No. 16 | 2013

Staff Memo

Financial Stability

Why regulate banks?

Henrik Borchgrevink, Ylva Søvik and Bent Vale

NB NORGES BANK

Staff Memos present reports and documentation written by staff members and affiliates of Norges Bank, the central bank of Norway. Views and conclusions expressed in Staff Memos should not be taken to represent the views of Norges Bank.

Staff Memo inneholder utredninger og dokumentasjon skrevet av Norges Banks ansatte og andre forfattere tilknyttet Norges Bank. Synspunkter og konklusjoner i arbeidene er ikke nødvendigvis representative for Norges Banks.

ISSN 1504-2596 (online only)

ISBN 978-82-7553-768-1 (online only)

Why regulate banks?

Henrik Borchgrevink, Ylva Søvik and Bent Vale[1], Financial Stability, 27 June 2013

1. Introduction

In this note we seek to provide a first guide to how the economic literature explains the rationale for regulating banks. The guide is not exhaustive with respect to academic papers covered. Neither do we elaborate on how banks should be regulated; we only mention key measures in existing regulation.

Regulation is justified by market failures. More precisely, regulation is beneficial when market outcomes are socially inefficient and regulation can improve efficiency in a way that outweighs the costs of regulation. The inefficiencies can come from market power or externalities and information asymmetries. Consumer protection and safe provision of goods and services that are important for the society are two common reasons given for regulation.

Banks provide services that are important for a market-based economy: They channel money from savers to borrowers and monitor the borrowers; they provide maturity transformation, i.e., transform short term deposits into long term loans; execute payments and distribute risks. Of course, also other sectors provide important services to society, but banks are special for two reasons that together can make a dangerous mix. First, the banking system is inherently unstable, for reasons to be explained in this note. Second, problems in the banking sector potentially affect the economy more broadly and more deeply than problems in most other sectors: A sudden halt in banking services can paralyze the payment system and short-term credit lines ("today's economy") and investments ("tomorrow's economy") – across the entire economy. Moreover, unlike most other sectors, the failure of one bank tends to weaken its competitors in the short run, because banks are exposed to each other both directly and indirectly. This means good substitutes for one bank's services can be hard to find swiftly when substitutes are most needed.

The instability of the banking system may lead to a smaller provision of bank services than what is socially optimal. Nevertheless, the safety nets that have been implemented to make banking more stable, can easily lead to higher provision of bank services and more risk-taking than what is optimal. The safety nets can thus foster new rounds of regulation. The starting point for this note is the rationale for the first round of regulation, the fundamental reasons for regulating banks. Section 2 presents an important mechanism that can lead to instability: A coordination problem stemming from the fact that banks borrow short term and lend long term. Section 3 discusses moral hazard in banks: Banks have incentives to choose too risky portfolios, partly because depositors are not able to effectively monitor their banks, and partly because of the safety nets set up to make banking more stable. Section 4 discusses another source of market failure, that bank behavior at the individual bank level can have important negative externalities for the macroeconomy.

2. Multiple equilibria when banks borrow short term and invest long term

Bank creditors must have sufficient trust in banks' ability to repay their debts for the banking system to work properly. Without trust the banking system can go from stability to instability – possibly in

[1] Thanks to comments from Ida Wolden Bache, Jin Cao, Arild J. Lund, Kjell Bjørn Nordal and Pål Winje.

very short time. This vulnerability arises because banks are typically funded with deposits which the depositor can withdraw anytime, while banks typically invest the deposits in assets with maturity several years ahead. These assets are normally illiquid. That is, if the assets are sold (liquidated) before maturity, the bank will incur a loss relative to the value of holding the assets until maturity. The loss is incurred because the market has a lower valuation of the asset than the bank, e.g. because the bank has better information about the asset. This valuation asymmetry is related to one of the most important roles of banks in the economy, namely banks' comparative advantage in screening and monitoring loans – on behalf of depositors. When liquidating assets before maturity entails a loss, investing short-term deposits in long-term assets make banks vulnerable to excessive deposit withdrawals (bank runs). A bank run forces the bank to liquidate assets prematurely. Banks are vulnerable to bank runs even if banks lend only relatively short term. For example, also 3 months trade financing loans funded with on-demand deposits can create runs.

