Shadow Banking and the Four Pillars of Traditional Financial Intermediation[*]

Emmanuel Farhi[†] and Jean Tirole[‡]

December 21st, 2017

Traditional banking is built on four pillars: SME lending, access to public liquidity, deposit insurance, and prudential supervision. This paper unveils the logic of the quadrilogy by putting core services to "special depositors and borrowers" at the heart of the analysis, and makes room for bank and depositor implicit and explicit guarantees. It analyzes how prudential regulation must adjust to the emergence of shadow banking. The model also rationalizes ring fencing between regulated and shadow banking and the sharing of liquidity in centralized platforms to counter syphoning and financial contagion.

Keywords: Retail and shadow banks, lender of last resort, deposit insurance, supervision, migration, ring fencing, CCPs.

JEL numbers: E44, E58, G21, G28.

1   Introduction

Traditional banking is built on four pillars: the commercial or retail bank lends to small and medium enterprises, is prudentially supervised and in exchange gets access to public liquidity and to deposit insurance. It caters to "special depositors", who want a liquid and safe vehicle for their savings, and to "special borrowers", the small and medium enterprises that need close oversight to secure financing. Other investors and borrowers have access and resort to financial markets. Other financial institutions traditionally have been left unregulated and could not claim access to deposit insurance and public liquidity.

This definition of retail banking (and by default of shadow banking) with exceptions has been left unquestioned in academia. Yet, the access to the discount window and other liquidity facilities on the one hand and to cheap deposits on the other can be priced and could be offered to the financial system as a whole. Besides, this conventional definition of retail banking is called into question by recent developments. Many shadow financial institutions (money market mutual funds, hedge funds and investment banks) gained access to public liquidity facilities during the 2008 crisis. Shadow banks in China lend to small and medium enterprises and cater to retail depositors through wealth management funds. Should we reconsider the time-honored SME lending/regulation/ public liquidity access/ deposit insurance quadrilogy?

To start answering this question, we unveil the logic of the conventional wisdom, which has traditionally been taken for granted. While our work is therefore intrinsically normative, we build on experience to identify the hazards facing financial systems: risk taking, migration, cross exposures between retail and shadow banking sectors, bailouts of banks and investors. Mechanism design calls for optimally devising institutions that best mitigate the combined hazards.

This paper first argues that there are basic complementarities between regulation and the other components of the quadrilogy. Through its monopoly privilege on taxation- i.e. its access to future earnings, the state has a special ability to create liquidity and therefore to provide insurance to banks and/or individuals when private markets are unable to provide insurance. However, deposit insurance and lender of last resort (LOLR) services are costly for society as they require the state to raise funds even in financial straits. Regulation lowers the cost of these put options on taxpayer money to the extent that it controls risk taking and reduces banking moral hazard. SME lending magnifies the benefit of regulation, as the fear of industrial disruption may trigger ex-ante-unwanted banking bailouts. Overall, the broad normative picture is one in which politically sensitive banking clients- small depositors and SMEs-, who are most affected by a banking failure, are served by a regulated entity and benefit from extended insurance from the state. For example, this picture chimes with the UK legislation, which is cast in terms of the continuity of provision of "core services" – to households and SMEs that lack non-bank alternatives.

The paper's second contribution is to develop a rationale for the two concepts of ring fencing and migration of transactions towards central clearing counterparties (CCPs)[1]. Ring fencing and CCPs feature prominently in a number of post-crisis reforms world-

---

[1]CCPs become the buyer to every seller and the seller to every buyer; they thereby ensure the future performance of open contracts. Under the Basel framework, clearing member banks operating through a "qualified CCP" will get preferential capital treatment.

wide, and, for the former, in the philosophy of the Glass-Steagall act (in force from 1933 through 1999 in the US) separating regulated commercial and unregulated investment banking. Nonetheless, to the best of our knowledge, these policies have not yet been subject to a formal analysis. To perform such an analysis, we introduce a rationale for cross-exposures among financial institutions: Imperfectly correlated liquidity shocks create scope for desirable liquidity pooling and therefore counterparty risk. We show that the provision of mutual insurance among financial intermediaries is subject to gaming in which either a regulated bank is only partially covered by its insurance counterparty and therefore holds "bogus liquidity" (the "AIG syndrom"[2]) or public liquidity is syphoned off to benefit a shadow banking entity (the "conduit syndrom"). We then study the optimal policy responses to such hazards.

The paper is organized as follows. Section 2 sets up the framework. There are three periods, 0, 1 and 2. At date 0, the state picks its optimal regulatory policy. The state can commit to its date-1 policy, but cannot commit not to renegotiate if there are mutual gains from trade. Concretely, this distinction introduces a difference between a bailout and the exercise of the LOLR and deposit insurance functions. A bailout is an uncontracted-for rescue of a bank or class of investors at date 1, perhaps unwillingly from the point of view of date 0; that is, the state might like to commit at date 0 not to engage in bailouts, but at date 1 finds it preferable to rescue distressed banks or investors, which of course does not object to the state's not keeping its promise. By contrast, LOLR and deposit insurance generate rescues that the state would not spontaneously undertake at date 1, and therefore requires the state's ability to commit.[3] The optimal regulatory contract specifies a right for the state to monitor the bank's activities (reducing moral hazard) if the bank joins the regulated sector, and a transfer, which embodies the prices of access to LOLR and to deposit insurance (the latter when a rationale for deposit insurance is introduced in Section

---

[2]The notion of "bogus liquidity" is documented by Yorulmazer (2013), who analyses the correlation between the insurer's default and the bank's shocks and argues that CDSs, which according to Basel regulation, can be counted as hedges and allow banks to free up regulatory capital, have been used to create a false sense of safety due to counterparty risk. There have also been concerns that regulated banks be dependent on investment funds for their short-term funding (see Jin-Nadal de Simone 2016 for evidence on the exposure of major European banks to investment funds).

[3]LOLR is often described as following Bagehot's dictum: To avert panics, central banks should lend early and freely (i.e. without limit), to solvent firms, against good collateral, and at "high rates". In practice, it is very difficult to distinguish illiquidity from insolvency, and LOLR ends up subsidizing financial institutions. We focus on this dimension and model LOLR as a commitment at date 0 to bail out banks at date 1. We could capture the notion that LOLR is only deployed against good collateral by introducing a new dimension on moral hazard which cannot be eliminated by regulation, whereby banks can take an unobservable action that increases the value of their collateral. In order to provide incentives for banks to take this action, the implicit LOLR contract would require committing to bail out banks only if the value of their collateral is high enough.

3). The conditions are optimally differentiated between the regulated and shadow sectors. In fact, the contract must satisfy the banks' "participation constraint", as the latter always have the option to operate in the shadow banking sector.

Banks then select in the menu designed by the regulator (in the basic model, in which they are ex-ante identical, they all select the same option). They then borrow from investors. Investors are risk neutral in Section 2; later on, we introduce special depositors who are risk averse. At date 1, banks may suffer a liquidity shock, i.e., require more money in order to continue their activity until date 2. A bank's probability of facing such a shock depends both on a macroeconomic shock and on a non-contractible date-0 choice by the banker. Regulation in our model consists in monitoring and reducing the bank's risk-taking and involves a cost (say, a compliance cost for the bank or a surveillance cost for the monitor). Thus in equilibrium, the regulated banks' asset portfolio is safer than the shadow banks', and regulated banks are also more leveraged[4]; but this difference is endogenous and does not result from any posited intrinsic difference between the two. Because liquidity shocks are correlated absent moral hazard and there is a shortage of outside private liquidity in the economy, the state is the only player who can rescue the banks in dire straits.

Date-1 rescues involve both a benefit and a cost. The benefit for the state is that it keeps the real economy going (in our model, the state does not care about the bankers per se). This is where SME lending enters the picture: Absent SME lending, the state would have no incentive to bail out banks, and there would be no moral hazard and no market failure. The cost is that bailouts create date-0 moral hazard. How this tradeoff is resolved depends on a second macroeconomic shock: The date-1 cost of public funds is random (alternatively, the social acceptability of bailouts could be state contingent). In the absence of prior contract between banks and the state, the state bails out the banks when the cost of public funds is low, but not when it is high.

In the absence of special depositors, the state can take on two functions: regulator, where regulation is to be understood as costly supervision that limits risk taking by the regulated banks; and lender of last resort, that allows the bank to receive liquidity assistance even in states in which the state would not have provided this liquidity spontaneously. Bankers' outside option is to migrate to the unregulated, shadow banking sector. We assume that the state has no efficiency advantage in monitoring the banks.[5] Our first key result is the complementarity between regulation and LOLR: offering LOLR

---

[4]Berg and Gider (2016) document that retail banks have much more leverage than other financial institutions, and that this is mainly explained by the fact that they hold lower risk portfolios.

[5]Our assumption that the fixed cost of monitoring is the same across potential monitors ignores incentive problems in monitoring, which will be discussed at the end of the paper.

is cheaper if the state simultaneously controls risk taking.[6]

Section 3 introduces special depositors. Section 3.1 considers the case of special depositors with infinite risk aversion; more precisely, they have Epstein-Zin preferences and are willing to pay a premium only for a totally safe asset. Like for banks the state is able to offer these depositors a unique service, which is the guarantee that they can dispose of their money even in bad macroeconomic states of nature. The state can charge for this service, and so a priori it is unclear whether the state wants to favor specific financial institutions and grant them exclusive access to deposits. However, for the same reason as for LOLR, the social cost of providing deposit insurance is lower in the regulated sector than in the riskier shadow banking sector. So deposit insurance and regulation are also complements.[7] Deposit insurance can be used to allow for the creation of safe deposits.

Section 3.2 assumes instead that special depositors' preferences, and the associated demand for safe assets, admit a Von Neumann-Morgenstern representation; special depositors need money at date 1 to accomplish or fulfill specific needs. We assume that the state cannot directly observe whether given savings belong to an ordinary or special depositor. The state may be tempted to make good on a financial claim that is held primarily by special depositors when the claim fails to deliver; to see that this is a real concern, just consider the nervousness of authorities when money market mutual funds break the buck, or life insurance companies that have promised a minimum return do not succeed in delivering this return.

We show that the shadow banking sector may cleverly use financial engineering so as to attract special depositors and create a put on taxpayer money. Another interesting feature of this variant is the phenomenon of clientele-dependant valuations: special depositors may in the absence of bailout prefer portfolio 1 to riskier portfolio 2, but nonetheless purchase (outbid ordinary depositors on) portfolio 2 that then becomes safer than portfolio 1 due to the investor bailout triggered by the special-depositor-heavy clientele. For the sake of that section, we show that there exists a "monotonic equilibrium" in which intrinsically safer portfolios are more attractive to special depositors and therefore more likely to trigger a bailout, and we focus on this equilibrium. Laissez faire is suboptimal for three

---

[6]It is interesting to note that in one important dimension, post-crisis reforms did not take this complementarity on board: The Fed now has much more regulatory authority over the broker-dealer subsidiaries of large bank holding companies, and subjects them to a large array of capital requirements and stress tests. But the Dodd Frank act has at the same time made it much harder for the Fed to act as an LOLR to these non-bank subsidiaries. Instead, our logic implies that a good reason to regulate the broker-dealers more stringently is precisely to be able to give them access to LOLR in times of liquidity crisis.

[7]As Peltzman (1989) argues: "The putative motive for this [government guarantee of deposits] subsidy is to use the banks as the government's agents for providing a cheap, liquid substitute for government money. The quid for this quo is that banks should refrain from using their access to the government guarantee simply to maximize profits." [We are grateful to John Vickers for referring us to Peltzman's work].

reasons over and above those that were analysed in the basic model of Section 2. Some ordinary depositors benefit from an investor bailout as they mix with special depositors. Second, moral hazard is even costlier than in the absence of special depositors, as it generates investor bailouts on top of bank bailouts. Third, shadow banks enjoy a large rent, as they sell the safe asset at a price that reflects the prospect of an investor bailout by the state. We show that all three problems can be resolved by incentivizing special depositors to remain in the regulated sector. Again deposit insurance and regulation co-vary.

Thus the overall picture is that the SME lending/regulation/ public liquidity access/ deposit insurance quadrilogy is a natural one due to the complementarity between regulation and each of its three other components. Note that this view is not inconsistent with the idea that deposits are a cheap form of funds for retail banks. To see this, observe that the state operates two transfers of opposite directions when providing regulation and deposit insurance: to prevent a bank from migrating to the shadow banking sector, the state must make a transfer to the bank; this transfer can be subtracted from the deposit insurance premium, thus transforming retail deposits into a cheap form of funds that is exclusively available to regulated banks.

Generalizing the preceding analysis, Section 4 assumes that financial institutions face liquidity shocks that are only partially correlated, and thereby introduces the possibility of liquidity pooling (such pooling is infeasible in our basic model because banking shocks are perfectly correlated).

