COMMERCIAL BANKING AND

SHADOW BANKING:

The Accelerating Integration of Banks and Markets and its Implications for Regulation

ARNOUĐ W. A. BOÔT AND ANJAN V. THAKOR

(prepared as revised version of Chapter 3 in The Oxford University Press Handbook, The Accelerating Integration of Banks and Markets and its Implications for Regulation, 3rd edition.)

3.1 INTRODUCTION

The financial sector has evolved rapidly over the last decades, with the impetus for change provided by deregulation and advances in information technology. Competition has become more intense. Interbank competition within domestic markets as well as across national
borders and competition from financial markets have gained importance. Both the institutional structure of financial institutions and the boundary between financial institutions and financial markets have been transformed. At no stage has this blurring of boundaries been more evident than during the events leading up to the financial crisis that began in 2007, events that have highlighted how large the shadow banking sector has become. Pozsar et al. (2010) estimate the size of the shadow banking system in the US at $16 trillion in 2010, but estimates (and measures) vary greatly (see Claessens et al., 2012). A major issue with shadow banking is that because it involves qualitative asset transformation, it is inherently risky and may pose systematic risk that threatens financial stability (FSB, 2016). There have been other developments that have the potential of creating unforeseen risks. For example, since the 2007-09 financial crisis, P2P lending has grown rapidly both in the US and Europe, raising questions about the role of non-intermediated credit relative to intermediated credit. This chapter reviews the literature related to these developments and uses it to examine the importance of this changing landscape for the structure of the financial services industry and the design and organization of regulation.

As we will argue, the increasingly intertwined nature of banks and financial markets is not without costs. In particular, as the financial crisis of 2007–2009 has illustrated, systemic risks may have become more prevalent. In this chapter, we seek to provide a fundamental

---

1 In its Global Shadow Banking Monitoring Report 2016, the Financial Stability Board covering 28 jurisdictions with over 80% of world GDP, reports $34 trillion assets in shadow banks. See their ‘narrow’ definition that confines shadow banking to activities posing financial stability risk; e.g. it excludes assets in institutions that are not susceptible to runs like pension funds, (unlevered) closed-end funds and insurance companies.

2 In the definition of Adrian and Ashcraft (2016), shadow banking consists of financial institutions that are involved in credit, maturity, and liquidity transformation (which could create financial stability risks), but without the access to public backstops that banks have. Gorton and Metrick (2012) define the shadow banking system as one consisting of the following key components: (i) money-market mutual funds or other institutional (market-based) lenders who replace depositors as a primary funding source for shadow banks; (ii) securitization of bank-originated loans, which permits the creation of asset-backed securities that then serve as collateral for the bank’s borrowing from mutual funds and other institutional lenders; and (iii) repurchase agreements (or repos), which represent the financial contract used by banks to raise funding from investors.
analysis of the underlying forces that could explain the evolution of the banking industry. We begin by discussing the key insights from the financial intermediation literature, including the potential complementarities and conflicts of interest between intermediated relationship banking activities and financial market activities (underwriting, securitization, etc.). While debt contracts dominate the financial intermediation literature, the impressive growth of private equity firms has turned the spotlight on equity. In a sense, one could interpret private equity (PE) as intermediation driven from the equity side. Given their economic functions as debt and equity intermediaries, respectively, how do banks and PE firms interact?

Our discussion reveals that the interaction between banks and PE firms is only one aspect of an increasing integration of banks and markets. Banks have a growing dependence on the financial markets not only as source of funding but also for hedging purposes and offloading risks via securitization, and possibly for engaging in proprietary trading. Financial market linkages often also imply that intra-financial sector linkages mushroom, for example, the asset-backed securities created by securitization can serve as collateral that financial institutions use to fund themselves in the shadow banking system. The multiple dimensions of bank dependence on markets generate both risk reduction and risk elevation possibilities for banks. For example, while hedging may reduce risk, proprietary trading, providing liquidity guarantees for securitized debt, and taking positions in credit default swaps can increase risk as well. This raises potential regulatory concerns. What do these developments imply for prudential regulation and supervision? Will the increasing interactions between banks and markets increase or decrease financial system fragility? The financial crisis of 2007–2009 suggests an increase in fragility, but how much can we generalize from this crisis? These questions have become particularly germane not only
because of growing banks–markets integration, but also due to the (up to recently) growing cross-border footprint of financial institutions.

These developments have also focused attention on the role of “gatekeepers” (Coffee, 2002), like credit rating agencies. While the financial intermediation literature has acknowledged the role of credit rating agencies as information processors and sellers for some time now (e.g., Ramakrishnan and Thakor, 1984; Allen, 1990), the literature has not discussed how rating agencies may affect the fragility of the financial sector through the important role they play as “spiders in the web of institutions and markets.” We take up this issue in our discussion.

The organization of the chapter is as follows. In Section 3.2, we focus on the economic role of financial intermediaries. The primary focus here is on the banks’ role in lending and how this compares to non-intermediated finance directly from the financial market. We will also analyze the effects of competition on the banks’ lending relationships. Does competition harm relationships and reduce their value and hence induce more transaction-oriented banking, or does competition augment the value of relationships? This discussion will summarize the key insights from the modern literature of financial intermediation. In Section 3.3 we discuss the increasingly interconnected nature of banks and financial markets, with a focus on securitization. This “technology” has been at the center of the 2007–2009 financial crisis. What are the future prospects for securitization? The proliferation of non-banking financial institutions, and particularly private equity firms, is discussed in Section 3.4. We will argue that much of this activity might be complementary to the role of banks, rather than threatening their raison d’être. Subsequently, in Section 3.5 we focus on the role of credit rating agencies. These agencies have been indispensable for the explosive growth (and temporary demise) of securitization. How will their role develop? We then discuss in Section 3.6 regulatory implications. Here we link the role of banks in
lending (as emphasized in our earlier discussions) to their role as providers of liquidity. This brings in the issue of fragility, which is at the heart of the current regulatory debate.