Diamond and Dybvig (1983) illustrate the vulnerability to bank runs through a game with multiple equilibria. These multiple equilibria are the source of market failure that causes instability: There exists a good equilibrium where depositors do not withdraw their money before the bank assets mature, except a fraction of the depositors who have a real need for their money earlier. In this good equilibrium, the bank can pay all depositors their money demands including interest at the point of withdrawal. As a consequence, it is rational for depositors not to withdraw earlier than their real need for money. But there also exists a bad equilibrium where depositors withdraw before they need their money, because they expect their bank to run short of money. This expectation is self-fulfilling; the bank must liquidate assets too early with a loss, to be able to pay out the deposits demanded. The loss entails that the depositors who withdraw late get less money back than the bank promised up front. As a consequence, it is rational for depositors to withdraw early. This coordination problem with multiple equilibria can lead sound banks into bankruptcy, implying a real economic loss to society. The loss stems from the liquidation of assets before the natural point of maturity, in addition to the arbitrary distribution of the loss among depositors. Diamond and Dybvig show that solvent banks can become insolvent because of a run on the bank. In practice, banks can of course become insolvent also without a bank run, when bank assets deliver too low payoffs at maturity. Uncertainty about a bank's solvency can ignite a bank run that leads to the bank becoming insolvent. The Diamond and Dybvig model considers consumer deposits. The coordination problem described is arguably even more prevalent for short-term wholesale financing.[2]

One way of reducing banks' vulnerability to runs is to abolish deposits that are available on demand, and allow banks to temporarily suspend conversion of deposits into cash when the bank is becoming short of liquid assets, or let banks instead be funded by marketable instruments with variable value similar to equity (see Jacklin (1993)). However, to abolish on demand deposits has negative effects. The fixed (nominal) value of deposits and their given rate of interest make deposits preferable as a savings instrument for risk-averse, non-sophisticated consumers. Moreover, equity contracts are less attractive for payment purposes because of the variable value. Banks' transformation of demand deposits to longer term assets with higher payoff, thus gives a benefit that can justify the cost of leaving banks vulnerable to runs or the cost of public measures that reduce this instability and make maturity transformation a viable business. As deposits are debt contracts that add to a bank's

[2] Run on a bank that *actually is insolvent* can, however, be efficient, and contribute to market discipline. See for instance Calomiris and Kahn (1991) or Allen and Gale (2007).

leverage, it also follows from this discussion that in order to reap the benefits of deposit contracts one has to accept that banks are highly leveraged.

Deposit insurance is a measure that can help avoid bank runs. When deposits are insured, there is less incentive to withdraw early. In fact, US banks mostly lent short term before more robust regulation, including deposit insurance, was established with the 1933 and 1935 Banking Acts. Until then, it was considered too risky to lend long term based on demand deposits.[3] These acts thus allowed for far more of socially valuable maturity transformation. Similarly to how deposit insurance can rule out the negative equilibrium in the Diamond and Dybvig model, central bank lending of last resort (LLR) may also eliminate the coordination problem. When liquidity is available from the central bank as an alternative to selling assets with a loss, the incentive for depositors and other creditors to run on a bank may vanish. Liquidity reserve requirements can also to some extent be used to reduce the vulnerability from banks borrowing short and lending long. Deposit insurance, LLR from the central bank and liquidity requirements are examples of regulations that can improve the market outcome. In that sense, the coordination problem described by Diamond and Dybvig is a market failure.

3. Moral hazard

When an agent undertakes actions that cannot be observed by other agents in the economy, and these actions confer risk on these other agents, it is called moral hazard. The standard textbook example of moral hazard is an insurance contract. The risk taken by the inured is transferred to the insurance company, but it cannot observe the risky actions of the insured. Thus, the latter may have incentives to take too much risk.

Shareholders of a firm with debt typically want the firm to take on high risk: If the firm succeeds, the shareholders earn a high return on equity because the equity is leveraged with debt. And if the firm fails, shareholders only lose the equity amount, due to limited liability. This phenomenon is prevalent in all firms with equity and debt. However, the problem is larger in banks, for reasons explained below.

Depositors are less capable of disciplining banks than creditors in other firms. Depositors are too small and too many to be able to coordinate efficient monitoring of a bank. Moreover, depositors will in general also lack the skills and information needed to monitor a bank. In fact, these shortcomings of depositors are closely linked to why banks exist: Banks screen and monitor loans on behalf of depositors – because depositors are not well equipped for the task (Diamond (1984)). With deposit insurance in place, the depositors in addition have weak incentives to monitor their banks.