Liquidity pooling among banks raises the issue of cross-exposures and contagion. Of primary interest in the policy debate are exposures of the regulated sector to the shadow banking sector, which played an important role in the 2008 financial crisis (many regulated entities were exposed to the failures of unregulated investment banks such as Lehman Brothers and AIG's holding). After the crisis a variety of structural reforms separating retail from investment banking were proposed, and for some implemented, so as to insulate basic banking services—and thereby depositors, the guarantee fund, and taxpayers—from investment banking risks (Volker in the US, Liikanen in Europe). In the UK, Vickers' rule creates a ring-fenced subsidiary (the retail bank) with a limited scope of activities: It can lend only to households and nonfinancial firms, trade high-quality securities and hedge the risk on corresponding exposures. All other activities are not allowed within the ring-fenced bank but can be performed by the rest of the bank (the investment bank). The ring-fenced bank has operational independence and is prohibited from providing support to the investment bank.

We provide conditions under which such ring-fencing is warranted. The idea is that the retail bank can abandon supervision-induced caution by taking exposures in the

shadow banking sector in a way it could not do in the regulated sector. More precisely, a regulated bank can use (otherwise desirable) liquidity pooling in order to foil the supervisory oversight in two ways. First, it can "secure" funds from the shadow sector that will not materialize in some contingencies in which they are needed by the regulated bank, creating contagion from the shadow sector to the regulated sector. Such "bogus liquidity" is exemplified by the failure of the CDSs provided by AIG to make good on their promises. Second, state-provided liquidity available to the regulated sector may be syphoned off to the shadow sector, so that the latter indirectly benefits from LOLR. Sticking to the 2008 crisis, a case in point is provided by the lines of credits awarded by the regulated banks to the conduits that they created in the securitization process.[8]

Ring fencing can help prevent such abuses. We first assume that when counterparties are both supervised, the regulator can learn the correlation structure between them (say through joint stress testing). It thereby can prevent the hazards described above; a simple regulation forcing regulated banks to co-insure through mutual lines of credit (which is a form of liquidity regulation) then delivers the second-best welfare level. In contrast, such an understanding is not available if one of the parties lies outside the regulated sphere, and liquidity pooling can then game the supervisory system.

Second, we make the opposite polar assumption that, unlike the counterparties, the regulator never learns the correlation structure. Ring fencing then no longer suffices to deliver the second best. Regulated banks can game the liquidity requirements and arrange bogus liquidity lines to each other, knowing that they will be protected by bailouts or LOLR anyway. They thereby maximize their put on taxpayer money. To restore the second best, the regulator can complement ring fencing with the requirement that liquidity pooling occur through a centralized exchange rather than bilaterally. This prevents banks from fine-tuning their liquidity provision at the expense of the taxpayer.

More generally one can imagine that banks are heterogeneous in their activities and so logically are not subject to exactly the same shocks. Heterogeneous activities also can help justify the existence of a shadow banking sector, which so far was a pure nuisance for the social planner (shadow banking defined the banks' reservation values and augmented their rents). It may be that some activities, such as SME and mortgage lending or plain interest-rate and exchange-rate derivatives, are sufficiently well-understood to be reasonably supervised by the state, while others involve very complex instruments such as bespoke derivatives, that either are poorly understood by the state or are extremely time-consuming to monitor and assess. In this case shadow banking is socially useful,

---

[8]Shin (2012) documents the staggering amount of funding of shadow banking by retail banks. A key aim of ring-fencing reforms is certainly to curtail the resulting exposure.

but is still a constraint on what the regulator can achieve. Ring-fencing on the other hand limits liquidity pooling and so therefore now comes with a meaningful tradeoff.

Sections 5 and 6 study extensions, robustness checks, and alleys for future research.

Relationship to the literature.   There are widely different views, both among economists and in the policy debate, about the social merits of shadow banking. The most positive view states that regulatory constraints stifle innovation, limit lending and distort markets; shadow banking then offers some breathing room and undoes a state failure. See for example Ordonez (2017) for an elaboration of this point in a model where banks are asymmetrically informed about their investment opportunities, and where migration into the shadow banking sector provides a way for the banks with the best opportunities to pursue them by avoiding blunt regulation. Feve-Pierrard (2017) provide some evidence of such migration in response to higher capital requirements. See also Buchak et al. (2017) who study the rise of fintech and non-fintech shadow banks in the residential lending market and find that financial technology can account for about 35% of shadow bank growth over the period 2007-2015.

Different strands of the academic literature articulate a more negative view. One branch of the literature stresses regulatory arbitrage: Shadow banking is then a (perhaps unavoidable) nuisance. The regulatory arbitrage view includes two possible subviews. In the first, retail banks evade capital requirements by providing liquidity support off-balance-sheet to shadow banks; Acharya et al (2013) find evidence that such regulatory arbitrage was a key motive behind setting up ABCP conduits, as losses from conduits remained with retail banks[9]. The underpricing of this absence of effective risk transfer was corrected by Basel 3, which put the corresponding exposures back on the retail bank's balance sheet. The second subview, spelled out for example in Acharya-Richardson (2009) and Claessens et al. (2012), involves capital requirement "evasion" by shadow banks, which face no capital adequacy requirement and yet receive public assistance. Shadow banks cut regulatory corners and have their cake and eat it too: They are free of constraints in normal times, and are bailed out if tail risk materializes.[10] Perhaps consistent with this view, the aforementioned study Buchak et al. (2017) also finds that the migration to shadow banking induced by the increasing regulatory burden faced by traditional banks account for 55% of shadow bank growth over the same period.

---

[9]See also Gorton-Metrick (2010) and Pozsar et al (2013).

[10]In the context of these two subviews, Farhi-Tirole (2012, 2017) and Di-Iasio-Pierobon (2012) emphasize strategic complementarities in regulatory arbitrage arising from a security in numbers due to the fact that bailouts are imperfectly targeted.

Another branch of the literature stresses behavioral factors: Shadow banks exploit neglected risk. Gennaioli et al (2012, 2013, 2015) assume that investors overweigh a favorable scenario upon good news and similarly overreact when bad news occur. Shadow intermediaries create false substitutes for truly safe bonds. Financial crises can be triggered by the repricing of risk following the sudden realization of the true risks embedded in these pseudo-safe assets. In Section 3.2 of our paper, shadow banks can create relatively (but not entirely) safe assets via financial engineering to attract special depositors but without exploiting the behavioral biases of the latter.

Finally, a last branch of the literature emphasizes comparative advantage.[11] For example, in Hanson et al (2015), households are willing to pay a premium for safe assets, as in Stein (2012).[12] Safe assets can be created in two ways; in the regulated sector through deposit insurance offered by the state in exchange of costly capital requirements; by an early exit option and the costly liquidation of assets in the shadow sector. In equilibrium, shadow banks therefore hold relatively liquid assets. The paper does not analyze optimal regulation, but identifies an externality in the unregulated sector, due to fire sales. This externality creates a tendency for the shadow banking sector to be too large compared to the regulated sector. Chrétien and Lyonnet (2017) pursue this logic by assuming that rather than outside investors, it is banks in the regulated sector that purchase the assets that are liquidated by shadow banks, and that they do so using cheap insured deposits. They study the resulting interactions between the two sectors. Relatedly, Gertler et al. (2016) build a model in which wholesale shadow banks borrow from regulated retail banks which in turn raise deposits from households. In their model, the relative size of the two sectors is determined by a tradeoff between assumed comparative advantages of wholesale banks in managing assets and of retail banks in overcoming agency frictions in fund borrowing. In a different vein, Moreira-Savov (2017) emphasize the coexistence of money (securitization products that are safe and liquid all the time) and shadow money (securitization products that are safe and liquid most of the time). In their model, compared to money, shadow money economizes on collateral but is more fragile. Periods of low uncertainty are associated with expansions in shadow money and economic booms, which come to an end when uncertainty increases, shadow money collapses, and the economy tanks.

---

[11]See e.g. Perotti (2014) for an early policy discussion.

[12]The demand for safe assets also figures prominently in Diamond (2017)'s theory of segmentation. In Diamond, firms tranche their liabilities so as to create relatively safe assets (debt), which are then held by banks. Banks transform these assets into really safe assets (deposits) through an equity add-on. In our model, only the state can create safe assets, but it finds it cheaper to do so if banks themselves hold relatively safe assets. The state then optimally piggybacks on the banks' balance sheets to do so.

Our model incorporates elements of these different branches of the literature. At its core is a problem of regulatory arbitrage, along the lines of the two corresponding sub-views mentioned above: Shadow banks avoid the capital requirements of the regulated sector and yet receive some public support in the form of bailouts; banks in the regulated sector must also be prevented by regulation from extending liquidity support to shadow banks. An extension of our model (see Sections 3.1 and 4) also incorporates a notion of comparative advantage: Some activities are simply too costly to regulate, perhaps because they are too complex, and so they are better performed by the shadow banking sector. Moreover, to the extent that the risks of the shadow banking sector are not perfectly correlated to those of the regulated sector, then allowing for the two sectors not only to co-exist, but also to share some risks, is desirable (see Section 4.5).

Few papers study optimal regulation in the presence of a shadow banking sector. Beguenau-Landvoigt (2017) solve for optimal capital requirements in a quantitative model where banks can migrate to the shadow banking sector in the presence of exogenous bailouts occurring with a higher probability in the regulated sector than in the shadow banking sector. The idea that regulation must account for the possibility of migration of banking activities can be found in earlier papers.[13] For example, Grochulski-Zhang (2014) analyze a model à la Diamond-Dybvig (1983), where regulation is motivated by a pecuniary externality arising from the possibility of private re-trades among banks as in Farhi et al. (2009), and introduce shadow banking as a nuisance in the form of a participation constraint which limits the scope of regulation. Similarly, Plantin (2015) sets up a model where a bank engages in excessive risk-taking and evades regulatory risk-monitoring through securitization and the granting of lines of credit to the resulting conduits. In his regulatory-evasion model, shadow banking is therefore a nuisance, and he shows that tightening capital requirements may spur a surge in shadow banking activity and reduce welfare. Harris et al. (2014) emphasize a different perverse effect of tighter regulation, namely that increased capital requirements can actually induce risk shifting in the regulated banking sector because of bailouts and because the competition of shadow banks is more intense for safe positive net-present-value projects than for risky negative-present value projects. In a different vein, Bengui-Bianchi (2014) analyze the optimal design of capital controls in a small open economy with pecuniary externalities when some possibility of evasion exists. In their model, tighter capital controls curb risk-taking in the regulated sector, increase it in the unregulated sector, and are overall desirable.

Our theory is unique in explaining the complementarities between regulation, LOLR, and deposit insurance, and in showing how the optimal deployment of these attributes

---

[13]See e.g. Hanson et al. (2011) for an early policy discussion.

endogenously gives rise to a regulated banking sector associated with the aforementioned attributes and a shadow banking sector devoid of them. Relative to the existing literature, our paper also makes forays into two new areas: the complementarity between the four classic markers of traditional banking, and the use of ring fencing and CCPs, adding two further markers. Finally, our paper emphasizes and distinguishes between bank bailouts and investor bailouts, and introduces the notion of clientele-contingent asset valuation.

2 Model

2.1 Environment

There are three periods indexed by $t = 0, 1, 2$, a single good, and three classes of players: investors, bankers, and the government (or "state", or "regulator"). There are two sources of aggregate uncertainty: the fiscal state which can be good (G) with probability $1 - p$ or bad (B) with probability p; and the liquidity state which can be high (H) or low (L). All uncertainty is resolved at date 1. There is no store of value in the economy.

Investors. Investors/consumers have risk-neutral preferences $\mathbb{E}[c_0^I + c_1^I + c_2^I]$ with no discounting over consumption in all three periods. They have large endowments in every period. As in Holmström and Tirole (1998), we assume that they cannot commit to pay any funds in the future. As a result, while they can save, they can neither borrow nor grant credit lines to bankers. This will imply that only the government, through its exclusive access to taxpayer money, can provide liquidity in bad times. This builds the foundations for the unique ability of the government to offer LOLR and deposit insurance.

Bankers. Bankers have sufficient net worth to finance a project requiring investment 1 at date 0. As shown in Figure 1, they have a fixed investment project. This project may be hit by a liquidity shock and require reinvestment 1 in order to prevent liquidation at date 1. If the project is liquidated, it yields a zero payoff. The payoff of the project if it is not liquidated has a pecuniary component $\rho_0 < 1$ which can be pledged to outside investors and a non-pecuniary component in the form of a private benefit b for bankers. Bankers value consumption at date 0 and the private benefit from running their investment project to completion according to $\mathbb{E}[c_0^B + bj]$ where $j \in \{0, 1\}$ is an indicator for the completion of the project. These preferences imply that the banker is incentivized through the threat
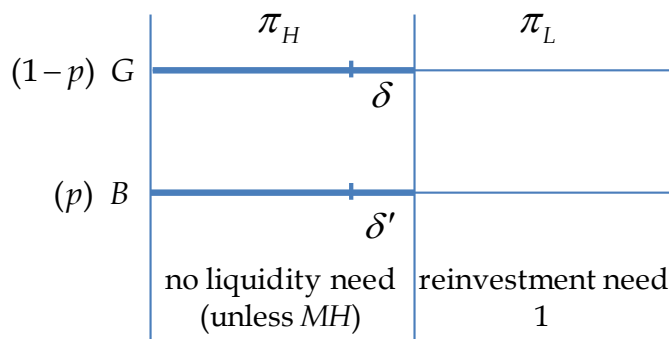
11

of termination rather than through a financial incentive scheme[14].

| 0 | 1 | If $j = 1$ | 2 |

- Banks raise funds and invest (size 1)
- Moral hazard?