3.2 UNDERSTANDING BANKS AS INFORMATION-PROCESSING INTERMEDIARIES

In this section we discuss two issues: (1) what is the key role of banks vis-à-vis markets? and (2) how does competition impinge on this role?

3.2.1 The Economic Role of Banks

We first discuss the role of banks in qualitative asset transformation—i.e., the process by which banks absorb risk to transform both the liquidity and credit risk characteristics of assets (see Bhattacharya and Thakor, 1993). For example, banks invest in risky loans but finance them with riskless deposits (e.g., Diamond, 1984; Ramakrishnan and Thakor, 1984; Millon and Thakor, 1985). They also invest in illiquid loans and finance them with liquid demandable deposits (e.g., Diamond and Dybvig, 1983). The theory of financial intermediation has placed special emphasis on the role of banks in monitoring and screening borrowers in the process of lending. Bank lending is typically contrasted with direct funding from the financial markets. What are the comparative advantages of bank loans over public capital market-bond financing?3

The most striking insight of the contemporary theory of financial intermediation is that banks are better than markets at resolving informational problems. The possession of better information about their borrowers allows banks to get closer to, and possibly more aligned

---

3 Much of the discussion that follows focuses on bank loans versus bond financing in the capital market, rather than equity financing in the market. In reality, we would expect the market to segment itself into some firms going for bank loans, some going for bond market financing, and some going for equity market financing. Boot and Thakor (1997) develop a theory that predicts the choice between bank loans and bond market financing. Brown, Martinson and Petersen (2017) provide evidence that better-developed stock markets support faster growth of high-tech industries, whereas better-developed bank-oriented credit markets foster growth in industries that rely on external financing for physical capital.
with their borrowers. Interestingly, a feedback loop is generated, as this proximity between 
the financier and the borrowing firm in bank lending arrangements may also help mitigate the 
information asymmetries that typically plague arm’s length arrangements in market 
transactions. This has several aspects. A borrower might be prepared to reveal proprietary 
information to its bank that it may have been reluctant to reveal to the financial markets 
(Bhattacharya and Chiesa, 1995). A bank might also gather information about prospective 
borrowers through their depository relationship with the bank, and may also have better 
incentives to invest in costly information acquisition. While costly, the substantial stake that 
it has in the funding of the borrower and the enduring nature of its relationship with the 
borrower—with the possibility of information reusability over time—increase the marginal 
benefit of information acquisition to the bank.5 Boot and Thakor (2000) analyze the 
economic surplus that relationship banking can generate.

Such borrower–lender proximity may also have a dark side. An important one is the 
hold-up problem that stems from the information monopoly that the bank may develop due 
to the spontaneous generation of proprietary information on borrowers. Such an 
informational monopoly may permit the bank to charge higher loan interest rates ex post 
(see Sharpe, 1990; Rajan, 1992; Boot, 2000, for a review). The threat of being “locked in,” 
or informationally captured by the bank, may dampen loan demand ex ante, causing a loss 
of potentially valuable investment opportunities. Alternatively, firms may opt for multiple-

---

4 Empirical evidence that depository information about potential borrowers is relevant to the bank is provided by 
Puri, Rocholl and Steffen (2017). That paper uses data on a million German loans to show that when a bank 
extends loans to those who have had a depository relationship with the bank (and continue to have it) exhibit 
lower default probabilities than those without depository relationships with the bank, consistent with one of the 
predictions in Donaldson, Piacentino and Thakor (forthcoming).

5 Ramakrishnan and Thakor (1984) and Millon and Thakor (1985) focus on pre-contract information 
asymmetries to rationalize the value that financial intermediaries add relative to markets. Diamond (1984) 
focuses on post-contract information asymmetries to rationalize intermediation. Coval and Thakor (2005) show 
that financial intermediaries can provide an institutional resolution of the problem of cognitive biases at the 
individual investor level, acting as a “belief’s bridge” between pessimistic investors and optimistic 
entrepreneurs. James (1987), Lummer and McConnell (1989), and Gande and Saunders (2005) provide 
empirical evidence on the informational value of bank financing. See also the “stories” provided by Berlin 
(1996) supporting the special role of banks.
bank relationships (see Carletti, Cerasi, and Daltung, 2007). This may reduce the informational monopoly of any individual bank, but possibly at a cost. Ongena and Smith (2000) show that multiple-bank relationships indeed reduce the hold-up problem, but can worsen the availability of credit (see Thakor, 1996, for a theoretical rationale).

