by a sponsor. Credible sponsor support means that even if all investors withdraw their funds from an MMF, the sponsor is able to provide sufficient liquidity to the MMF such that it can serve all investors. Recall that we use the narrative that MMFs are literally offering demand-deposit contracts. In practice, an MMF issues equity shares, and its sponsor guarantees stable NAV for theses shares, i.e., it promises to buy these shares at face value in case of liquidity problems.

We now relax the assumption that the guarantee is always credible. We explicitly model the credibility of the guarantee by assuming that the sponsors have m units of liquidity per unit of investment in the MMF that they can provide in case of a crisis. Moreover, we keep the assumption of existing liquidity guarantees. We show that providing $mG(s^*)$ units only credibly prevents a run on MMFs if this amount is sufficient to fill the liquidity shortfall in case of a run of investors on MMFs, which in turn triggers a run of MMFs on shadow banks.

Proposition 6. *Assume that the economy is in the second-best steady state equilibrium described in Proposition 3. Assume further that all shadow banks (ABCP conduits) are granted liquidity guarantees by commercial banks, and that per unit of investment, MMFs receive m units of liquidity support from their sponsor. A run of investors on MMFs may occur whenever*

$$G(s^*) > \frac{\max[A, \ell] + 1}{1/2\left[\sqrt{\pi^2 + 4(1-\pi)R} - \pi\right] + 1 - m} = \nu > \vartheta.$$

If the $\nu > G(s^*) > \vartheta$, investors never run on MMFs. However, MMFs might run on shadow banks, which then draw on the sponsor support.

Proof. Observe that once an MMF needs liquid funds because investors withdraw unexpectedly, it will stop rolling over ABCP. Now, whenever the shadow banking sector exceeds the critical threshold ϑ, a run of MMFs on shadow banks is self-fulfilling as shadow banks will make losses only in this case. This therefore is a necessary condition for a run by investors on MMFs. If it is not satisfied, MMFs are always able to fulfill their obligations by stopping the rollover of ABCP, making it a weakly dominant strategy for patient investors not to withdraw early. However, it is not a sufficient condition.

Observe that the resulting liquidity shortfall for the MMFs is given by

$$\left[1/2\left[\sqrt{\pi^2 + 4(1-\pi)R} - \pi\right] + 1\right] G(s^*) - \max[A, \ell] - 1.$$

Therefore, a run of investors on MMFs constitutes an equilibrium only if

$$mG(s^*) < \left[(1/2\left[\sqrt{\pi^2 + 4(1-\pi)R} - \pi\right] + 1\right] G(s^*) - \max[A, \ell] - 1.$$

□

The result builds on the fact that sponsor support is like a liquidity backstop. If there is a run by MMFs on shadow banks, MMFs will make losses. This additionally triggers a run of investors on MMFs if the sponsor is not able to cover these losses. Again, losses depend on the fire-sale price. The fire-sale price in turn depends on the amount of assets sold in case of a run by MMFs on shadow banks, which is determined by the size of the shadow banking sector. If the shadow banking sector is so large that runs by MMFs on shadow banks occur, but not so large that losses cannot be covered by the sponsors, investors do not run. This is the case for $\nu > G(s^*) > \vartheta$. In turn, if the shadow banking sector size exceeds ν, a run by MMFs on shadow banks will always be accompanied by a run of investors on MMFs because sponsor support is insufficient to cover losses in case of a run.

6. Discussion

The main contribution of this paper is to show how regulatory arbitrage-induced shadow banking can sow the seeds of a financial crisis. We illustrate how shadow banking activities undermine the effectiveness of a safety net that is installed to prevent a financial crisis. Moreover, we show how regulatory arbitrage may even induce the safety net to be costly for the regulator (or taxpayer) in case of a crisis.

Our model features multiple equilibria. The key mechanism giving rise to multiple equilibria is cash-in-the-market pricing in the secondary market for shadow banks' long-term securities which results from limited availability of arbitrage capital. Cash-in-the-market pricing leads to depressed fire-sale prices if there are too many assets on the market. The amount of assets is thus crucial in determining whether shadow banking is fragile or not. In turn, the amount of assets sold in case of a run on shadow banks is determined by the size of the shadow banking sector. Therefore, multiple equilibria only exist if the shadow banking sector is large.

As indicated earlier, our model lacks certain features that might be considered relevant that should be considered in future research. First, a financial crisis is a purely self-fulfilling phenomena in our model, while fundamental values do play a role in reality. However, our paper is an attempt to demonstrate how the structure of the financial system can set the stage for severe fragility: Because of maturity mismatch in a large shadow banking sector without access to a safety net, small shocks can lead to large repercussions. Second, by focusing on regulatory arbitrage as the sole reason for the existence of a shadow banking, we ignore potential positive welfare effects of shadow banking and securitization, such as catering to the demand for liquid assets or improving risk allocation. However, the fragility that arises in the context of regulatory arbitrage arguably also exists for other types of banking activities outside the regulatory perimeter.