- Liquidity shock (needs 1) to continue or not (needs nothing).
- Continue: $j = 1$ Stop: $j = 0$ (no payoff).

- Private benefit b
- Pledgeable income $\rho_0 < 1$

Figure 1: Timing for bank's operations.

Bankers can exert effort on two separable dimensions to reduce the probability of a liquidity shock in each of the fiscal states G (probability $1 - p$) and B (probability p). If they do not exert effort on a given dimension, they gain a small non-pecuniary benefit. In the computations below, we will take this benefit to be zero for notational convenience, but absolutely nothing hinges on this; what matters is that the prospect of a rescue (either a bailout or the use of the LOLR facility) makes banks less mindful of risk. If banks exert effort on both dimensions, the probabilities of a liquidity shock conditional on each fiscal state at date 1 are both equal to $\pi_L = 1 - \pi_H$. If they shirk along the G dimension, then the probability of a liquidity shock conditional on state G is $\pi_L + \delta$. If they shirk along the second dimension, then the probability of a liquidity shock conditional on state B is $\pi_L + \delta'$. See Figure 2. We assume that bankers want to invest at date 0, which is guaranteed by their willingness to invest in the shadow banking sector.[15]

We assume that conditional on exerting effort, liquidity shocks are perfectly correlated across bankers. The extra realizations of liquidity shocks conditional on shirking are also assumed to all be correlated across bankers. These assumptions, which we relax later in Section 4, ensure that bankers cannot benefit by accumulating liquidity in the form of financial claims on other bankers, and that the government cannot economize on taxpayer money by forcing banks to co-insure.

A monitoring technology prevents shirking along each dimension at pecuniary cost c

---

[14]We could alternatively posit that the bankers need to be incentivized through a monetary incentive scheme so as to exert some date-1 effort or to not take a private benefit at that date. This would require assuming a non-negligible cost of effort or private benefit, but would not qualitatively change the analysis: The key feature, shared by all models of liquidity management, is that the surplus associated with continuation is not fully pledgeable to investors and so refinancing problems may emerge.

[15]Because we will make an assumption that guarantees bailouts in state G but not in state B this assumptions amounts to: $[(1 - p) + p\pi_H]b + [\pi_H - (1 - p)\delta]\rho_0 > 1$.

Figure 2: Stochastic structure and moral hazard.

per banker, with $c > [\delta(1-p) + \delta'p]\rho_0$. This assumption ensures that the sole benefit of raising more funds at date 0 is not enough to cover the monitoring costs. This monitoring technology can indifferently be operated by the private sector or by the government, although we will later argue that the government may be the natural monitor because it has a financial stake in cautious bank behavior regardless of who monitors.

Government. The government runs a balanced budget. As shown in Figure 3, it can tax investors and rebate revenues to them in every period. It can use the taxation proceeds to rescue bankers at date 1. Because public funds will be assumed to be costly, a rescue operation optimally transfers resources to bankers after wiping out all legacy investors, who can therefore be seen as "bailinable". The government also runs a lender of last resort (LOLR) scheme which commits at date 0 to bail out bankers at date 1. It sets taxes (or subsidies) τ on bankers at date 0 which depend on their choices regarding the adoption of the monitoring technology and of the LOLR scheme. It will be optimal to rebate proceeds to investors, and thus we assume so.

We assume that there are efficiency costs from collecting extra taxes from investors at date 1. They take the form of iceberg costs such that collecting one unit of revenues at date 1 results in a reduction of consumption of investors by λ_1 units. Furthermore, λ_1 is stochastic with a value of $\lambda > 1$ in state G and $\Lambda > \lambda$ in state B. One can think of λ_1 as the marginal social cost of public funds (MSCPF). Rebating one unit of revenues to investors does not entail any efficiency cost. Put differently, λ_1 being the shadow cost of *extra* taxes, the government's buying shares in the firms brings no social benefit. [A state-contingent MSCPF is one way of creating uncertainty about bailouts; alternative approaches to the same effect include uncertainty about the government resolve not to rescue bank, about the political pressure not to do so, or about the enforcement of state aid regulations.]

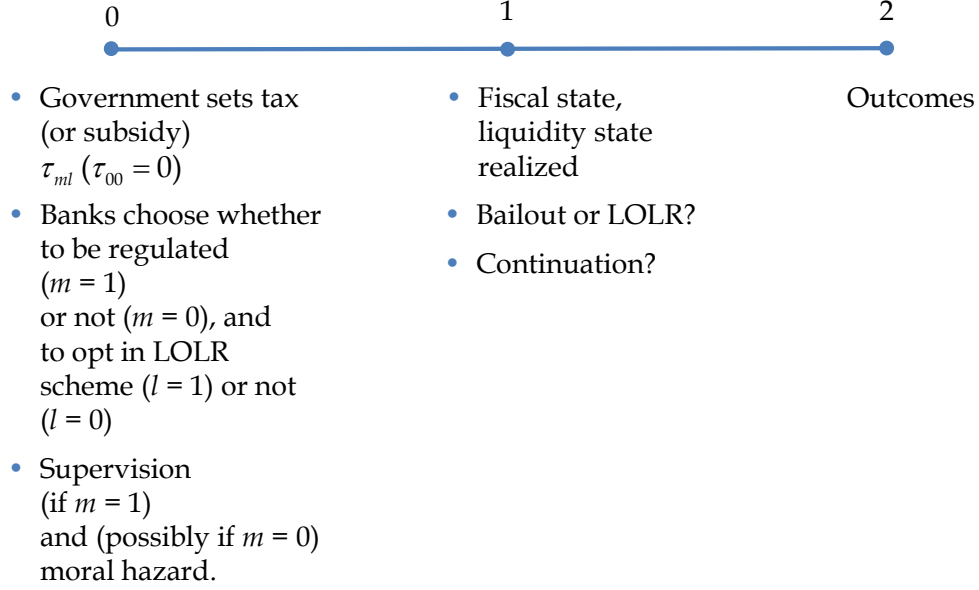The government puts a social welfare weight of 1 on investors. It also puts social

Figure 3: Policy timing.

welfare weights on bank stakeholders (to be interpreted as the SMEs that will not be refinanced due to their reliance on the banking relationship: See footnote [17] below), but no social welfare weight on bankers, with a resulting weight[16] on the completion of a project of β with

$$\Lambda(1 - \rho_0) > \beta > \lambda(1 - \rho_0).$$

Because the liquidity shortfall in case of a liquidity shock is $1 - \rho_0$ after initial investors have been bailed in, this assumption guarantees that bankers hit by a liquidity shock receive a discretionary bailout in state G but not in state B. This also means that ex post at date 1 in state B, the government would like to renege on any promise of bailout that would be part of the LOLR scheme, and is only prevented to do so by its commitment at date 0. This social welfare function can be given credit-crunch foundations, following Holmström-Tirole (1997) to which we refer for a more elaborate model[17].

In principle, the LOLR option could be made contingent on the liquidity state H or L.

---

[16]More generally the social weight put on bankers could be positive, as long as it is smaller than 1. In that case, β is the total weight put on bankers and on bank stakeholders.

[17]For example, at date 0, each bank makes an investment in knowledge/staff so as to be able to invest in a mass 1 of firms, each with investment need 0 or 1 at date 1 and no net worth. The bank monitors firms (or shirks) at both dates 0 and 1. At date 1, the firms monitored at date 0 by the bank need 1 unit of cash each with probability $\pi_L - \delta s_G^i$, in state G (and similarly in state B), where $s_G^i = 1$ in case of date-0 bank shirking, and 0 otherwise. At date 2, firms succeed or fail (then return 0). Success is guaranteed if none of the banker, the managers and the workers shirks. Otherwise, success accrues with probability 0. Shirking at date 1 for a banker brings benefit b, shirking for a firm manager brings benefit b_f and shirking by the workers yields b_w. There is no payoff beyond the incentive payoffs $b + b_f + b_w$ of these stakeholders. We then get $\beta = \beta_f b_f + \beta_w b_w$, where β_f and β_w are the welfare weights for firms and workers

However, we assume that this aggregate state is not verifiable. We justify this assumption by imagining that the only verifiable information regarding aggregate states comes in the form of reports by the government. In our model, the government would always prefer to slant its report ex post at date 1 in order to minimize its LOLR liability. As a result, only non-contingent schemes promising to unconditionally bail out bankers are feasible. This restriction affects only one of our results. In its absence, offering LOLR to regulated and shadow banks would be equally costly for the state: the quadrilogy would become a trilogy, the theory making no prediction on whether LOLR is better targeted at regulated or shadow banks (so it would be offered to both or to none).[18]

Timing. The timing is as follows. At date 0, the government decides on a set of taxes (subsidies if negative) τ_{ml} to be paid by bankers depending on their choices at date 0, with $\tau_{00} = 0$. Bankers then decide whether ($m = 1$) or not ($m = 0$) to be monitored, and whether ($l = 1$) or not ($l = 0$) to opt in the LOLR scheme which guarantees that they will be rescued for sure at date 1. Bankers always have the option of opting out of both schemes and not to pay any taxes. If a bank decides to be monitored, we say that it operates in the regulated sector, otherwise we say that it operates in the shadow banking sector. Formally the transfer paid by a regulated bank, τ_{1l}, is a monetary transfer, but more generally it stands for the cost of various requirements, liquidity and solvency ratios, that will determine how much public money is at stake; our model captures these features in a simple manner.

Bankers then raise funds from investors, pay taxes, and invest. Finally, bankers decide whether or not to exert effort. At date 1, the MSCPF λ_1 and liquidity shocks are realized. Each individual banker who has opted in the LOLR scheme and who experiences a liquidity shock receives the promised bailout. Each individual banker who has not opted in the LOLR scheme and who experiences a liquidity shock may receive a bailout at the discretion of the government. Then bankers who experience a liquidity shock optimally dilute legacy investors and raise additional funds to cover the liquidity cost. If they succeed, the project continues and otherwise it is liquidated. At date 2, payoffs are realized.

---

[18]Were the liquidity state verifiable, the regulator could perfectly detect moral hazard. In state B, the LOLR option could be made contingent on the absence of shirking (there would be no change in state G, as the bank is then bailed out anyway). Note also that if moral hazard could not be perfectly detected by the state (which would be the case if there were idiosyncratic as well as aggregate shocks, and this even if the aggregate liquidity state is verifiable), then regulation would lower the state's cost of LOLR, and so the quadrilogy would re-emerge.

2.2 Equilibrium

At date 0, the amount of funds that a banker can raise from investors if he is expected to exert effort in both dimensions is $\rho_0 \pi_H$. If he is expected to shirk along the G dimension, this amount shrinks by $\rho_0(1-p)\delta$. If he is expected to shirk along the B dimension, it shrinks by $\rho_0 p \delta'$. We always assume that enough funds can be raised to complement their net worth A to finance the investment outlay of 1.

When a banker is hit by a liquidity shock at date 1, he needs to raise resources equal to 1, but even by fully diluting legacy investors, he can only raise ρ_0. Either he receives a bailout in the form of an injection of funds[19] $1 - \rho_0$ and then he can continue, or else he must liquidate his project. Because $\Lambda(1-\rho_0) > \beta > \lambda(1-\rho_0)$, a banker hit by a liquidity shock is bailed out at date 1 in the G state[20], but is rescued in the B state only if it benefits from LOLR.

Bankers' outside option. One can think of each banker as having the outside option of operating in the shadow banking sector without LOLR and without paying any tax. Let us first consider the private tradeoff associated with monitoring. The latter costs c, but boosts pledgeable income and enables the shadow bank to lever up. Because $c > \delta(1-p)\rho_0$, such a banker prefers not to pay the monitoring cost, always exerts effort along the B dimension but not along the G dimension. Similarly, because $c > \delta(1-p)\rho_0 + \delta' p \rho_0$, a banker operating in the shadow banking sector with LOLR prefers not to pay the monitoring cost and exerts no effort along any dimension.

Taxes. As we already alluded to above, we use the indicator variable m (for "monitoring") to denote whether a banker is in the regulated sector ($m = 1$) or the shadow banking sector ($m = 0$). We use the indicator variable l to denote whether a banker benefits from LOLR ($l = 1$) or not ($l = 0$). The outside option corresponds to $(m, l) = (0, 0)$.

The outside option can be thought of as determining a reservation utility, giving rise to individual rationality constraints and influencing the overall level of utility of bankers. Given our notation, this simply imposes $\tau_{00} = 0$. The government still has enough flexibility to induce any pattern of bankers' utility across choices (m, l) with the appropriate choice of taxes τ_{ml}. It can therefore induce any given equilibrium configuration (m, l).

---

[19]The government would never give a larger bailout since it places no direct welfare weight on bankers.