Another aspect is that relationship banking could accommodate an intertemporal smoothing of contract terms (see Boot and Thakor, 1994; Allen and Gale, 1995, 1997) that would entail losses for the bank in the short term that are recouped later in the relationship. Petersen and Rajan (1995) show that credit subsidies to young or “de novo” companies may reduce the moral hazard problem and informational frictions that banks face in lending to such borrowers. Banks may be willing to provide such subsidized funding if they can expect to offset the initial losses through the long-term rents generated by these borrowers. The point is that, without access to subsidized credit early in their lives, “de novo” borrowers would pose such serious adverse selection and moral hazard problems that no bank would lend to them. Relationship lending makes these loans feasible because the proprietary information generated during the relationship produces “competition-immune” rents for the bank later in the relationship and permits the early losses to be offset. The importance of intertemporal transfers in loan pricing is also present in Berlin and Mester (1999). They show that rate-insensitive core deposits allow for intertemporal smoothing in lending rates. This suggests a complementarity between deposit taking and lending. Moreover, the loan commitment literature has emphasized the importance of intertemporal tax subsidy schemes in pricing to resolve moral hazard (see Boot, Thakor, and Udell, 1991; Shockley and Thakor, 1997) and also the complementarity between deposit taking and commitment lending (see Kashyap, Rajan, and Stein, 2002).

---

6 One strong implication of the Boot and Thakor (1994) theory is that the gains from relationship lending will take some time to be manifested. Recent empirical evidence in support of this prediction is provided by Lopez-Espinosa, Mayordomo and Moreno (2017) who document that the gains from relationship lending accrue only when the relationship is of longer duration than two years.
The bank–borrower relationship also displays greater contractual flexibility than that normally encountered in the financial market. This flexibility inheres in the generation of hard and soft proprietary information during a banking relationship. The information gives the bank the ability to adjust contractual terms to the arrival of new information and hence encourages it to write “discretionary contracts” ex ante that leave room for such ex post adjustments. This is in line with the important ongoing discussion in economic theory on rules versus discretion, where discretion allows for decision making based on more subtle—potentially non-contractible—information (see for example Simons, 1936; Boot, Greenbaum, and Thakor 1993).

The papers by Stein (2002), and Berger et al. (2005) highlight the value of “soft information” in lending. Soft information could be an example of more subtle and non-contractible information. On this issue, two dimensions can be identified. One dimension is related to the nature of the bank–borrower relationship, which is typically long term, with accompanying reinforcing incentives for both the bank and the borrower to enhance the durability of the relationship. This allows for implicit—non-enforceable—long-term contracting. An optimal information flow is crucial for sustaining these “contracts.” Information asymmetries in the financial market, and the non-contractibility of various pieces of information, would rule out long-term alternative capital market funding sources as well as explicit long-term commitments by banks. Therefore, both the bank and the borrower may realize the added value of their relationship, and have an incentive to foster the relationship.7

The other dimension is related to the structure of the explicit contracts that banks can write. Because banks write more discretionary contracts, bank loans are generally easier to

---

renegotiate than bond issues or other public capital market funding vehicles (see Berlin and Mester, 1992). Such renegotiability may be a mixed blessing because banks may suffer from a “soft-budget constraint” problem: borrowers may realize that they can renegotiate ex post, which could give them perverse ex ante incentives (see Dewatripont and Maskin, 1995; Bolton and Scharfstein, 1996). The soft-budget-constraint problem is related to the potential lack of toughness in enforcing contracts due to the ex post distribution of “bargaining power” linked with relationship banking proximity (see Boot, 2000). In practice, one way that banks can deal with this issue is through the priority structure of their loan contracts. If the bank has priority/seniority over other lenders, it could strengthen the bank’s bargaining position and allow it to become tougher. These issues are examined in Diamond (1993), Berglöf and von Thadden (1994), and Gorton and Kahn (1993).

The bank could then credibly intervene in the decision process of the borrower when it believes that its long-term interests are in jeopardy. For example, the bank might believe that the firm’s strategy is flawed, or a restructuring is long overdue. Could the bank push for the restructuring? If the bank has no priority, the borrower may choose to ignore the bank’s wishes. The bank could threaten to call the loan, but such a threat may lack credibility because the benefits of liquidating the borrower’s assets are larger for higher-priority lenders, and the costs from the termination of the borrower’s business are higher for lower-priority lenders. When the bank loan has sufficiently high priority, the bank could credibly threaten to call back the loan, and this may offset the deleterious effect of the soft-budget constraint. This identifies a potential advantage of bank financing: timely intervention. Of course, one could ask whether bondholders could be given priority and allocated the task of timely intervention. Note that bondholders are subject to more severe information asymmetries and are generally more dispersed (i.e., have smaller stakes). Both characteristics make them ill-suited for an “early intervention” task.
3.2.2 Intermediation and Competition

Since relationship banking is an integral part of the economic services provided by banks and generates rents for banks, it also potentially invites multiple-bank entry, which then generates interbank competition. An interesting question this raises is how competition might affect the incentives for relationship banking. While this may ultimately be an empirical question, two diametrically opposite points of view have emerged theoretically. One is that competition among financiers encourages borrowers to switch to other banks or to the financial market. The consequent shortening of the expected “life span” of bank–borrower relationships may induce banks to reduce their relationship-specific investments, thereby inhibiting the reusability of information and diminishing the value of information (Chan, Greenbaum, and Thakor, 1986). Banks may then experience weaker incentives to acquire (costly) proprietary information, and relationships may suffer. There is empirical evidence that an increase in relationship length benefits the borrower. Brick and Palia (2007) document a 21-basis point reduction in the loan interest rate due to a one-standard deviation increase in relationship length.