Despite the simple nature of our model, we can still draw some conclusions. Our key finding is that the size of the shadow banking sector plays a crucial role for the stability of the financial system. However, the actual quantities of shadow banking activities are not completely clear to academics and regulators. Therefore, a first important implication of our model is that the size of the shadow banking sector (or, more precisely, the magnitude of maturity mismatch in the shadow banking sector) and the interconnectedness of banking and shadow banking should be variables that regulating authorities keep track of. The *Global Shadow Banking Monitoring Report* (Financial Stability Board (2013)) displays a very valuable step in the right direction. Still, the report calls for devoting even more resources to tackling concrete data issues. Our model can be taken as an argument in support of this view.

We make a strong case for why regulatory arbitrage poses a severe risk to financial stability. However, it would be wrong to conclude that regulation should thus be reduced.[17] One needs to keep in mind that – under the presumption that regulation is in place for a good reason – it is not regulation itself that poses a problem, but the circumvention of regulation. If the regulator insures depositors in order to eliminate self-fulfilling runs of depositors, she may need to impose some regulation on banks in order to prevent moral hazard. Regulatory arbitrage may eventually reintroduce the possibility of runs. However, this does not alter the fact that it is a good idea to aim at preventing runs in the first place.

Under the premise that regulatory arbitrage cannot be prevented at all, our model indicates that financial stability may not always be reached by providing a safety net and regulating banks. One may consider a richer set of policy interventions that go beyond safety nets and regulation. E.g., the government or the central bank may have the ability to intervene on the secondary market in case of a crisis. However, such interventions are likely to give rise to different problems as they may change incentives ex-ante, e.g., they may give rise to excessive collective maturity mismatch as in Farhi and Tirole (2012). A richer model than ours would be needed to analyze such effects consistently.

In turn, under the premise that regulatory arbitrage can be prevented or can be made more difficult, we argue that it should be prevented or at least reduced. Given that regulatory arbitrage can be very costly in terms of creating systemic risk, it should be made very costly to those who are conducting it. While this may sound self-evident at first, a glimpse at the history of bank regulation and its loopholes should be a reminder that regulatory arbitrage and the associated risks have not always been a major concern.[18]

[17] There are also argument against strict regulation, building on reputation concerns or charter value effects, see, e.g., Ordoñez (2013).

[18] See, e.g., Jones (2000) for an early analysis of how the Basel requirements were circumvented.

References

Acharya, Viral V., Philipp Schnabl, and Gustavo Suarez (2013). "Securitization without Risk Transfer". *Journal of Financial Economics* 107, 515–536.

Adrian, Tobias and Adam Ashcraft (2012). "Shadow Banking: A Review of the Literature". In: *The New Palgrave Dictionary of Economics.* Palgrave Macmillan.

Allen, Franklin and Douglas Gale (1994). "Limited Market Participation and Volatility of Asset Prices". *American Economic Review.*

Bolton, Patrick, Tano Santos, and Jose A. Scheinkman (2011). "Outside and Inside Liquidity". *Quarterly Journal of Economics* 126, 259–321.

Brunnermeier, Markus K. (2009). "Deciphering the Liquidity and Credit Crunch 2007-2008". *Journal of Economic Perspectives* 23, 77–100.

Claessens, Stijn, Zoltan Poszar, Lev Ratnovski, and Manmohan Singh (2012). "Shadow Banking: Economics and Policy". *IMF Staff Discussion Note.*

Coval, Joshua, Jakub Jurek, and Erik Stafford (2009). "The Economics of Structured Finance". *Journal of Economic Perspectives* 23, 3–26.

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

Dang, Tri Vi, Gary B. Gorton, and Bengt Holmström (2013). "Ignorance, Debt and Financial Crises". *Unpublished, Yale SOM.*

Diamond, Douglas W. and Philip H. Dybvig (1983). "Bank Runs, Deposit Insurance, and Liquidity". *Journal of Political Economy* 91, 401–419.

Farhi, Emmanuel and Jean Tirole (2012). "Collective Moral Hazard, Maturity Mismatch and Systemic Bailouts". *American Economic Review* 102, 60–93.

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

Financial Stability Board (2013). *Global Shadow Banking Monitoring Report 2013.* Basel: Bank for International Settlement.

Gennaioli, Nicola, Andrei Shleifer, and Robert W. Vishny (2013). "A Model of Shadow Banking". *Journal of Finance* 68, 1331–1363.