[20]Would the government want to commit not to rescue the bank in state G if it could? Clearly yes if it can identify occurrences of moral hazard, since a credible commitment not to rescue the bank eliminates moral hazard in that state and increases welfare by $\delta(1-p)[\lambda(1-\rho_0) + \rho_0]$. But even if cannot identify moral hazard and so there are states of nature in which bailouts are vindicated, the government would want to commit not to rescue the bank in state G, provided that $\delta(1-p)[\lambda(1-\rho_0) + \rho_0] > \pi_L[\beta - (1-\rho_0)]$.

Bankers' welfare. We adopt the outside option with $(m,l) = (0,0)$ as a baseline.[21] The welfare of the representative banker (gross of taxes) is given by

$$\tilde{U}_{ml} = U_{00} + m[\delta(1-p)\rho_0 - c] + l[p\pi_L b - \delta' p\rho_0] + ml[\delta' p\rho_0].$$

This equation compares the utility of the banker depending on his choice (m,l) to his outside option $(0,0)$.

Suppose first that the representative banker is in the regulated sector but without LOLR so that $(m,l) = (1,0)$. On the one hand, he bears the monitoring cost c, which can be thought of being passed through to him by the government. On the other hand, he can raise additional revenues $\delta(1-p)\rho_0$ because he now exerts effort along the G dimension.

Suppose next that the representative banker is in the shadow banking sector but benefits from LOLR so that $(m,l) = (0,1)$. On the one hand, he loses revenues $\delta' p\rho_0$ since he now shirks along the B dimension because of the additional bailout guarantee from LOLR and investors are bailed in when the bank is illiquid but continues. On the other hand, he can continue his project more often since he is now rescued when he experiences a liquidity shock in the B state, which increases his utility by $p\pi_L b$.

Finally, suppose that the representative banker operates in the regulated sector and benefits from LOLR so that $(m,l) = (1,1)$. Then he experiences a benefit of $\delta' p\rho_0$ over and above the separate net benefits from being in the regulated sector and of having access to LOLR, because regulation prevents shirking along the B dimension induced by access to LOLR.

The welfare of the representative banker net of taxes is given by

$$U_{ml} = \tilde{U}_{ml} - \tau_{ml}.$$

We denote by $\tilde{\tau}_{ml}$ the taxes that make the bankers exactly indifferent between (m,l) and $(0,0)$. They are given by

$$\tilde{\tau}_{ml} = \tilde{U}_{ml} - \tilde{U}_{00} = \tilde{U}_{ml} - U_{00}.$$

Total welfare. Because the government places no direct welfare weight on bankers, the optimal policy for a government wishing to induce the equilibrium configuration (m,l) is to set taxes equal to infinity except for $(0,0)$ and (m,l) for which $\tau_{00} = \tilde{\tau}_{00} = 0$ and $\tau_{ml} = \tilde{\tau}_{ml}$. We now compute the resulting total welfare taking into account the welfare

---

[21]If g denotes the (small) private benefit associated with shirking along the G dimension, then $U_{00} = A - 1 + g + \rho_0[\pi_H - (1-p)\delta] + b[1 - p\pi_L]$. Note though that for notational convenience we have taken $g = 0$.

effects of taxes on the representative banker, the welfare of bank stakeholders, and the fiscal externalities from bailouts on investors:

$$W_{ml} = W_{00} + m[\delta(1-p)[\rho_0 + \lambda(1-\rho_0)] - c]$$
$$+ l[p\pi_L[b + \beta - (1-\rho_0)\Lambda] - \delta'p[\rho_0 + \Lambda(1-\rho_0)]]$$
$$+ ml[\delta'p[\rho_0 + \Lambda(1-\rho_0)]].$$

As above, we adopt the outside option with $(m,l) = (0,0)$ as a baseline.[22]

Suppose first that the representative banker is in the regulated sector but without LOLR, so that $(m,l) = (1,0)$. Because the representative banker experiences a net utility gain of $\delta(1-p)\rho_0 - c$, the tax that he bears and the corresponding transfer to investors must be adjusted resulting in net welfare gain $\delta(1-p)\rho_0 - c$. Because he now exerts effort along the G dimension, bailout costs are reduced by $\delta(1-p)\lambda(1-\rho_0)$.

Suppose next that the representative banker is in the shadow banking sector but with LOLR so that $(m,l) = (0,1)$. Because the representative banker experiences a net utility gain of $p\pi_L b - \delta'p\rho_0$, the tax that he bears and the corresponding transfer to investors must be adjusted, resulting in net welfare gain $p\pi_L b - \delta'p\rho_0$. Because he now benefits from LOLR, he continues even in the B state, resulting in a net welfare gain $p\pi_L[\beta - (1-\rho_0)\Lambda]$, and he now shirks along the B dimension, resulting in additional bailout costs $\delta'p\Lambda(1-\rho_0)$.

Finally, suppose that the representative banker is in the regulated sector with LOLR so that $(m,l) = (1,1)$. Because the representative banker experiences a net utility gain of $\delta'p\rho_0$ over and above the separate net benefits from being in the regulated sector and of having access to LOLR, the tax that he bears and the corresponding transfer to investors must be adjusted resulting in net welfare gain $\delta'p\rho_0$. Because regulation prevents shirking along the B dimension induced by access to LOLR, there is a reduction in rescue costs resulting in a net welfare gain of $\delta'p\Lambda(1-\rho_0)$. The undoing of the moral hazard created by LOLR is captured by the cross term in ml, and is the source of the complementarity between regulation and LOLR.

---

[22]We have $W_{00} = [\beta - \lambda(1-\rho_0)](1-p)(\pi_L + \delta) + \beta\pi_H$.

2.3 LOLR and Regulation Complementarities

A fundamental observation is that there are complementarities between regulation (m) and LOLR (l):

$$\frac{\partial^2 \tilde{U}_{ml}}{\partial m \partial l} = \delta' p \rho_0 > 0 \quad \text{and} \quad \frac{\partial^2 W_{ml}}{\partial m \partial l} = \delta' p [\rho_0 + \Lambda(1 - \rho_0)] > 0,$$

where we use the notation $\partial^2 X_{ml} / (\partial m \partial l) \equiv (X_{11} - X_{10}) - (X_{01} - X_{00})$.

Proposition 1 (Complementarity between regulation and LOLR). *The marginal benefits of LOLR are increasing in regulation, both from the private perspective of the representative banker, and from the social perspective of total welfare.*

In fact, because the complementarities arising from the fiscal externalities of LOLR $\delta' p \Lambda (1 - \rho_0)$ add up to those arising from the taxes absorbing banker's surplus $\delta' p \rho_0$, we have $\partial^2 W_{ml} / (\partial m \partial l) > \partial^2 \tilde{U}_{ml} / (\partial m \partial l)$ so that the complementarities are stronger from a social than from a private perspective.

From now on, we will assume that the traditional banking system is optimal:[23]

Assumption 1. $W_{11} = \max_{\{m,l\}} W_{ml}$.

Corollary 1 (trilogy). *Under Assumption 1, the equilibrium configuration features banks lending to SMEs being regulated and enjoying LOLR.*

When indicated, we will also assume that it is not optimal to extend LOLR to the shadow banking sector:

Assumption 2. $W_{00} \geq W_{01}$.

That $W_{00} \geq W_{01}$ means that compared to the default point of a shadow banking sector with no LOLR, it is detrimental to extend LOLR to the shadow banking sector. A necessary and sufficient condition is that moral hazard be severe enough in the second dimension $\delta' > \pi_L [b + \beta - (1 - \rho_0)\Lambda] / [\rho_0 + (1 - \rho_0)\Lambda]$.

Under this parameter configuration the equilibrium features regulation and LOLR with $(m, l) = (1, 1)$. Extending LOLR is beneficial, but only coupled with regulation. Bankers enter a *quid pro quo* with the government whereby they accept being regulated

---

[23]That $W_{11} \geq W_{10}$ means that there is a further benefit from extending LOLR to the regulated sector. A necessary and sufficient condition is that the benefits of continuation for the bankers be high enough: $b > (1 - \rho_0)\Lambda - \beta$. Whether $W_{10} \gtrless W_{00}$ depends on whether compared to the default, it is beneficial to subsidize banks to operate under the regulatory umbrella. A necessary and sufficient condition is that the cost of regulation be low enough compared to the MSCPF $c < \delta(1 - p)[\rho_0 + (1 - \rho_0)\lambda]$.

in exchange for access to an LOLR policy that goes beyond what they could count on ex post with pure discretionary bailouts. This quid pro quo is a manifestation of the complementarities between regulation and LOLR highlighted above.

Note that the shadow banking sector then only acts as a constraint on the tax $\tau_{11} = \tilde{\tau}_{11} = \tilde{U}_{11} - \tilde{U}_{00}$ that the government can levy on bankers. The sign of this tax is not pinned down by these assumptions. A necessary and sufficient condition for $\tilde{\tau}_{11} > 0$ is that the benefits of continuation for the bankers be high enough $b > [\delta' p \rho_0 + c - \delta(1 - p)\rho_0]/p\pi_L$. When it is the case, bankers also contribute a fee in exchange for the access to LOLR that accrues to the regulated sector.

3 Special Depositors

Special depositors can be formalized in at least two ways and we explore both. First, in Section 3.1, following Gennaioli et al (2012, 2013), Stein (2012) and Caballero-Farhi (2017), we assume that special depositors are ex-ante risk averse (à la Epstein-Zin)—actually infinitely risk averse in the version below—at date 0; so they are willing to pay for the certainty of receiving their return for sure.[24] Second in Section 3.2, we model special depositors in a way similar to banks: special depositors have future projects (buying a house, sending children to college) for which they need cash at date 1 and so they are de facto ex-post risk averse. They lose the benefit of the project if their bank defaults on their savings and they are not bailed out by the government.

The two approaches deliver similar insights: complementarities between regulation and deposit insurance, resulting in economies of scope in regulation; quadrilogy in the sense that in equilibrium with optimal policy, banks are regulated, have access to deposit insurance and LOLR, and lend to SMEs. There are three main differences between the two approaches. First, the government may ex post spontaneously bail out special depositors in the second approach, but not the first. Second, in the second approach but not in the first, special depositors are willing to take some risk and so shadow banks may try to lure them through quasi-deposit offerings. Third, securities have clientele-dependent valuations because bailouts depend on the composition of ownership of each security in the second approach but not in the first.

---

[24]See also Malherbe-McMahon (2017) for a model with risk-averse households where only bank equity and deposits are traded, and where deposit insurance leads to risk-shifting.

3.1 Ex-Ante Risk Aversion: Deposit Insurance and Regulation Complementarities

We introduce, on top of the ordinary, risk-neutral depositors, a continuum of special depositors with mass μ. Each special depositor is willing to pay $1 + \theta$ for exactly one unit of safe deposits between dates 0 and 1. Importantly, these deposits must be absolutely safe in order to be valued.[25]

Since bankers experience liquidity shocks with some probability, they cannot supply safe deposits on their own. They must rely partly on the government. It may therefore be optimal for the government to run a separate deposit insurance scheme on top of its LOLR scheme. This deposit insurance scheme promises, in exchange for a fee which depends on bankers' choices, to make whole every special depositor who has deposited funds in a given bank in case this bank experiences a liquidity shock.

We alter the game described in Section 2 by adding a decision by the government at date 0 to charge a fee per unit of insured special deposit τ^{SD}_{ml} which is contingent on the choice (m, l) of each banker, as well as a decision by the bankers to post a price q at which they are willing to supply deposits. A unit of deposit insurance guarantees that the depositor is made whole at date 1 even when its bank is hit by a liquidity shock.

Finally and for conciseness, we assume that $\rho_0 \geq \mu$.[26] That is, in the absence of liquidity shock, the banks' pledgeable income suffices to cover the special deposits demand. Thus deposit insurance, like LOLR, optimally operates only when the banking sector is in distress.

As in the baseline model of Section 2, the government can induce any equilibrium configuration (m, l, d), where $d \in \{0, 1\}$ is an indicator variable denoting whether $(d = 1)$ or not $(d = 0)$ special depositors are served. It simply needs to charge, on top of the tax τ_{ml} computed in Section 2, a fee τ^{SD}_{ml} per unit of deposit which depends on (m, l). To induce the equilibrium configuration (m, l, d), if $d = 0$, it is optimal for the government to set the fee equal to infinity, and if $d = 1$, it is optimal for the government to set the fee equal to infinity except for (m, l) in which case it should be $\tilde{\tau}^{SD}_{ml}$, where $\tilde{\tau}^{SD}_{ml}$ is the insurance fee per unit of deposit that makes bankers indifferent between accepting and refusing special deposits.

---

[25]The exact microfoundations are as follows. Special depositors live for two periods, at dates 0 and 1. Their utility at date 0 is given by a form of recursive utility given by $c^{SD}_0 + (1 + \theta) \min_0 \{\min\{c^{SD}_1, 1\}\}$ where the notation $\min_0$ denotes a minimum over states of the world and $\theta > 0$. Their utility at date 1 is given by $\min\{c^{SD}_1, 1\}$. Special depositors are infinitely risk averse but have an infinite intertemporal elasticity of substitution up to a satiation point at which it becomes zero.