Moreover, increased credit market competition could also hurt relationship lending by imposing tighter constraints on the ability of borrowers and lenders intertemporally to share surpluses (see Petersen and Rajan, 1995). In particular, it becomes more difficult for banks to “subsidize” borrowers in earlier periods in return for a share of the rents in the future. Thus, the funding role for banks that Petersen and Rajan (1995) see in the case of young corporations (as already discussed) may no longer be sustainable in the face of sufficiently
high competition. This implies that interbank competition may have an ex post effect of diminishing bank lending.\(^8\)

Another way in which competition can hurt relationship lending is through consolidation. An extensive empirical literature focuses on the effect of consolidation in the banking sector on small-business lending. This consolidation may in part be a response to competitive pressures. The effects on small-business lending, however, are not clear-cut. Sapienza (2002) finds that bank mergers involving at least one large bank result in a lower supply of loans to small borrowers by the merged entity. This could be linked to the difficulty that larger organizations have in using “soft information” (Stein, 2002; Berger et al., 2005). However, Berger et al. (1998) show that the actual supply of loans to small businesses may not go down after bank mergers, since they invite entry of “de novo” banks that specialize in small-business lending (see also Strahan, 2007).

The opposite point of view is that competition may actually elevate the importance of a relationship-orientation as a distinct competitive edge. The idea is that competition pressures profit margins on existing products and increases the importance of financier differentiation, and more intense relationship lending may be one way for the bank to achieve this. Boot and Thakor (2000) formalize this argument to show that a more competitive environment may encourage banks to become more client-driven and customize services, thus generating a stronger focus on relationship banking.\(^9\) They distinguish between “passive” transaction lending and more intensive relationship lending by banks. Transaction lending competes head-on with funding in the financial market. Greater interbank competition results in banks

---

\(^8\) Berlin and Mester (1999) provide a related, albeit different argument. Their analysis suggests that competition forces banks to pay market rates on deposits, which may impede their ability to engage in the potentially value-enhancing smoothing of lending rates.

\(^9\) In related work, Hauswald and Marquez (2006) focus on a bank’s incentives to acquire borrower-specific information in order to gain market share, and Dinç (2000) examines a bank’s reputational incentives to honor commitments to finance higher-quality firms. Song and Thakor (2007) theoretically analyze the effect of competition on the mix between relationship and transaction lending, and focus on fragility issues raised by the bank’s desire to match core deposit funding with relationship lending and purchased money funding with transaction lending.
engaging in more relationship lending, but each relationship loan has lower value to the borrower. By contrast, greater competition from the capital market leads to a lower volume of relationship lending, but each relationship loan has greater value. In this context, it is also interesting to note that Berger et al. (2008) find empirically that bank ownership type (foreign, state-owned, or private domestic) affects the bank’s choice between transaction and relationship lending.

Relationships may foster the exchange of information, but may simultaneously give lenders an information monopoly and undermine competitive pricing. As discussed above, the informational monopoly on the “inside” lender’s side may be smaller if a borrower engages in multiple-banking relationships. This would mitigate the possibilities for rent extraction by informed lenders and induce more competitive pricing (see Sharpe, 1990; Petersen and Rajan, 1995). Transaction-oriented finance, however, may give banks little incentive to acquire information but is potentially subject to more competition. This suggests that markets for transaction-oriented finance may fail when problems of asymmetric information are insurmountable without explicit information acquisition and information-processing intervention by banks. This argument is used by some to highlight the virtues of (relationship-oriented) bank-dominated systems (e.g., Germany and Japan) vis-à-vis market-oriented systems. This is part of the literature on the design of financial systems (see Allen, 1993; Allen and Gale, 1995; Boot and Thakor, 1997). One objective of this literature is to evaluate the economic consequences of alternative types of financial system architecture.

What this discussion indicates is that the impact of competition on relationship banking is complex; several effects need to be disentangled. However, empirical evidence (see Degryse and Ongena, 2007) seems to support the Boot and Thakor (2000) prediction that the orientation of relationship banking adapts to increasing interbank competition, so higher
competition does not drive out relationship lending. Despite this adaptation, there is also evidence that in recent years the geographic distance between borrowers and lenders has increased (see DeYoung, Glennon, and Nigro, 2008). The latter might point at an increasing availability of data and data processing capacity which might challenge relationship banking. New specialized lenders have arisen that seek to replace relationship lenders and traditional credit scoring with sophisticated algorithms based on Big Data mining (data analytics). While still in its infancy, such analysis predicts creditworthiness by analyzing buying habits, lifestyle choices and all manner of opportunistic demographic correlates. One could envision similar developments enabling P2P lending as well.10

3.3 Bank Lending, Securitization, and Capital Market Funding

Much of our focus in the previous section was on interbank competition. Nonetheless, banks also face competition from the capital market. The standard view is that banks and markets compete, so that growth in one is at the expense of the other (see Allen and Gale, 1995; Boot and Thakor, 1997). In this context, Deidda and Fattouh (2008) show theoretically that both bank and stock-market development have a positive effect on growth, but the growth impact of bank development is lower when there is a higher level of stock-market development. They also present supporting empirical evidence. What this shows is that the dynamics of the interaction between banks and markets can have real effects. How banks and markets interact is therefore of great interest.