Goodhart, Charles A. E., Anil K Kashyap, Dimitrios P. Tsomocos, and Alexandros P. Vardoulakis (2012). "Financial Regulation in General Equilibrium". *NBER Working Paper* 17909.

——— (2013). "An Integrated Framework for Analyzing Multiple Financial Regulations", 109–143.

Gorton, Gary B. (2012). *Misunderstanding Financial Crises.* Oxford University Press.

Gorton, Gary B. and Andrew Metrick (2010). “Regulating the Shadow Banking System”. *Brookings Papers on Economic Activity*, 261–312.

——— (2012). “Securitized Banking and the Run on Repo”. *Journal of Financial Economics* 104, 425–451.

——— (2013). “Securitization”. In: *Handbook of the Economics of Finance*. Ed. by Milton Harris George M. Constantinides and Rene M Stulz. Vol. 2, Part A. Handbook of the Economics of Finance. Elsevier, 1–70.

Gorton, Gary B. and George Pennacchi (1990). “Financial Intermediaries and Liquidity Creation”. *Journal of Finance* 45, 49–71.

——— (1995). “Banks and Loan Sales: Marketing Nonmarketable Assets”. *Journal of Monetary Economics* 35, 389–411.

Gorton, Gary B. and Nicholas S. Souleles (2006). “Special Purpose Vehicles and Securitization”. In: *The risks of financial institutions*. Ed. by Rene M. Stulz and Mark Carey. University of Chicago Press.

Hellwig, Martin (1994). “Liquidity Provision, Banking, and the Allocation of Interest Rate Risk”. *European Economic Review* 38, 1363–1389.

Holmström, Bengt and Jean Tirole (1997). “Financial Intermediation, Loanable Funds, and the Real Sector”. *Quarterly Journal of Economics* 112, 663–691.

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, 26–47.

Jones, David (2000). “Emerging Problems with the Basel Capital Accord: Regulatory Capital Arbitrage and Related Issues”. *Journal of Banking & Finance* 24, 35–58.

Kacperczyk, Marcin and Philipp Schnabl (2009). “When Safe Proved Risky: Commercial Paper during the Financial Crisis of 2007-2009”. *Journal of Economic Perspectives* 24, 29–50.

Krishnamurthy, Arvind, Stefan Nagel, and Dmitry Orlov (2013). “Sizing up Repo”. *Journal of Finance, forthcoming.*

Martin, Antoine, David Skeie, and Ernst-Ludwig von Thadden (2014). “Repo Runs”. *Review of Financial Studies* 27, 957–989.

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.*

Pozsar, Zoltan, Tobias Adrian, Adam Ashcraft, and Hayley Boesky (2013). “Shadow Banking”. *Federal Reserve Bank of New York Economic Policy Review*, 1–16.

Qi, Jianping (1994). “Bank Liquidity and Stability in an Overlapping Generations Model”. *Review of Financial Studies* 7, 389–417.

Segura, Anatoli (2014). “Why did Sponsor Banks Rescue their SIVs? A Signaling Model of Rescues”, 1–47.

Shleifer, Andrei and Robert W. Vishny (1997). “The Limits of Arbitrage”. *Journal of Finance* LII, 35–55.

——— (2011). “Fire Sales in Finance and Macroeconomics”. *Journal of Economic Perspectives* 25, 29–48.

Stein, Jeremy C. (2010). “Securitization, Shadow Banking & Financial Fragility”. *Daedalus* 139, 41–51.

Thadden, Ernst-Ludwig von (1998). “Intermediated versus Direct Investment: Optimal Liquidity Provision and Dynamic Incentive Compatibility”. *Journal of Financial Intermediation* 7, 177–197.

A. Moral Hazard and Regulatory Costs

In the model described above, regulatory costs enter as an exogenous parameter γ. In this section, we extend the model by a few aspects to provide a foundation for this assumption. We show that once a bank is covered by a safety net that is in place to prevent self-fulfilling runs (e.g., a deposit insurance), the bank will not be disciplined by investors and will have incentives to invest in a riskier project with private benefits. The regulator thus needs to impose a minimum capital requirement in order to ensure that the bank behaves diligently. As raising capital is assumed to be costly for the bank (e.g., due to dilution costs), the overall return a bank will make will be reduced by the regulation. We recommend reading this part only after having finished reading Section 3.

Let us assume that commercial banks as well as shadow banks are run by owner-managers. Assume that bank managers receive some constant private benefit w (per unit of deposits) as long as their (shadow) bank is operating. If the bank goes bankrupt, the manager loses his bank and his income. The manager discounts the future at rate $\delta < 1$, his discounted income over his (infinite) lifetime is given by $w/(1-\delta)$. Now assume that next to the short asset and the long asset described in the beginning of Section 2, bank managers also have access to an additional production technology that we call "private asset". This private asset is similar to the long asset, but it has the property that, with some probability α, the asset defaults completely. In addition, this asset produces some private benefit b (per unit) for the bank manager. We assume that the long asset associated with a private benefit is never socially optimal, i.e.,

$$R > (1-\alpha)R + b.$$

This structure is reminiscent of how moral hazard is introduced by e.g. Holmström and Tirole (1997).