[26]When $\rho_0 > \mu$, banks also attract ordinary depositors. In fact, some banks may only attract ordinary depositors. It is also optimal to regulate these banks and to grant them access to LOLR.

Deposit insurance fees. We denote by q (to be determined below) the price of a unit of special deposit, that is the amount of money a banker collects at date 0 in exchange of a promise to pay whole one unit at date 1. We have

$$\tilde{\tau}_{ml}^{SD} = q - [\pi_H - (1-p)\delta] + m[-(1-p)\delta] + l[p\delta'] + ml[-p\delta'].$$

Suppose that the representative banker operates in the shadow banking sector without LOLR so that $(m,l) = (0,0)$. The benefit for the banker of attracting one unit of special deposit is the price q that he can charge for it. The cost is the expected repayment which is simply the probability $\pi_H - (1-p)\delta$ that the banker is not hit by a liquidity shock, since that money could have been pledged to ordinary depositors instead. In contrast, the repayment when the banker is hit by a liquidity shock is shifted to the taxpayer via the deposit insurance scheme. The deposit insurance fee $q - [\pi_H - (1-p)\delta]$ guaranteeing indifference extracts the net benefit. The other terms $-m[(1-p)\delta] + l[p\delta'] - ml[p\delta']$ simply reflect the changes in the net benefit for different values of (m,l). These changes come from changes in the cost via changes in the probability of liquidity shocks arising from moral hazard along the G and B dimensions. For example, if the representative banker operates in the shadow banking sector with LOLR so that $(m,l) = (0,1)$, then the probability of a liquidity shock is $\pi_L + (1-p)\delta + p\delta'$ instead of $\pi_L + (1-p)\delta$ since LOLR leads to moral hazard along the B dimension; and if the representative banker operates in the regulated sector with or without LOLR so that $(m,l) = (1,1)$ or $(m,l) = (1,0)$, then the probability of a liquidity shock is π_L since monitoring eliminates moral hazard along both dimensions.

Fiscal cost of deposit insurance. We can also compute the fiscal cost of deposit insurance for a mass μ of potential deposits and $d \in \{0,1\}$:

$$\mu d \left[[\pi_L[(1-p)\lambda + p\Lambda] + \delta(1-p)\lambda] + m[-\delta(1-p)\lambda] + l[\delta'p\Lambda] + ml[-\delta'p\Lambda] \right].$$

Suppose that the representative banker operates in the shadow banking sector without LOLR so that $(m,l) = (0,0)$. The fiscal cost of deposit insurance per unit of deposit is then $\pi_L[(1-p)\lambda + p\Lambda] + \delta(1-p)\lambda$. The other terms $m[-\delta(1-p)\lambda] + l[\delta'p\Lambda] + ml[-\delta'p\Lambda]$ reflect the changes in the fiscal cost for different values of (m,l). These changes come from changes in the cost via changes in the probability of liquidity shocks arising from moral hazard along the G and B dimensions.

Total welfare. We assume that the government puts a social welfare weight at most 1 on special depositors so that if special deposits are created through deposit insurance, then it is optimal for the government to set fees so as to extract all the surplus from special depositors, leading to $q = 1 + \theta$. Compared to the baseline model of Section 2, total social welfare must be adjusted to reflect the net fiscal cost of deposit insurance, which consists of the fiscal cost of deposit insurance net of the deposit insurance fees

$$W_{mld} = W_{ml} + \mu d \Big[[1 + \theta - [\pi_H - (1-p)\delta] - \pi_L[(1-p)\lambda + p\Lambda] - \delta(1-p)\lambda] $$
$$+ m[\delta(1-p)(\lambda - 1)] + l[-\delta' p(\Lambda - 1)] + ml[\delta' p(\Lambda - 1)] \Big].$$

Special Deposits and Regulation Complementarity. The complementarity between regulation (m) and LORL (l) identified above is unchanged. But now with three equilibrium dimensions (m, l, d) instead of two in the baseline model of Section 2, there are other patterns of complementarities and substituabilities across dimensions.

First, in the shadow banking sector ($m = 0$), LOLR and deposit insurance are substitutes, while in the regulated sector ($m = 1$), there are no interactions between LOLR and deposit insurance:

$$\frac{\partial^2 W_{0ld}}{\partial l \partial d} = -\delta' p(\Lambda - 1) < 0 \quad \text{and} \quad \frac{\partial^2 W_{1ld}}{\partial l \partial d} = 0.$$

This is because in the shadow banking sector, LOLR induces moral hazard along the B dimension, which increases the net fiscal cost of deposit insurance. In the regulated sector, LOLR does not result in any moral hazard and so the net fiscal cost of deposit insurance is independent of LOLR.

Second, whether ($l = 1$) or not ($l = 0$) there is LOLR, regulation and deposit insurance are complements:

$$\frac{\partial^2 W_{m0d}}{\partial m \partial d} = \delta(1-p)(\lambda - 1) > 0 \quad \text{and} \quad \frac{\partial^2 W_{m1d}}{\partial m \partial d} = \delta' p(\Lambda - 1) > 0.$$

This is because regulation mitigates moral hazard either along the B dimension or along the G dimension depending on whether or not LOLR is present, resulting in lower net fiscal costs of deposit insurance.

Proposition 2 (complementarity between supervision and deposit insurance). *Whether or not the representative bank is provided with LOLR assistance, the social benefit of deposit insurance is higher if the bank is supervised. In contrast, LOLR and deposit insurance are substitutes*

in the shadow banking sector and are independent in the regulated banking sector.

The complementarity between regulation and LOLR and that between regulation and deposit insurance are the signature of *economies of scope* in regulation: regulation facilitates both LOLR and deposit insurance.[27]

Corollary 2 (quadrilogy). *Under Assumption 1, a necessary and sufficient condition for the equilibrium configuration to feature regulation, LOLR, and deposit insurance* $[(m,l,d) = (1,1,1)]$ *is that the willingness to pay of special depositors be greater than the social cost of providing deposit insurance in the regulated sector:*

$$1 + \theta > \pi_H + \pi_L[(1 - p)\lambda + p\Lambda]. \tag{1}$$

Bankers then enter a *double quid pro quo* with the government whereby they accept being regulated in exchange for LOLR which allows them to weather liquidity shocks, as well as for deposit insurance which allows them to attract and service special depositors.

Observed Heterogeneity. Assume now that banks differ in how hard there are to monitor. For example, monitoring complex exposures on OTC markets is harder than assessing the risk on plain vanilla rated municipal bonds or even a (diversified) loan portfolio. One can imagine that there is a distribution of banks, each associated with a pattern of banking activities and characterized by its monitoring cost.[28] The monitoring cost is c for a fraction x of banks and $+\infty$ for the remaining fraction. The latter banks are necessarily in the shadow sector, while the former have a choice. We assume that the pledgeable income when banks are liquid is sufficient to repay deposits: $x\rho_0 \geq \mu$. So this extended version boils down to the previous one for $x = 1$. A bank's value of the monitoring cost is observable by the supervisor.

Heterogeneity of activities creates the possibility of co-existence of regulated and shadow banks in equilibrium.

Corollary 3 (heterogeneity). *Make Assumptions 1 and 2, and assume that (1) holds. Then, there are two sectors in the economy. A fraction x of the banks are regulated and enjoy deposit insurance and LOLR* ($m = l = d = 1$); *the remaining fraction* $1 - x$ *is in the shadow banking sector and has no public support* ($m = l = d = 0$); *special depositors are serviced by regulated banks.*

---

[27]This rationale for the co-existence of lending and deposit-taking is distinct from the one articulated by Kashyap et al. (2002). They also emphasize economies of scope but arising from a different mechanism: the need for a pool of safe and liquid assets for these two imperfectly correlated activities.

[28]Actually the banks can be ex-ante identical. What matters is that activities differ in their surveillance cost. One of the strengths of our modeling is that we don't presume that some banks cannot do certain things.

3.2 Ex-Post Risk Aversion: Financial Engineering, Safe Asset Creation, and Implicit Depositor Guarantees

So far, our framework created scope for bailouts solely of banks. In practice, governments may bring liquidity support and guarantees to funds that lie outside the traditional prudential sphere, but fail to deliver their promises to individual investors. A case in point is money market mutual funds that break the buck, that is, report a share value below a dollar. Individual investors have come to consider that such funds are the equivalent of safe banking savings accounts, but intermediaries do not have to pay for deposit insurance[29]. Another example is provided by life insurance, which promises a floor return equal to the initial investment, and even sometimes a strictly positive minimum return. Such investments pose a problem to authorities, as they may be tempted to bail out the investors in the same way they may bail out the banks. In China, wealth management products, usually sold by banks, are close substitutes for deposits and individual investors often presume that they are implicitly guaranteed; and trusts linked to or owned by banks raise money from businesses and individuals frustrated by the low cap imposed by the government on interest rates on bank deposits.

This section's insight is that the shadow banking sector may cleverly use financial engineering so as to attract special depositors and thereby create a put on taxpayer money. We first focus on the shadow banking sector, which will provide the most interesting insights of this section.

Pure shadow banking. Like in Section 3.1, there is a substantial mass of ordinary depositors, who are risk neutral and do not discount the future, and a mass μ of special depositors. Special depositors are risk neutral over $[0,1]$ at date 1 and have an investment opportunity for sure: they obtain $1 + \theta \in (\lambda, \Lambda)$ per unit of date-1 income, as long as this income does not exceed 1. They do not value date-1 revenues above 1. The condition $\theta > \lambda - 1 > 0$ implies that value may be destroyed if special depositors don't have 1 at date 1, while the condition $1 + \theta < \Lambda$ guarantees that there is no gain from trade in bailing out special depositors in state of nature B.

Thus, if special depositors invest q in a portfolio with a contingent return at most 1 in each state of nature and expected payoff Z, their valuation of the portfolio is $(1 + \theta)Z - q$, while the ordinary depositors value the same portfolio at $Z - q$. Without loss of generality we will consider only portfolios that deliver at most 1 in each contingency.

Later on, we will distinguish for each portfolio k the private fundamental, X_k and the

---

[29] Recent regulatory reforms in the US aim at addressing this presumption of implicit guarantee.

public bailout, Y_k so that $Z_k = X_k + Y_k$. Let $(\mu_k, 1 - \mu_k)$ denote the fractions of special depositors and ordinary depositors holding in equilibrium asset or portfolio k.

The social planner puts weight 1 on depositors, special or ordinary. The condition for a bailout of depositors holding portfolio k at date 1 in state of nature G is

$$\mu_k(1 + \theta) + (1 - \mu_k) \geq \lambda \Leftrightarrow \mu_k \geq \mu^*.$$

If this condition is satisfied, then the government at date 1 makes up for the shortfall in return to bring the return to 1.

There are three states of liquidity for (here shadow) banks (see Figure 4). With probability π_H, banks have no liquidity needs; for convenience, we locate moral hazard in that state of nature: shirking generates a liquidity need equal to 1 with probability δ in state G and δ' in state B; so the probability of no illiquidity becomes $\pi_H - \delta$ (respectively $\pi_H - \delta'$). Let $\pi_1 \equiv \pi_H - (1 - p)\delta$. With probability $\pi_L x$, all banks are short of liquidity: they have no date-1 revenue and need 1. Let $\pi_2 \equiv \pi_L x + (1 - p)\delta$. With probability $\pi_3 \equiv \pi_L(1 - x)$, banks need to reinvest an amount $\rho_0 - r$, so they have free cash flow r; we assume that $\mu < r < \rho_0 < 1$. Note that for $x = 1$, the stochastic structure coincides with that of Section 2.
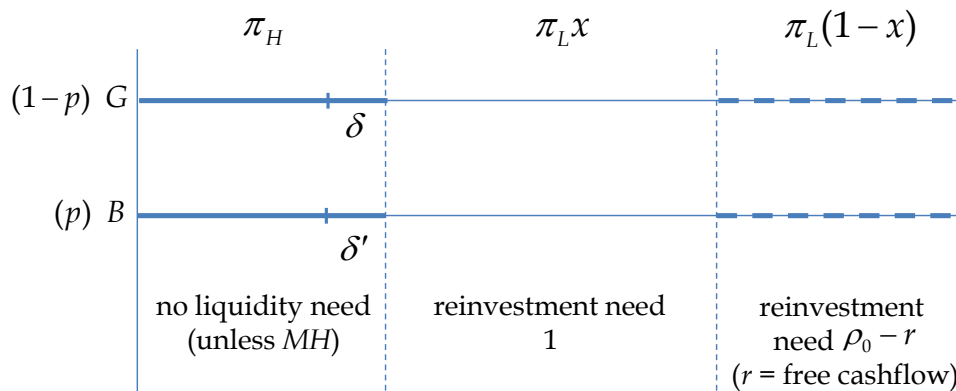
$$\pi_1 \equiv \pi_H - (1-p)\delta \quad \text{and} \quad \pi_3 \equiv \pi_L(1-x)$$

Figure 4: States of nature.

The timing goes as follows: (1) Securities are put on the market. (2) Investors (special depositors and ordinary depositors) bid for individual securities- truthfully so as they are infinitesimal. (3) The auctioneer allocates securities so as to maximize gains from trade.[30] Investors pay the market clearing price.