In contrast to the standard view that they compete, the observations in the previous section suggest that there are also potential complementarities between bank lending and

10 See chapter 18 in Greenbaum, Thakor and Boot (2016).
capital market funding. We argued that prioritized bank debt may facilitate timely intervention. This feature of bank lending is valuable to the firm’s bondholders as well. They might find it optimal to have bank debt take priority over their own claims, because this efficiently delegates the timely intervention task to the bank. The bondholders will obviously ask to be compensated for their subordinated status. This—ignoring the timely intervention effect—is a “wash.” In other words, the priority (seniority) and subordination features can be priced. That is, as much as senior debt may appear to be “cheaper” (it is less risky), junior or subordinated debt will appear to be more expensive, and there should be no preference for bank seniority, other than through the timely bank-intervention channel. Consequently, the borrower may reduce its total funding cost by accessing both the bank-credit market and the financial market.11 A theoretical analysis of complementarity appears in Song and Thakor (2010) who show that banks and markets exhibit three forms of interaction: competition, complementarity and co-evolution.

Another manifestation of potential complementarities between bank lending and capital market activities is the increasing importance of securitization, this being an example of the unbundling of financial services. Securitization is a process whereby assets are removed from a bank’s balance sheet, so banks no longer permanently fund assets when they are securitized; instead, the investors buying asset-backed securities provide the funding. Asset-backed securities rather than deposits thus end up funding dedicated pools of bank-originated assets. More specifically, the lending function can be decomposed into four more primal activities: origination, funding, servicing, and risk processing (Bhattacharya and

---

11 The complementarity between bank lending and capital market funding is further highlighted in Diamond (1991) and Hoshi, Kashyap, and Scharfstein (1993). Diamond (1991) shows that a borrower may want to borrow first from banks in order to establish sufficient credibility before accessing the capital markets. Hoshi, Kashyap, and Scharfstein (1993) show that bank lending exposes borrowers to monitoring, which may serve as a certification device that facilitates simultaneous capital market funding. In related theoretical work, Chemmanur and Fulghieri (1994) show that the quality of the bank is of critical importance for its certification role. This suggests a positive correlation between the value of relationship banking and the quality of the lender. See Petersen and Rajan (1994) and Houston and James (1996) for empirical evidence.
Thakor, 1993). Origination subsumes screening prospective borrowers, and designing and pricing financial contracts. Funding relates to the provision of financial resources. Servicing involves the collection and remission of payments as well as the monitoring of credits. Risk processing alludes to hedging, diversification, and absorption of credit, interest rate, liquidity, and exchange-rate risks. Securitization decomposes the lending function such that banks no longer fully fund the assets, but continue to be involved in other primal lending activities. One potential benefit of securitization is better risk sharing (see Gorton and Pennacchi, 1995 for an economic rationale for bank loan sales and securitization). The proliferation of securitization may, however, also be induced by regulatory arbitrage—for example, as a vehicle to mitigate capital regulation. And a third benefit is highlighted by Boot and Thakor (1993), who show that the pooling of assets and tranching of claims in securitization achieve both a diversification of idiosyncratic information and the creation of information-sensitive claims that increase the issuer’s revenues from selling these securities.

Central to the extensive academic work on securitization is the idea that it is not efficient for originators to completely offload the risks in the originated assets. The originating bank needs to maintain an economic interest in the assets in order to alleviate moral hazard and induce sufficient effort on the originating bank’s part in screening and monitoring. What this implies is that, even with securitization, banks do not become disengaged from the assets they originate. Banks still continue to provide the services involved in screening and monitoring borrowers, designing and pricing financial claims, and providing risk-management and loan-servicing support. As such, securitization preserves those functions that are at the core of the raison d’être for banks. This militates against the notion that securitization effectively lessens the importance of banks.

Boyd and Gertler (1994) have argued that the substitution from on-balance-sheet to off-balance-sheet banking induced by securitization may have falsely suggested a shrinking role
for banks. Indeed, by keeping banks involved in their primal activity of pre-lending borrower screening, securitization preserves much of the banks’ value added on the asset side.

Up to the 2007–2009 financial crisis, securitization was rapidly gaining in importance. In fact, prior to the summer of 2007, securitization became prevalent for ever-wider types of credits, including business credits that were previously thought to be difficult to securitize because of their information opaqueness. Also, a rather new market for securitization involving asset-backed commercial paper (ABCP) conduits emerged as a significant force. As the subprime crisis of 2007 has shown, these developments are not without problems. The structure of real-world securitization transactions appears to have taken a rather fragile form. In particular, it is important to note that much of the securitization leading up to the crisis involved the financing of long-term assets with short-term funding, which induced substantial liquidity risk. While this liquidity risk was sometimes mitigated by liquidity guarantees (e.g., stand-by letters of credit and refinancing commitments), the underwriting institutions often underestimated the risks involved and overstretched themselves. Recent events may cast doubt on the optimality of such strategies. Also, because the originating institutions appeared to have retained minimal residual risk, monitoring incentives may have been compromised (see Mian and Sufi, 2009). The eagerness of banks to securitize claims—and keep the repackaging “machine” rolling—may have also adversely impacted the quality of loans that were originated through a dilution of banks’ screening incentives due to lower retained residual risks (e.g., subprime lending; see Keys et al., 2010).

---

12 Most noteworthy are the bankruptcies among German Lander banks that were involved in providing liquidity guarantees.
13 Securitization is facilitated in part by credit enhancement, including partial guarantees by the arranger of a securitization transaction (and/or he holds on to the most risky layer of the transaction). In the recent credit crisis, this disciplining mechanism broke down; residual risk with the arranger was minimal or framed as liquidity guarantees to off-balance-sheet vehicles without appropriately realizing the inherent risks. The marketability of securitized claims has also been facilitated by accreditation by credit rating agencies. However, the role of rating agencies has been called into question during the subprime lending crisis, see section 3.5.
The 2007–2009 financial crisis brought securitization almost to a grinding halt. However, the risk diversification that securitization can accomplish appears to be of more than just ephemeral importance. Thus, we expect securitization to re-emerge, albeit possibly in a form that entails lower levels of liquidity risk, as well as lesser moral hazard in screening (loan underwriting standards) and monitoring. A caveat is that some of the activity in securitization might have been induced merely by capital arbitrage. With the stronger regulatory scrutiny following the financial crisis, we would expect such securitization to be discouraged.