Even more professional creditors often do not have full incentives or sufficient information to monitor banks. The incentive problem arises because the creditors might expect that the government will guarantee the bank's debt. This implicit guarantee stems from governments' inclination to bear the costs of saving a bank in distress rather than face the costs of letting the bank fail. When a failing bank is forced to close, its borrowers may try to switch to other non-failing banks. However, these borrowers risk encountering higher interest rates and possibly no loan offer at all, from the other non-failing banks, as these banks may also often be under financial stress when there

[3] In 1913 57 percent of US bank loans matured in less than 90 days. In 1940 the proportion was 30 percent, see Klebaner (1974).

are bank failures. Furthermore, they have less information on the financial soundness of the borrowers of a failed bank. For the economy, the costs of such a credit crunch can be of a large magnitude. The government will thus often have incentives to rescue banks rather than letting them fail.

Banks have higher leverage than non-financial firms, i.e. banks have less capital (equity) relative to total assets. The higher leverage, the stronger the incentives for shareholders to let banks take on high risk, as explained above. One obvious reason for higher leverage in banks is that banks offer deposit contracts. Access to on demand deposits is a good provided by banks, but deposits are debt contracts and therefore contribute to a bank's leverage. Another reason for high leverage is that bank debt is cheaper than equity because of the deposit insurance and the implicit government guarantee: Depositors and creditors will not demand high risk premiums even from risky banks, when their loans to the bank are considered guaranteed.[4]

Government support may also benefit shareholders more directly, e.g. when the government buys troubled assets from a distressed bank where shareholders have not been wiped out. Expectations of such support will give shareholders incentives to choose more risky portfolios and higher leverage.

Shareholders will choose management and incentive schemes in attempt to align the management's incentives with the incentives of shareholders. As a consequence, also the management of a bank will typically have incentives to take on too much risk in the bank. If the bank succeeds, the management will have high remuneration and possibly also high non-pecuniary benefits from the success. However, punishment of the management if the bank fails is contained to management losing their jobs, bonuses etc. That is obviously a too low punishment to induce bank managements to internalize the society's costs of financial distress in a bank.

These moral hazard issues entail higher than optimal risk in banks and can justify regulation. Removing the guarantees (deposit insurance and implicit government guarantees) will remove a fundamental reason for moral hazard in banks. But removing deposit insurance will increase the probability of bank runs (confer section 2). To remove the implicit government guarantee is not credible unless the authorities have other credible tools for tackling financial distress in a systemically important bank. The type of regulation that most efficiently can reduce moral hazard due to the guarantees is therefore crisis management frameworks that let shareholders and creditors bear the losses, while the bank's provision of credit and other important services are sustained. There are also other types of regulation that can reduce moral hazard. Capital requirements will increase the capital lost in case of failure and are therefore expected to reduce the incentive to have high risk ("more skin in the game"). Moral hazard stemming from deposit guarantee may to some extent be countered by risk-based deposit insurance premiums. Designating a public authority to monitor banks on behalf of depositors will help contain risk, since depositors are not well equipped to monitor banks. And regulation of the remuneration schemes of bank managers can induce managers to behave more prudently, e.g. by rewarding management more for profit in the long term and less for short term profits.

[4] The value of a guarantee on a bank's debt is increasing in the riskiness of the bank, as showed by Merton (1977)

4. Pecuniary externalities

The examples of externalities that are presented below are *pecuniary* externalities. A pecuniary externality occurs when the actions of one agent, for instance through sale of an asset, reduces the price of assets held by other agents in the economy. These are externalities that at the outset only affect the distribution of wealth between agents, and as such they are irrelevant for efficiency in a first best situation. However, pecuniary externalities can affect efficiency in the presence of market imperfections, like for instance asymmetric information and/or incomplete markets. In credit markets such imperfections will often occur. For example, if actions by other agents cause the value of a borrower's collateral to deteriorate, the borrower may become credit constrained and his profitable investment will not be carried out. As a result, a pecuniary externality has caused an efficiency loss in the real economy. Banks that suffer losses to their capital may also become constrained in funding markets and be forced to reduce their lending. If their borrowers do not have easy access to other sources of credit these borrowers may have to forego profitable investments. Hence, banks' pecuniary losses may cause losses in the real economy. When a pecuniary externality is likely to cause inefficiencies of macroeconomic proportions, intervention by regulators can be justified.