---

[30]This will yield a competitive equilibrium. It is the equivalent of least-cost-dispatch in day-ahead electricity generation auctions. Competitive Treasury bill auctions also allocate bills to the highest bidder.

Endogenous sorting and clientele-dependent valuations. An interesting property of this model of asset valuation is that the sorting condition is endogenous and that asset values are determined by a fixed-point reasoning, with a possible multiplicity of equilibrium values. The yield on a portfolio k does not depend solely on the fundamental X_k, but also on the bailout option Y_k. The existence of a bailout depends on investor composition, which itself depends on total return Z_k and therefore on the existence of a bailout. So valuations depend on the assets' clienteles, which depend on valuations.[31]

To see that the sorting condition is endogenous, take two portfolios, 1 and 2, such that $X_1 > X_2$, but $Z_1 = X_1 < Z_2 = (1-p) + pX_2$; that is, holders of portfolio 2 are bailed out in state G, but holders of portfolio 1 are not. In the absence of bailout, special depositors prefer portfolio 1 to portfolio 2 whenever ordinary depositors do also; in the presence of bailout, special depositors prefer portfolio 2 to portfolio 1 whenever ordinary depositors do also. This suggests that multiple valuation equilibria will in general coexist. One could even envision the possibility of "disembodied bailouts", in which the government would rescue investors who hold an intrinsically valueless asset. Because of the government guarantee, such an asset would be attractive to special depositors (if p is small enough), which in turn would trigger a bailout[32].

To the best of our knowledge, equilibria with investor bailouts and endogenous sorting have not been studied in the economics literature. To avoid such reversals of the sorting condition, we impose the following monotonicity condition: If $X_1 > X_2$, then $\mu_1 \geq \mu_2$.[33] Then intrinsically safer portfolios are more attractive to special depositors and therefore more likely to trigger a bailout, so $Z_1 > Z_2$.

An equilibrium with implicit guarantees. We can now describe the following laissez-faire equilibrium. Let the financial system re-package banks' debt to create the safest possible fundamental: The corresponding security yields exactly 1 both when banks have no liquidity needs and when they do but they have free cash flow. The strategy of repackaging to create relatively safe assets captures the essence of shadow banking institutions

---

[31]In particular, two identical securities may be priced differently contingent on the realization of a sunspot: one held by ordinary depositors and benefiting from no implicit guarantee; and another with a sufficient number of special depositors among its holders and fetching a much higher price.

[32] There are instances of disembodied bailouts, but primarily outside the financial realm. Take the case of farmers in France. They by and large do not hedge in futures markets, and expect a government bailout if prices drop substantially. This bailout is limited as the government is trying to address the most pressing needs (the "special farmers" are here the most destitute or the most indebted). A similar observation applies to minimum retirement income provided by the government (again it is set at very low levels, so as to address the most pressing needs).

[33]We impose this on -and off- the equilibrium path (security 1 or 2 or both might not be purchased).

such as money market mutual funds. Here the highest fundamental is

$$X^* = (1 - p)\pi_1 + p\pi_H + \pi_3.$$

The maximal number of such securities is r. Assume further that

$$\frac{\mu}{r} \geq \mu^*.$$

In equilibrium, all special depositors buy this security at price equal to

$$q^* = (1 - p) + p(\pi_H + \pi_3),$$

because the marginal investor is an ordinary depositor, and the ownership share of these securities vindicates an investor bailout in state G. No other security benefits from an implicit guarantee, as it is purchased by ordinary depositors (and priced at its intrinsic value).

Can financial intermediaries deviate and create a larger number of securities that trigger depositor bailouts? To do so requires pledging less than $X < X^*$. The corresponding securities will not be selected by the special depositors, whose pecking order is to first select securities with the highest fundamental, $X = X^*$, which are in excess supply when it comes to addressing the needs of special depositors. The new securities therefore will attract only ordinary depositors, and will not benefit from a depositor bailout in state G. Financial engineering to attract ordinary depositors does not raise profit.

Optimal Regulation. The possibility of implicit investor guarantee in the shadow banking sector does not affect our analysis, except for the reservation utilities. In the equilibrium in the shadow banking sector described above, all securities are priced by ordinary depositors.[34] Special depositors' utility in the shadow banking sector is increased by implicit guarantees. And so the government must also be friendlier to depositors when offering deposit insurance: deposits are no longer that cheap.

Laissez faire is suboptimal for three reasons over and above those that were analyzed in the basic model of Section 2, though. First, in the equilibrium above, some ordinary depositors- with mass $r - \mu$, benefit from an investor bailout in state G (and pay for it). With probability $(1 - p)(\pi_L x + \delta)$, they receive 1 at cost λ for the state. So the equilibrium involves a deadweight loss $(r - \mu)(1 - p)(\pi_L x + \delta)(\lambda - 1)$. Second, moral hazard is costly as it increases the probability of illiquidity in state G by $(1 - p)\delta$ and hence of investor

---

[34]We have $U^{\text{pure shadow banking}} = A - 1 + [[\pi_H - (1 - p)\delta]\rho_0 + r\pi_3] + r(1 - p)(\pi_L x + \delta) + b(1 - p\pi_L)$.

bailouts (on top of bank bailouts); therefore, and as earlier, the state wants to bring the banks within the regulated sector. Third, the shadow banks enjoy a large rent. They sell the safe asset at a price that reflects the prospect of an investor bailout by the state in state of nature G.

All three problems can be resolved in the following way, that maximizes welfare. First, regulated banks are monitored, eliminating moral hazard, and benefit from LOLR. Second, the state purges the unregulated investor market from special depositors by attracting them in the regulated sector. So only ordinary depositors remain in the unregulated market and the shadow banking sector does not benefit from investor bailouts. Third, special depositors are attracted in the regulated sector with deposit insurance. The regulated banks offer to those who choose it (the special depositors) the following explicit deposit insurance: They are fully covered in state G and are covered in state B with probability $\pi_H + \pi_3$. To make sure that special depositors are attracted by this offer rather than an alternative offer by the shadow banking sector, the state can offer a bit of extra deposit insurance in state B, so that special depositors are then bailed out with probability $(1 - p)\pi_L x$.[35]

The reservation utility of bankers (and hence their equilibrium utility) is determined by the outside option of shadow banking and is the same as in Section 2.[36] Unlike in Section 3.1, special depositors are able to capture a rent $[(1 - p) + p(\pi_H + \pi_3)](\theta - 1)$ because the securities that they purchase are priced by ordinary depositors who are the marginal investors. This rent is irrelevant for welfare because they have a welfare weight of 1. Ordinary depositors do not obtain any rent.

As in Section 3.1, we can extend the model by assuming that some banks cannot be regulated, ensuring the co-existence of a regulated sector and of a shadow banking sector. But special depositors locate in the regulated sector as they are attracted there by deposit insurance.

Overall, as in Section 3.1, we still obtain complementarities between regulation and LOLR and between regulation and deposit insurance, resulting in economies of scope in regulation. But now the shadow banking sector has the possibility of attracting special depositors, and it is only the extension of deposit insurance in the regulated sector at the appropriate price that prevents their migration in the shadow banking sector. Moreover,

---

[35]The state offers deposit insurance with probability ϵ in state B when it is needed, where ϵ is arbitrarily small. Special depositors are covered in state B with probability $\pi_H + \pi_3 + \epsilon$. The total cumulative probability of a bailout of special depositors in states G and B is $(1 - p)\pi_L x + p\epsilon$.

[36]This is true up to an additive constant which arises from the slight model modification giving bankers extra liquidity r in some state, which increases their utility in all equilibria by $\pi_3 r$. Note that unlike earlier, a bank that deviates in the shadow banking sector get less than its utility in the pure shadow banking equilibrium: $U_{00} = U^{\text{pure shadow banking}} - r(1 - p)(\pi_L x + \delta)$.

special depositors enjoy some rents.[37]

Proposition 3 (complementarity supervision/ deposit insurance and quadrilogy). *Under ex-post risk aversion for special depositors, deposit insurance is offered only in the regulated sector to attract special depositors and thereby economize on investor bailouts and rents to banks. The quadrilogy obtains. In contrast to the case of special depositors with ex-ante risk aversion, the government may ex post spontaneously bail out special depositors, special depositors are willing to take some risks, securities have clientele-dependent valuations because bailouts depend on the composition of ownership of each security, and shadow banks use financial engineering to try to woo special depositors.*

4   Contagion, Liquidity Syphoning, and Ring Fencing

So far liquidity shocks were perfectly correlated and so there was no rationale for liquidity pooling and therefore for counterparty risk. In practice, liquidity pooling occurs through credit default swaps, interest and FX swaps, lines of credit, guarantees, money market instruments, and other varieties of financial instruments. To capture liquidity pooling and its regulatory consequences, we allow liquidity shocks to be imperfectly correlated. We assume that banks are able to recognize the patterns of correlation. The regulator cannot assess the correlation when one of the two counterparties is in the shadow banking sector: Figuring out the correlation (which requires knowing the two types) requires at the very least the supervision of both balance sheets. When both counterparties are regulated, we look at the polar cases in which the regulation learns (say, through joint stress tests) or does not learn the pattern of correlation of the two institutions. Imperfect regulatory knowledge will create opportunities for gaming, whose consequences for prudential supervision we examine.

4.1   Setup and Optimal Liquidity Sharing

Setup.   For expositional simplicity, there is no date-2 pledgeable income ($\rho_0 = 0$), and no special depositors. The model is otherwise an extension of the perfect-correlation basic model. We interpret liquidity shocks in the following way: each bank receives a date-1 revenue equal to 1 or 0, and has a reinvestment need of 1 or 0 at date 1. We can thus

---

[37]Making deposit insurance in the regulated sector too expensive can trigger a migration of special depositors to the shadow banking sector. Similarly, discouraging banks from lending to SMEs can encourage the latter to enter relationship banking with shadow banks. Both were probably at work in China and in the U.S. in recent years.

assume that at date 1, there are four possible types of banks: $(0,1)$ for a bank with no liquidity and a reinvestment need; $(1,1)$ for a bank with liquidity and a reinvestment need; $(1,0)$ for a bank with liquidity and no reinvestment need; and $(0,0)$ for a bank with no liquidity and no reinvestment need. There are three aggregate states of nature within each aggregate state G and B:

- "No Illiquidity" (NI): With probability $\pi_H - \pi$, banks have no reinvestment need and no liquidity so that they are all of ex-post type $(0,0)$. If a bank shirks, then with probability $\pi_H - \pi - \delta$ (resp. $\pi_H - \pi - \delta'$), it has no reinvestment need and no liquidity so that it is still of ex-post type $(0,0)$, and with probability δ (resp. δ'), it has a reinvestment need and no liquidity so that it is of ex-post type $(0,1)$ if the aggregate state is G (resp. B).

- "Widespread Illiquidity" (WI): With probability π_L, banks have a reinvestment need and no liquidity so that they are all of ex-post type $(0,1)$.

- "Insurance Opportunity" (IO): With probability π, there are four equiprobable possible aggregate unverifiable sub-states in which exactly one quarter of the banks are of each of the four possible ex-post types. We assume that there are banks are equidistributed in $4! = 24$ different ex-ante types (known at date 0) of banks corresponding to the different possible assignments of the four possible ex-post types to the equiprobable aggregate sub-states. We denote the ex-ante type of a bank according to its ex-post types in the different aggregate sub-states. For example a bank of ex-ante type $[(0,1),(1,0),(0,0),(1,1)]$ is of ex-post type $(0,1)$ in the first aggregate sub-state, $(1,0)$ in the second one, $(0,0)$ in the third one, and $(1,1)$ in the last one. A bank and its potential banking counterparties know its type at date 0.

When $\pi = 0$, the model is strictly isomorphic to our baseline model in which there is no scope for liquidity sharing. When $\pi > 0$, there are opportunities for liquidity sharing between the different ex-ante types of banks in the "insurance opportunity" state. We will study different liquidity sharing arrangements. Throughout, we maintain the assumption that any extra liquidity that remains in a bank at date 1 after liquidity has been shared and investments have been made can be pledged to investors at date 0 and the corresponding revenues used to increase the consumption of the banker at date 0.

Optimal liquidity sharing. From now on, we assume that $\pi > 0$. In each of the aggregate sub-states within the IO state, there is exactly $1/2$ units of aggregate liquidity, and $1/2$ units of reinvestment need, and so there is sufficient aggregate liquidity, but some

banks have a net liquidity deficit while others have surplus. The most socially efficient way to share this liquidity is to distribute it so that each bank with a unit of reinvestment need has one unit of liquidity on hand. This ensures that all reinvestment needs in the IO state are met at the least possible fiscal cost (equal to zero). In effect, it suffices to share liquidity in pairs of banks which "natural insurance counterparties" in the two aggregate states of the world G and B, where two banks of are said to be natural insurance counterparties if whenever the ex-post type of one of the bank is $(0,1)$, that of the other one is $(1,0)$. For example, two ex-ante types are natural insurance counterparties for the ex-ante type $[(0,1),(1,0),(0,0),(1,1)]$: $[(1,0),(0,1),(0,0),(1,1)]$ and $[(1,0),(0,1),(1,1),(0,0)]$.