Another effect of the interaction between banks and markets is that as markets evolve and entice bank borrowers away, banks have an incentive to create new products and services that combine services provided by markets with those provided by banks. This allows banks to follow their customers to the market rather than losing them. There are numerous examples. For instance, when a borrower goes to the market to issue commercial paper, its bank can provide a backup line of credit. In similar spirit, Drucker (2005) shows that junk-rated firms and companies in local lending relationships are more likely to select an integrated (universal) commercial investment bank when they expect to issue public debt in the future. This revealed preference for commercial investment bank relationships by firms that issue informationally sensitive securities suggests that there might be benefits for banks to use private information from lending in investment banking. A similar picture emerges if one looks at US banking following the 1999 Financial Services Modernization Act. It appears that information collected through the banks’ commercial lending businesses may have reduced the costs of underwriting debt and equity (see Drucker and Puri, 2005; Schenone, 2004). While this suggests a potential for value creation, an extensive literature has focused on the potential conflicts of interest related to banks combining lending and capital market activities; particularly, conflicts of interest in universal banking. Much of
earlier work is motivated by the Glass–Steagall regulation in the US (see Kroszner and Rajan, 1994; Puri, 1996; Ramirez, 2002). Typical findings are reassuring, i.e. conflicts were found to be limited.

In more recent work, a somewhat more critical picture has emerged; the problems with securitization, as already discussed, are a good example. Moreover, as Boot and Ratnovski (2016) show, combining relationship banking with financial market-oriented transaction activities (like trading) might undermine the commitment needed for relationship banking. More specifically, the ability to shift resources to trading activities within financial institutions may undermine relationship banking activities by violating (implicit) funding commitments to those borrowers. This might be particularly acute because trading activities are typically more readily scalable than relationship banking activities; i.e. the latter depend on more long term engagements leading to more cultivated relationships. This suggest that combining banking and trading activities could lead to lack of commitment and loss of franchise value. Consistent with this, Laeven and Levine (2007) find that banks that combine lending and non-lending activities lose value relative to engaging in these activities separately (see also Schmid and Walter, 2009).

The impetus for market-based activities grows stronger as interbank competition puts pressure on profit margins from traditional banking products, and the capital market provides access to greater liquidity and lower cost of capital for the bank’s traditional borrowers. As a consequence, there is a natural propensity for banks to become increasingly integrated with markets, and a sort of unprecedented “co-dependence” emerges that makes banking and capital market risks become increasingly intertwined.14 A discussion of

14 Innovations integrating banks and markets went far beyond securitization. For example, OTC derivatives, especially credit default swaps, showed in the period preceding the 2007-09 crisis enormous growth, outpacing real investment by a factor of twelve (Posen and Hinterschweiger, 2009). For further insights, see also Shleifer and Vishny (2010).
whether this is desirable and what the regulatory implications might be is given in section 3.6.

3.4 BANKS, EQUITY, AND PRIVATE EQUITY FIRMS

The emergence of non-banking financial institutions such as PE firms is considered by some to be a (further) signal for the diminishing role of banks. However, we will argue that these developments are rather complementary to the role of banks. Let us first discuss the role that PE firms play.

The arguments above about the need for banks to have seniority suggest a natural economic inhibiting of investments by banks in the equity of corporations. Equity “softens” a bank’s incentive to intervene for much the same reasons as does junior debt. So, while the emphasis of corporate finance theory on agency problems would suggest that it might be efficient for the bank to have both debt and equity claims on a corporation, this seems not to be advisable from a timely intervention point of view. This might explain why equity intermediation has largely been in the hands of PE firms and/or bulge-bracket global investment banks that typically engage less in relationship banking and focus more on transactions and the associated capital market activities.

Some more observations can be made about PE firms. Their activities could be viewed as intermediation driven from the equity side. That is, PE firms attract funding from a group of investors (“partners”) and invest the funds as equity in businesses. They are extensively involved in monitoring and advising these businesses. How different is this from the role that banks play as debt intermediaries? To address this question, note first that banks do occasionally take equity positions in their role as venture capitalists, particularly for later-stage financing where there is a prospect for developing a valuable relationship on the
lending side. Thus, banks participate in venture capital financing with higher probability if there is a greater likelihood of subsequent lucrative lending activity (Hellmann, Lindsey, and Puri, 2008). However, this may create a weakness in the participation decision. Bank-affiliated private equity investments do on average worse than non-affiliated investments (see Fang, Ivashina and Lerner, 2010). Banks may also have (participations in) PE subsidiaries that operate independently from the other businesses of the bank. However, this somewhat limited role as an equity financier does not mean that it would be efficient for the bank permanently to become an integrated provider of debt and equity finance, a “one-stop” financier of sorts. In particular, equity as a junior security may undermine a bank’s bargaining power and thus compromise its role in timely intervention. Also, soft-budget constraint problems may then (re)emerge.