When a bank is in distress and does not pay the full amount on its liabilities, the counterparties of the bank incur losses. This is direct contagion of financial distress. Banks tend to have many and large exposures to other banks, so called interconnectedness. An important reason for interconnectedness is liquidity distribution among banks in the interbank market. The interbank market provides liquidity insurance to the participating banks, as shown in Bhattacharya and Gale (1987). Because of the high level of interconnectedness in the banking system, direct contagion between banks can be a threat to financial stability. Typically, a bank will not take into account losses incurred on counterparties, so this is an externality in the banking system

Contagion can also be indirect through lowered market prices when a bank in distress liquidates assets rapidly (fire sales). When bank A sells assets at a low price, the market value of similar assets in bank B may be reduced. Fire sales can thus harm banks that are not directly exposed to the bank in distress. During the financial crisis 2008-2009, fire sales contagion proved more important than direct contagion. With more marked-to-market accounting and with more leveraged banks, such indirect contagion is likely to be more pronounced. Moreover, indirect contagion will be more important the more similar assets the different banks are holding.

Banks that have similar portfolios will also be vulnerable to the same external shocks. Banks tend to become more similar when each bank diversifies its portfolio. In particular, big, diversified banks operating in the same markets will be similar. Wagner (2010) has shown how diversification makes each individual bank safer, but simultaneously the banking system becomes more vulnerable.

Similarly, non-financial borrowers of the banks can be vulnerable to the same shocks, if they have pledged relatively homogenous collateral. One example is growth in bank lending to the non-financial sector with housing as loan collateral. Due to incomplete markets, borrowers are unable to insure against adverse shocks to the value of the assets that they have pledged as collateral. The larger the accumulated debt by the non-financial sector, the more vulnerable the agents are to an adverse shock. If there is such an adverse shock to the economy the value of these assets deteriorates. If many agents become credit constrained by the fall in collateral value, the real

economic losses can be severe. Rational expectations are not sufficient to avoid these losses as individual lenders or borrowers have no incentive to internalize the pecuniary externality they confer onto other agents through their lending and borrowing decisions, see Bianchi (2011).

There can also be indirect contagion through provision of funding between banks: If one bank incurs unexpected losses it will tend to reduce its lending and hence reduce the supply of funding through the interbank market. This tightens the funding supply to other banks, and may reduce the lending from these banks. If banks expect such tightening to occur they may precautionary hoard liquidity, which amplifies the tightening. Reduced liquidity in these funding markets can also lead to reduced liquidity in other asset markets where banks invest. Reduced liquidity in these markets can also contribute to reduce the funding liquidity of banks, for instance through increased haircuts on collateral in repos.[5] This interaction between funding liquidity and market liquidity has been coined liquidity spirals (Brunnermeier and Pedersen (2009)), and was one of the main drivers of the 2007-08 crisis.

Because pecuniary externalities can make the banking system vulnerable even when individual banks seem sufficiently robust, bank regulators must monitor risks also at the macro level and have regulation measures to address vulnerabilities at the system level. Examples of such macroprudential regulation are liquidity requirements, capital requirements (both in general and higher capital requirements when systemic risk is rising), and to require banks to clear derivatives through a central counterparty to reduce contagion risks in the OTC derivative market.

[5] Funding liquidity is the ease with which funding can be aquired. See Brunnermeier and Pedersen (2009) for a more thorough discussion of these terms.

References

Allen and Gale (2007): Understanding Financial Crisis. Oxford University Press

Bhattacharya and Gale (1987): Preference Shocks, Liquidity and Central Bank Policy. In *New Approaches to Monetary Economics*, W. Barnett and K. Singleton (eds.), Cambridge University Press, 69-88.

Bianchi (2011): Overborrowing and Systemic Externalities in the Business Cycle, American Economic Review, American Economic Association, vol. 101(7), pages 3400-3426, December.

Brunnermeier and Pedersen (2009): Market liquidity and funding liquidity. Review of Financial Studies 22 (6) June, 2201-2238.

Calomiris and Kahn (1991): The Role of Demandable Debt in Structuring Optimal Banking Arrangements. American Economic Review 81 (3), 497-513

Diamond (1984): Financial intermediation and delegated monitoring. Review of Economic Studies 51 (3), 393-414.

Diamond and Dybvig (1983): Bank runs, deposit insurance, and liquidity. Journal of Political Economy 91 (3), 401-419.

Freixas and Rochet (2008): Microeconomics of Banking, Second edition, MIT Press

Jacklin (1993): Market rate versus fixed rate demand deposits. Journal of Monetary Economics 32 (1993), 237-258.

Klebaner (1974): *Commercial banking in the United States: a history*, Hinsdale Illinois, Dryden Press.

Merton (1977): An analytic derivation of the cost of deposit insurance and loan guarantees: an application of modern option pricing theory. Journal of Banking and Finance 1, 3–11.

Wagner (2010): Diversification at financial institutions and systemic crises. Journal of Financial Intermediation 19, 373-386.