The analysis is then similar that of the baseline model. There are some subtle differences, which we highlight by using hats to denote utility and welfare values in the augmented model.

First, no matter what the original equilibrium (m,l) is, the binding off-equilibrium migration to the shadow banking sector is a collective deviation involving any number of pairs of natural insurance counterparties. These natural insurance counterparties then share liquidity in the aggregate state (B, IO), and not in the other states—in particular, they do not want to co-insure themselves in the aggregate state (G, IO) since this would imply losing their put on taxpayer money.

Second, $\hat{\tilde{U}}_{00} = \tilde{U}_{00} + (1-p)\frac{\pi}{4}$ since banks in the shadow banking sector are left with extra liquidity if their ex-post type is $(1,0)$ in aggregate state G.[38] In addition, we have $\hat{\tilde{U}}_{01} = \tilde{U}_{01} + \frac{\pi}{4}$ since banks in the shadow banking sector are left with extra liquidity if their ex-post type is $(1,0)$ in aggregate states G and B. And for $(m,l) \in \{(1,1),(1,0)\}$, we have $\hat{\tilde{U}}_{ml} = \tilde{U}_{ml}$. This in turn implies that $\hat{\tilde{\tau}}_{01} = \tilde{\tau}_{01} + p\frac{\pi}{4}$ and $\hat{\tilde{\tau}}_{ml} = \tilde{\tau}_{ml} - (1-p)\frac{\pi}{4}$.

Third, $\hat{W}_{00} = W_{00} - (1-p)\frac{\pi}{4}\lambda$ because banks in the shadow banking sector need a bailout with a higher probability in the G state. In addition $\hat{W}_{01} = W_{01} - [(1-p)\lambda + p(\Lambda - 1)]\frac{\pi}{4}$ because banks in the shadow banking sector need a bailout with a higher probability in the G state and more LOLR in the B state, and because the before-tax utility is higher relative to the outside option. And finally, for $(m,l) \in \{(1,1),(1,0)\}$, we have $\hat{W}_{ml} = W_{ml} - (1-p)\frac{\pi}{4}$ because the outside option is higher. Since we have assumed that $W_{11} = \max_{m,l}\{W_{ml}\}$ and $W_{00} > W_{01}$ in the baseline model, this remains true in the augmented model so that $\hat{W}_{11} = \max_{m,l}\{\hat{W}_{ml}\}$ and $\hat{W}_{00} > \hat{W}_{01}$.

---

[38]For simplicity, we conduct all our analysis under the simplifying assumption that intermediate liquidity positions are impossible: ex post, a bank must have a either 0 or 1 unit of liquidity. These indivisibilities rule out the possibility that a shadow bank secures just enough liquidity to be bailed out in some sub-states in aggregate state B. Dispensing with this assumption would complicate the analysis since it would require the characterization of the equilibrium liquidity sharing arrangement in the shadow banking sector. It would result in a higher value of the outside option $\hat{\tilde{U}}_{00}$, a reduction of $\hat{W}_{ml}$ for $(m,l) \neq (0,0)$ by the same amount, and a larger reduction in $\hat{W}_{00}$.

4.2 Implementing Optimal Liquidity Sharing: Mutual Credit Lines

Credit lines. We now turn to the decentralization of this socially efficient outcome. We show that it can be implemented through a set of credit lines. A credit line is a (possibly state-contingent) call option to borrow a given amount of money from the counterparty at a prespecified rate of interest. We assume that a banker cannot exercise a credit line and swindle the money, which thus has to be re-invested in the bank; this prevents a bank with sufficient liquidity to nonetheless call on a credit line and consume the money. This allows us to capture actual institutions in a simple manner.

We introduce the following financial contract: A credit line allows its beneficiary to draw 1 unit upon request in exchange of a corresponding IOU. Similarly, lending by the state under a bailout or through LOLR involves an IOU (which is also the case in practice). At date 1, each bank observes the revenue of a counterparty (the grantor) from which it may receive money through a credit line, and can threaten to take the credit line grantor into a bankruptcy proceeding if the latter does not abide by its commitment; but if the grantor has no revenue, there is no bankruptcy proceeding (to ensure that, one may envision a small cost of bankruptcy proceedings). Date 1 has three sub-periods as summarized in Figure 5:

- 1.1 Each bank receives its revenue (0 or 1). It can then call on the private credit line or the LOLR assistance it has secured at date 0. In either case, it thereby contracts a loan that it will have to reimburse at date 1.3.

- 1.2 If its reinvestment need, if any, is covered by available resources, the bank reinvests. If there is a reinvestment need and no sufficient resources, then the bank attempts to be bailed out by the state; if the request is granted, the bank reinvests; it the request is denied, its activity stops.

- 1.3 The bank reimburses its loan, if any. If it does not have resources, the debt is forgiven (the debtor is indifferent, but he would prefer continuation if there were an epsilon pledgeable income or an epsilon cost of legal action), and the bank operates until date 2.

We suppose that these credit lines can be made contingent on the aggregate states G and B.[39]

When banks that are natural insurance counterparties grant each other such lines of credit, the occurrence of bank distress is minimized, allowing the state to economize on

---

[39]Recall that the sub-states are unverifiable. On the other hand the ability to contract on sub-states would not affect gaming opportunities.

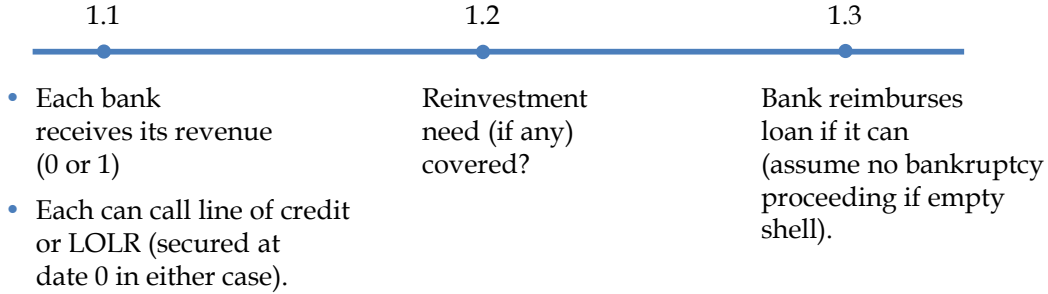| 1.1	1.2	1.3
• Each bank receives its revenue (0 or 1) • Each can call line of credit or LOLR (secured at date 0 in either case).	Reinvestment need (if any) covered?	Bank reimburses loan if it can (assume no bankruptcy proceeding if empty shell).

Figure 5: Timing within date 1.

taxpayer money (for a bailout in state G and a LOLR rescue in state B). Thus, to reach maximal welfare, the regulator must not only induce banks to join the regulated sector and give them access to LOLR (under Assumption 1), but also ensure that the banks are hedged through proper risk transfer schemes. This raises the issue of potential gaming of the hedging function, to which we now turn.

Implementation of optimal liquidity sharing and the necessity of ring fencing. We first suppose that the regulator can observe at date 0 the ex-ante type of each bank *in the regulated sector*, as well as their portfolios of credit lines. Banks can observe at date 0 the ex-ante type of another bank in the shadow banking sector, but the regulator cannot in the sense that a bank in the shadow banking sector can portray itself to the regulator to be of any possible ex-ante type at date 0.

Implicit in the optimal allocation is a form of ring fencing of liquidity which makes sure that banks in the regulated sector do not trade liquidity with banks in the shadow banking sector. Such ring fencing is necessary. Indeed, without ring-fencing, the equilibrium would unravel.

Consider, starting from an equilibrium with banks in the regulated sector and LOLR $((m,l) = (1,1))$, an individual deviation by a bank of ex-ante type $[(0,1),(1,0),(0,0),(1,1)]$. This deviating bank could join the shadow banking sector and sign a mutual credit line arrangement for state B with a regulated bank of ex-ante type with first element in the set $\{(1,0),(1,1)\}$, so that it receives one unit of liquidity when it has a reinvestment need but no liquidity (when its ex-post type is $(0,1)$), and with second and fourth elements in the set $\{(0,0),(1,0),(1,1)\}$, so that it does not send net liquidity when it has liquidity, whether or not it has a reinvestment need (when its ex-post type is $(1,0)$ or $(1,1)$). The regulated bank would be indifferent between signing this mutual credit line arrangement (with no ex-ante transfer) and signing a mutual credit line arrangement with a natural insurance counterparty in the regulated sector. The shadow bank would achieve a utility

level higher than $\hat{U}_{00}$ by an amount $p\frac{\pi}{4}$.

The shadow bank is better off because it receives liquidity from the regulated bank whenever it needs it in the B state when it cannot count on a bailout, but never sends liquidity to the regulated bank: The credit line granted by the regulated bank is legitimate whereas that granted by the shadow bank is "bogus". The reason the regulated bank is indifferent is that it benefits from LOLR and therefore is fully insulated from the associated liquidity loss. Overall, liquidity is "syphoned away" from the regulated sector to the shadow banking sector which piggy-backs on LOLR. Because it raises the outside option, this marginal deviation therefore upsets the original equilibrium, and leads to a lower welfare value for $\hat{W}_{11}$ by an amount $p\frac{\pi}{4}$.[40] We collect all these results in a proposition.

Proposition 4 (Mutual Credit Lines and Ring Fencing). *The socially efficient arrangement, which has all banks in the regulated sector and, under Assumption 1, enjoying access to LOLR, can be decentralized in the following way if the regulator can monitor correlations within the regulated sector. The regulator forces each bank to sign mutual credit lines with one of its "natural insurance counterparties" in the regulated sector in the two aggregate states of the world G and B. Mandating mutual credit lines by itself is not sufficient, ring fencing is required.*

4.3 Monitoring Correlations Within the Regulated Sector

The necessity of monitoring correlations within the regulated sector. An important assumption supporting the implementation of optimal liquidity sharing via mutual credit lines described above is that the regulator can observe the ex-ante bank type of a regulated bank and hence monitor the correlations and thereby understand the exposures created by credit lines. Suppose that we do away with this assumption but maintain the ring fencing assumption. Then in an equilibrium in the regulated sector and LOLR with $(m,l) = (1,1)$, the only stable pairings would be: for a regulated bank of ex-ante type $[(0,1),(1,0),(0,0),(1,1)]$ to sign mutual credit line arrangements with a regulated bank of any ex-ante type with first element in the set $\{(0,0),(0,1),(1,1)\}$, so that it does not receive liquidity when it has a reinvestment need but no liquidity (when its ex-post type is $(0,1)$), and with second element in the set $\{(0,0),(1,0),(1,1)\}$, so that it does not send liquidity when it has liquidity, whether or not it has a reinvestment need (when its ex-post type is $(1,0)$ or $(1,1)$), both in states G and B (with no ex-ante transfer). In other

---

[40]This result shows that ring fencing is required. We do not attempt to solve equilibrium welfare without ring fencing, which requires analyzing all possible collective deviations, in a context where the price of credit lines is endogenous to the size and composition of the deviating coalition. We content ourselves with noting that the analysis above holds as long as the shadow banks form a minority and therefore are on the short side of the insurance market, reaping the gains from trade.

words, ring fencing gets rid of liquidity syphoning since liquidity is only redistributed within the regulated sector, but does not eliminate the breakdown of liquidity insurance via bogus credit lines within the regulated sector, for which monitoring of correlations within the regulated sector is required. This would lead to a lower welfare value for $\hat{W}_{11}$ by an amount $[(1-p)(\lambda-1)+p(\Lambda-1)]\frac{\pi}{4}$.

Central counterparty clearing houses. If correlations cannot be monitored in the regulated sector, an alternative arrangement to implement optimal liquidity sharing is to set up a central counterparty clearing house (CCP). Banks in the regulated sector are forced to participate in the CCP and banks in the shadow banking sector are banned from participating. The CCP grants an unconditional (for both aggregate states G and B) credit line to all banks in the regulated sector and banks in the regulated sector grant an unconditional credit lines to the CCP. This guarantees the efficient distribution of liquidity within the regulated sector and blocks any syphoning of liquidity to the shadow banking sector by preventing banks from fine-tuning their liquidity provision at the expense of the taxpayer.[41]

Proposition 5 (CCPs). *The implementation of the socially efficient arrangement with mutual credit lines with natural insurance counterparties of Proposition 4 requires the monitoring of correlations within the regulated sector. If correlations cannot be observed, an alternative implementation can be used: mandating participation in CCP in the regulated sector and banning participation from the shadow banking sector.*

4.4 Heterogeneity and coexistence of the two sectors

The arguments above apply in a model where a representative fraction $1-x$ of banks have a monitoring cost equal to ∞. Banks with infinite monitoring cost operate in the shadow banking sector.

We continue to denote by (m,l) the configuration applying to banks with finite monitoring cost c. Under the same assumptions as above, in the second best: banks with finite monitoring costs are monitored in the regulated sector, benefit from LOLR, and pool their liquidity among natural insurance counterparties in aggregate states G and B; banks with infinite monitoring costs are not monitored in the shadow banking sector, do not benefit from LOLR, and pool their liquidity among natural insurance counterparties only in aggregate state B.