At a more general level, one could ask whether the monitoring role of PE firms substitutes for the lending-related monitoring of banks. It might. Note, however, that equity and debt are fundamentally different securities. The type of monitoring needed will differ significantly. What will be true, however, is that the increasing involvement of PE investors induces banks to partner with these investors (often as providers of loans). In a sense, banks start building relationships with PE firms rather than the firms that the PE investors take equity positions in. This is not without risks since it may affect the added value of banks in timely intervention vis-à-vis the (underlying) borrower and even the banks’ incentives to be involved in this. However, to the extent that PE firms are an integral part of the capital market, this development too makes the involvement of banks in the capital market deeper and more intricate. Such complexity is further exacerbated by the emergence of other intermediaries such as hedge funds, particularly because of the growing importance of hedge funds as direct lenders. See Brophy, Ouimet, and Sialm (2009), who point out that hedge funds have emerged as “lenders of last resort,” providing finance to firms that banks
do not typically lend to. This is part of the growing importance of the shadow banking sector as a source of financing.

3.5 ROLE OF CREDIT RATING AGENCIES

Credit ratings are a fascinating part of today’s financial markets. Their importance is evident from the behavior of market participants. However, academic researchers have generally been skeptical about their incremental value, largely because of the absence of a theory of rating agencies. In the literature on financial intermediary existence, bank debt offers monitoring advantages that would not be available in the financial market. The typical argument for the lack of monitoring in the capital market is that free-rider problems among investors prevent effective monitoring. Boot, Milbourn, and Schmeits (2006) have shown that credit rating agencies (CRAs) add a monitoring-type element to the financial market, and thereby play a role as a “focal point” to resolve coordination failures among multiple dispersed investors (creditors). The CRA’s ability to resolve such coordination failure arises from the effect of its actions—the assigned rating and the “credit watch” process—on firm behavior via the conditioning of investors’ investment decisions on the assigned rating. Da Rin and Hellmann (2002) showed that banks could also resolve a multiple-equilibria problem among borrowers by helping coordinate the investment decisions of these borrowers. The role that Boot, Milbourn, and Schmeits (2006) give to CRAs has some similarity to this.

This role of CRAs in resolving coordination failures in the financial market qualifies the distinction between public debt and bank financing. The mechanism is, however, less “direct” than in the case of bank financing: the credit rating (and particularly the threat of a downgrade) induces good firm behavior rather than preventing bad behavior through direct intervention. Apart from bank loans, the non-bank private debt market also offers a potentially more direct alternative than credit rating agencies in the public debt market. In
fact, private debtors often impose more discipline than banks and hence serve even riskier borrowers (Carey, Post, and Sharpe, 1998).

Another mechanism that links banks and CRAs is the certification role of bank loans. Datta, Iskandar-Datta, and Patel (1999) show that the monitoring associated with bank loans facilitates borrowers’ access to the public debt market. This certification role of banks therefore complements what CRAs do. As rating agencies become more sophisticated and reliable, the certification role of banks diminishes in importance, causing bank borrowers to migrate to the capital market. In this sense, CRAs intensify the competition between banks and markets. But CRAs also pull banks into the capital market. For example, banks originate loans that they securitize, and then seek ratings for the securitized pools from CRAs. The ratings, in turn, facilitate the ability of banks to sell (securitized) asset-backed securities in the capital market.

One reason why credit ratings do not precisely reflect the credit risk of the rated debt instrument is that ratings are “coarse” relative to underlying default probabilities—there are only a little more than twenty ratings, but default probabilities lie in a continuum. This raises the question of why such coarseness exists. Goel and Thakor (2015) provide a theory in which they rationalize coarse credit ratings in a cheap-talk framework and show that coarseness addresses a truthful reporting (by the CRA) issue when the CRA has multiple objectives (issuers versus investors) pulling against each other. Thus, coarse ratings emerge in equilibrium even though coarseness has negative real effects.15

This largely positive interpretation of CRAs is clouded somewhat by recent negative publicity. In the 2001 crisis surrounding Enron, CRAs were accused of being strategically

---

15 See also Lizzeri (1999), in whose model, some pooling of credit types is induced by the profit maximization objective of the CRA. See Sangiorgi and Spatt (2017b) for an overview of the literature on credit ratings.
sluggish in downgrading.\textsuperscript{16} More recently, CRAs have been blamed (in part) for the sub-prime crisis in which they were allegedly too lenient in rating the senior tranches in securitization transactions. Allegations have been made about conflicts of interest for CRAs, arising from the fact that structured finance is a source of ever-increasing income for CRAs, which then corrupts their incentives for accurately rating the issuers involved in structured finance (Cantor, 2004). In this context, Coffee and Sale (2008) point out that it is naïve to think that reputation-building incentives alone would keep credit rating agencies in check.

Of particular concern are the so-called “rating triggers.” For example, some debt contracts may dictate accelerated debt repayments when the rating falls. The consequences of such accelerated debt repayments might, however, be so severe as to cause rating agencies to become reluctant to lower the ratings of those borrowers in a timely manner. Complications also arise from the role played by the so-called “monoliners.” These are insurers who traditionally guaranteed municipal bonds but now also guarantee the lowest-risk (best) tranches in securitization transactions. These insurers are virtually indispensable in the sense that the viability of many forms of securitization is predicated on this type of “reinsurance.” However, the ability of the monoliners to issue credible guarantees (and hence their role in securitization) depends on these institutions themselves having AAA ratings. This potentially generates an indirect chain-reaction mechanism for CRAs. In rating (and monitoring) the monoliners, CRAs affect the viability of the securitization market. Thus, the impact of CRAs is both direct (rating securitization tranches) and indirect (rating the monoliners). The potential failure of such monoliners would have a significant effect on the value of various structured finance products and induce an additional chain reaction.