If the manager invests in the private asset instead of the long asset, the bank still offers the same demand deposit contract as in the standard case. The bank can serve its depositors with probability $1-\alpha$, but with probability α it defaults. We assume that investors can observe what the manager is doing. However, this monitoring is associated with private costs for the investors. We assume that these monitoring costs vary across investors and each investor i has some monitoring costs s_i which are drawn from $G(s)$. These monitoring costs are equivalent to the switching costs introduced in the main part of the paper.

There are three different environments that a (shadow) bank can operate in: In the first environment, the manager holds no equity and his depositors are not protected by a deposit insurance. In the second environment, the manager does not hold (inside) equity either, but his depositors are protected by a deposit insurance. In the third environment, the manager does hold an equity position e.

Let us consider the case without equity and without deposit insurance first. The absence of deposit insurance induces investors to monitor the manager and to withdraw (or not deposit) their funds if the manager misbehaves. Therefore, the manager will behave diligently.[19]

In contrast, in the presence of a deposit insurance scheme, investors do not care about what the manager is doing. If the manager has no "skin in the game" (i.e., if he has no inside equity), he chooses the private asset iff

$$b > [1 - (1 - \alpha)\delta] \frac{w}{(1 + \delta)}.$$

If this inequality is satisfied, the deposit insurance becomes tested, i.e. has to cover claims, with probability α. The regulator therefore has an incentive to ensure diligence of the manager by regulating him. While there are multiple ways to regulate a bank manager, we assume that the regulation requires the bank to hold a minimal amount of equity e per unit of deposits.

This changes the manager's incentives. Because he now has "skin in the game", he will behave diligently whenever

$$e > b - [1 - (1 - \alpha)\delta] \frac{w}{(1 + \delta)}.$$

By choosing an equity requirement $\bar{e} \equiv b - [1 - (1 - \alpha)\delta]w/(1 + \delta)$ per unit of deposit, the regulator can ensure diligence.

Formally, incorporating this moral hazard and the resulting regulation into the framework of banking and shadow banking in Section 3 works in the following way: There exists a sector of commercial banking which is covered by a safety net and regulated to prevent moral hazard. Bank managers have to raise equity, and this is costly, e.g., due to dilution costs. We define the cost of raising e units of equity to be the regulatory cost γ.

There also exists an unregulated shadow banking sector which is not subject to this capital requirement. However, there is a securitization cost of ρ. In addition, investors who choose to deposit their funds with the shadow banks have to spend the monitoring cost s_i. As shown in Section 3, only investors with costs $s_i < s^*$ will choose to invest in the shadow banking sector.

[19] For tractability, we abstract from the strategic interaction between investors which arises from monitoring having positive externalities.

B. Heterogeneous Investors

In the main part of the paper, we assumed that all investors have the same size (endowment of one unit). The heterogeneity among investors consists in the switching costs s_i that are distributed according to some distribution function G in the population. We argued that there are several reasons why investors have heterogeneous switching costs, and that the only necessary feature of the model is that switching costs relative to the investors' budget is heterogeneous.

In this appendix we want to show that we obtain qualitatively similar results if we assume that all investors have identical switching costs, but different endowments. For simplicity, let us assume that the investors' endowment is either high, $x_i = x_h$, or low, $x_i = x_l$. The fraction of "large investors", i.e., with a high endowment, is given by p. The switching cost is assumed to be identical across investors, $s_i = s$. For convenience, we assume that switching costs are monetary, i.e., the utility from receiving c units from a shadow bank is given by $u(c - s)$.

For an investor with endowment x_i, the expected utility of depositing at a commercial bank is given by

$$EU_b(x_i) = \pi u(x_i r_b) + (1 - \pi)u(x_i r_b^2),$$

while the utility from depositing at a shadow bank is given by

$$EU_{sb}(x_i) = \pi u(x_i r_{abcp} - s) + (1 - \pi)u(x_i r_{abcp}^2 - s).$$

Again, an investor chooses the shadow bank if $EU_{sb}(x_i) > EU_b(x_i)$. If the endowments and the switching costs are such that $EU_{sb}(x_h) > EU_b(x_h)$ and $EU_{sb}(x_l) < EU_b(x_l)$, all large investors choose the shadow banking sector, while all small investors stay with the commercial banks. The size of the shadow banking sector is thus given by the fraction of "large investors", p. And the size of the commercial banking sector is thus given by the fraction of "small investors", $1 - p$.