---

[41]With the mandatory use of a CCP, do we still need ring fencing? Abandoning ring fencing could make the CCP solution collapse if "rogue" shadow banks (outside of the model) show up with a very different risk profile that makes them predominantly receivers of liquidity.

As in Proposition 4, the underlying liquidity arrangement can be implemented by mandating regulated banks to sign mutual credit lines between natural insurance counterparties within the regulated sector. This requires both ring fencing and the monitoring of correlations within the regulated sector. There are a few interesting differences. For example, it is interesting to note that there are additional costs from relaxing ring fencing since the associated liquidity syphoning via bogus credit lines decreases welfare not only by increasing the outside option of operating in the shadow banking sector[42], but also by directly increasing the fiscal costs of LOLR in the regulated sector. Alternatively, as in Proposition 5, the underlying liquidity arrangement can be implemented with a CCP.

4.5 Shadow Banking as Diversification and the Costs of Liquidity Segregation

We now modify the setup to point out a potential cost of segregating liquidity across sectors. Consider the model with a fraction $1 - x$ of banks with an infinite regulation cost. We modify the stochastic structure of the economy, only in state IO. We assume that the ex-ante type of a bank is perfectly correlated with its monitoring cost. Banks with a finite monitoring cost are all of ex-ante type $[(0,1),(1,0),(0,0),(1,1)]$, and banks with an infinite monitoring cost are all of ex-ante type $[(1,0),(0,1),(0,0),(1,1)]$.

This means that there is no scope for liquidity sharing neither within the regulated sector nor within the shadow banking sector. Liquidity sharing can only be implemented across the regulated and shadow banking sectors via cross exposures. When $x < 1/2$, banks with finite monitoring costs are on the short side of the liquidity market. In this case, there is no change in the outside option of operating in the shadow banking sector. Mandating liquidity sharing via mutual credit lines increases welfare by reducing the fiscal cost of bailouts in both sectors and of LOLR in the regulated sector. When $x > 1/2$, there is a tradeoff: Mandating liquidity sharing via mutual credit lines (with potential ex-ante transfers) across sectors reduces welfare by increasing the outside option of operating in the shadow banking sector on the one hand, but increases welfare by reducing the fiscal cost of bailouts in both sectors and of LOLR in the regulated sector.

---

[42]As long as $1 - x < 1/2$, we can assume that the banks with infinite monitoring cost reap all the gains from liquidity sharing if the underlying trades are possible since they are on the short side of the liquidity market.

5 Discussion

Comparative advantage in supervision. So far monitoring has been assumed to be perfectly contractible, and we have been agnostic as to whether it was best performed by a public and private entity. In practice, monitoring is performed both by the public sector (banking supervision) and by the private sector (holders of shares and bailinable securities, rating agencies), and in both cases it is potentially subject to moral hazard and capture. So it may be useful to look at the stakes. The following is not a full treatment, but it indicates some alleys for research on the political economy of monitoring.

Under shadow banking, the stake in monitoring for the bailinable claims is $(1 - p)\delta\rho_0$ while that for the state is $(1 - p)\delta(1 - \rho_0)\lambda$. So depending on the parameters, the private or the public sector may have a higher incentive in monitoring. By contrast, if the bank is financed by retail depositors protected by deposit insurance, the state's stake is higher and the investors' stake lower (actually equal to 0 if the financing is entirely through retail deposits). LOLR in turn further increases the state's stake. Together with the complementarities unveiled in Propositions 1 and 2, the case for public supervision seems particularly strong for traditional banking.

Bailing out the bank or its counterparties? A nagging question faced by all models of contagion is that of why shadow financial institutions such as AIG or the large investment banks were rescued in 2008, despite the fact that they had no retail depositors and by and large were not lending to SMEs. The standard answer to this is that authorities were concerned about systemic risk, coming in two guises: regulated financial institutions' exposure to the distressed entities and the possibility of fire sales. But such arguments for "indirect rescues" rather than direct ones are not fully convincing. For example, AIG could have been allowed to default and its regulated counterparties could have received assistance from the government. Somehow, it must be the case that the government has imperfect information about the real exposures (say, implied by complex derivative products). Our model, like other existing models, does not allow for opaqueness obviating the direct rescue of regulated institutions. This is an open and important line for future research.

Universal banks and regulating institutions vs. activities. We have not analyzed the question of banks involved in different activities, some traditionally thought of as belonging to the regulated sector and some traditionally thought of as belonging to the shadow banking sector. An interesting question is whether regulation should be performed at the

activity level or at the institution level. Our model suggests an argument for regulating institutions rather than activities to the extent that liquidity can be reallocated more easily inside universal banks than through arms-length financial transactions. Opening-up the black box of financial institutions and tackling the question of firm boundaries is an important area where future research will be needed.[43]

6 Conclusion

We studied the optimal regulation of banks when supervision reduces moral hazard and the riskiness of balance sheets and financial intermediaries can migrate to shadow banking in response to regulatory requirements. We did not posit that shadow or retail banks had a comparative advantage, and rather derived differences in their behavior from equilibrium considerations. The first key insight is the complementarity between regulation and the forms of insurance provided by the state: LOLR to banks and deposit insurance to depositors. Insurance is costly and supervision helps reduce the risk that its promises are called upon. Our analysis makes room for both bank and depositor implicit and explicit guarantees, and introduces the new concept of clientele-dependent valuation of bank securities. Second, we provide the first formal rationale for ring fencing and for incentivizing the migration of transactions towards CCPs. To this purpose, we showed how imperfect regulatory information may lead to gaming of hedging among financial intermediaries, resulting in banks being only partially covered as they hoard "bogus liquidity" and in public liquidity being syphoned off to the shadow sector. Overall the picture emerging from the analysis is an hexalogy: prudential supervision of banking goes hand in hand with servicing special borrowers (SMEs) and special lenders (retail depositors), LOLR, deposit insurance, incentivized migration to CCPs and ring fencing.

Needless to say, this sharp picture is only meant to stress natural covariations. Reality as always is more complex than the model suggests. The unique features associated with the traditional banking sector themselves impose costs, leading to a finer overall picture. We hope that future work will sharpen this analysis.

---

[43]First attempts at studying the costs and benefits of separation of traditional lending and investment banking and therefore the merits of universal banking are provided by Shy and Stenbacka (2017) and Vickers (2017).

References

Acharya, V., and M. Richardson, eds., (2009) *Restoring Financial Stability: How to Repair a Failed System*, John Wiley and Sons.

Acharya, V., Schnabl, P., and G. Suarez (2013) "Securitization without Risk Transfer," *Journal of Financial Economics*, 107(3): 515–536.

Beguenau, J. and T. Landvoigt (2017) "Financial Regulation in a Quantiative Model of the Modern Banking Sytem", mimeo Stanford.

Bengui, J. and J. Bianchi (2014) "Capital Flow Management when Capital Controls Leak," Minneapolis Fed Working Paper.

Berg, T., and J. Gider (2016) "What Explains the Difference in Leverage between Banks and Non-Banks?" *Journal of Financial and Quantitative Analysis*(forthcoming).

Buchak, G., Matvos, G., Piskorski, T., and A. Seru (2017) "Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks," Chicago Booth Working Paper.

Caballero, R.J., and E. Farhi (2017) "The Safety Trap," forthcoming *Review of Economic Studies*.

Chrétien, E., and V. Lyonnet (2017) "Traditional and Shadow Banks", HEC and INSEE CREST mimeo.

Claessens, S., Z. Pozsar, L. Ratnovski and M. Singh (2012) "Shadow Banking: Economics and Policy," IMF Staff Discussion Note, 12/12, http://dx.doi.org/10.2139/ssrn.2187661.

Di Iasio, G. and F. Pierobon (2012) "Shadow Banking, Sovereign Risk and Collective Moral Hazard," http://dx.doi.org/10.2139/ssrn.2060839.

Diamond, W. (2017) "Safety Transformation and the Structure of the Financial System," job Market Paper, Harvard University.

Diamond, D. W., and P.H. Dybvig, (1983) "Bank Runs, Deposit Insurance, and Liquidity," vol. 91(3), pages 401-419, June.

Farhi, E., Golosov, M., and A. Tsyvinski (2009) "A Theory of Liquidity and Regulation of Financial Intermediation," *Review of Economic Studies* vol. 76(3), pages 973-992.

Farhi, E. and J. Tirole (2012) "Collective Moral Hazard, Maturity Mismatch, and Systemic Bailouts," *American Economic Review*, 102(1): 60–93.

Farhi, E. and J. Tirole (2017) "Deadly Embrace: Sovereign and Financial Balance Sheets Doom Loops," *Review of Economic Studies*, forthcoming.

Feve, P. and O. Pierrard (2017) "Financial Regulation and Shadow Banking: A Small-Scale DSGE Perspective", TSE working paper 17-829.

Gennaioli, N., Shleifer, A., and R. Vishny (2012) "Neglected Risks, Financial Innovation, and Financial Fragility," *Journal of Financial Economics* 104 (3): 452–68.

Gennaioli, N., Shleifer, A., and R. Vishny (2013) "A Model of Shadow Banking," *The Journal of Finance*, 68, 1331–1363.

Gennaioli, N., Shleifer, A., and R. Vishny (2015) "Neglected Risks: The Psychology of Financial Crises," *American Economic Review: Papers & Proceedings*, 105(5): 310–314.

Gertler, M., Kiyotaki, N and A. Prestipino (2016) "Wholesale Banking and Bank Runs in Macroeconomic Modelling of Financial Crises," *Handbook of Macroeconomics*, edited by John B. Taylor and Harald Uhlig, 2: 1345–1425, chapter 16.

Gorton G. and A. Metrick (2010) "Regulating the Shadow Banking System," *Brookings Papers on Economic Activity*, 41: 261–312.

Grochulski, B. and Zhang, Y. (2015) "Optimal Liquidity Regulation with Shadow Banking," FRB Richmond Working Paper No. 15–12, http://dx.doi.org/10.2139/ssrn.2673540.

Hanson, S., Kashyap, A.K., and J. Stein (2011) "A Macroprudential Approach to Financial Regulation", *Journal of Economic Perspectives*, vol. 25(1), pages 3-28, Winter.

Hanson, S., Shleifer, A., Stein, J. and R. Vishny (2015) "Banks as Patient Fixed-Income Investors," *Journal of Financial Economics*, 117: 449–469.

Harris, M., Opp, C. and M. Opp (2014) "Higher Capital Requirements, Safer Banks? Macroprudential Regulation in a Competitive Financial System," Wharton Working Paper (revised 2017).

Holmström, B., and J. Tirole (1997) "Financial Intermediation, Loanable Funds and the Real Sector," *Quarterly Journal of Economics*, 112(3), 663–691.

Holmström, B. and J. Tirole (1998) "Private and Public Supply of Liquidity," *Journal of Political Economy*, 106(1): 1–40.

Jin, X. and F. Nadal de Simone (2016)? Tracking Changes on the Intensity of Financial Sector?s Systemic Risk?, BCL WP 102.

Kashyap, A.K., Rajan, R., and J.C. Stein (2002) "Banks as Liquidity Providers: An Explanation for the Coexistence of Lending and Deposit-Taking", *Journal of Finance*, vol. 57(1), pages 33-73, 02.

Malherbe, F. and Michael McMahon (2017) "Financial Sector Origins of Economic Growth Delusion," mimeo.

Moreira, A. and A. Savov (2017) "The Macroeconomics of Shadow Banking," forthcoming *Journal of Finance*.

Ordoñez, G. (2017) "Sustainable Shadow Banking," *AEJ: Macroeconomics*, forthcoming.

Peltzman, S. (1989) ?The Economic Theory of Regulation after a Decade of Deregulation? Brookings Papers: Microeconomics 1-41.

Perotti, E. (2014) "The Roots of Shadow Banking," in Claessens, Stijn et al. (eds), *Shadow Banking,* Federal Reserve Bank of Chicago International Banking Series.

Plantin, G. (2015) "Shadow Banking and Bank Capital Regulation,", *Review of Financial Studies*, 28(1): 146–175.

Pozsar, Z., T. Adrian, A. Ashcraft and H. Boesky (2013) "Shadow Banking," *Economic Policy Review*, 19(2): 1–16.

Shin, H. (2012) ?Global Banking Glut and Loan Risk Premium? IMF Economic Review, 60: 155?192.

Shy, O. and R. Stenbacka (2017) ?Ring-Fencing, Lending Competition, and Taxpayer Exposure?, available at SSRN: https://ssrn.com/abstract=2874316 or http://dx.doi.org/10.2139/ssrn.2874316.

Stein, J. (2012) "Monetary Policy as Financial Stability Regulation," *Quarterly Journal of Economics*, 127, 57–95.

Vickers, J. (2017) ?Economics of Structural Reform of Banks?, Thys Visser Lecture 3.

Yorulmazer, T. (2013) "Has Financial Innovation Made the World Riskier? CDS, Regulatory Arbitrage and Systemic Risk," discussion paper.