\textsuperscript{16} As an illustration consider the following discussions in the US Senate: “On March 20, 2002, the Senate Committee held a hearing entitled ‘Rating the Raters: Enron and the Credit Rating Agencies’… The hearing sought to elicit information on why the credit rating agencies continued to rate Enron a good credit risk until four days before the firm declared bankruptcy…” (US Senate Hearings, 2002). Similarly, US Senate Staff Report (2002): “in the case of Enron, credit rating agencies displayed a lack of diligence in their coverage and assessment of Enron.” See also Cantor (2004) and Partnoy (1999).
among players active in the structured finance market, including investors. This further underscores the increasing interlinkages in the financial markets. Other concerns are related to the oligopolistic nature of the industry, and the importance that ratings have due to regulation. The latter includes the exclusivity given to a few rating agencies via the “Nationally Recognized Statistical Rating Organization” (NRSRO) classification, weakened in the 2006 Credit Rating Agency Reform Act, but also the references to external ratings in the Basel II capital regulation framework.

Under the Dodd–Frank Act 2010, the legal liability for CRAs has been elevated. Whether this will result in credit ratings that more accurately reflect credit risks is an open question.17

### 3.6 Regulation and the Second Raison D'Être for Banks: Liquidity Creation

In Section 3.2, we discussed the role of banks as information processors and delegated monitors. That information processing and monitoring referred to credit risk primarily. But banks also perform another important function, which is the provision of liquidity. The typically way this is framed is that banks invest in illiquid assets (loans) but finance themselves largely with highly liquid demand deposits, and through this intermediation process create liquidity in the economy. Liquidity is then created because depositors who invest in illiquid projects through the bank have liquid claims (demand deposits) that they would not have had if they had invested directly in those projects. The actual operations of banks, however, would have them make loans while simultaneously creating a matching

---

17 The Dodd-Frank Act also repealed the exemption given to CRAs in Regulation FD that allowed firms to have undisclosed material discussions with rating agencies. The institutional feature of rating shopping (i.e. firms may choose to hide ratings) is another element in the effectiveness of CRAs. See Sangiorgi and Spatt (2017a,b).
deposit in the borrower’s bank account, thereby creating new money (see Bank of England, 2014). This alternative framing, would still lead banks to provide liquidity to the economy.18

In the process of creating liquidity, banks expose themselves to the risk of unanticipated deposit withdrawals and become fragile. Our discussion of this issue in this section will focus on “institution-driven fragility,” manifested in the classic run on an individual bank, as well as “market-driven fragility,” that refers to risks that come primarily via the financial market and interbank linkages, and appear to be more systemic. We will discuss how the increasing integration of banks into financial markets allows banks to shift some of their traditional risks to the markets, and what this implies for financial system stability and regulation. Issues related to the economics of bank regulation are covered in Bhattacharya, Boot, and Thakor (1998, 2004).

3.6.1 Fragile Banks as Liquidity Providers

In the classical interpretation, a financial crisis is directly linked to the notion of bank runs. In a fractional reserve system with long-term illiquid loans financed by (liquid) demandable deposits, runs may come about due to a coordination failure among depositors (Diamond and Dybvig, 1983). Even an adequately capitalized bank could be subject to a run if the deadweight liquidation costs of assets are substantial. Regulatory intervention via lender of last resort (LOLR) support, deposit insurance, and/or suspension of convertibility could all help, and perhaps even eliminate the inefficiency. In fact, such intervention can be justified because of its potential to expunge the negative social externalities arising from the possible contagion effects associated with an individual bank failure. While these implications arise theoretically in a rather simple and stylized setting, many have generalized this simple setting by allowing for asymmetric information and incomplete contracts; see Rochet (2004) for a

18 In this spirit, Donaldson, Piacentino and Thakor (forthcoming)) have banks create private money by making loans that go beyond their stock of physical deposits, thereby creating “funding liquidity” and allowing the economy to invest more in real projects than its initial endowment.
The general conclusion is that fragility is real, and information-based runs are plausible. In particular, Gorton’s (1988) empirical evidence suggests that bank runs are not sunspot phenomena (as in Diamond and Dybvig, 1983), but are triggered by adverse information about economic fundamentals. More importantly, the banking crises stemming from such runs have independent negative real effects (see Dell’Ariccia, Detragiache, and Rajan, 2008). Also relevant in this context is the large literature that has now developed on banks and liquidity (see, e.g., Acharya and Schaefer, 2006; Acharya, Gromb, and Yorulmazer, 2007; Brunnemeier and Pedersen, 2009).

Given that bank runs are triggered by adverse information that depositors have about the financial health of banks, one might think that a simple solution would be to make banks safer by, for example, imposing higher capital requirements. Calomiris and Kahn (1991) first argued that the threat of bank runs may be a valuable disciplining device to keep bank managers honest, since a greater diversion of bank resources for personal consumption can increase the likelihood of a bank run. Building on this argument, Diamond and Rajan (2001) have suggested that financial fragility created by high bank leverage may play an important role in inducing banks to create liquidity, and thus a reduction in fragility through higher bank capital may lead to lower liquidity creation. Acharya and Thakor (2016) show that this link between bank leverage and liquidity creation has a dark side in that it causes higher bank leverage to generate higher systemic risk via spillover effects, namely inefficient “contagious liquidations” of healthy banks due to the observed liquidations of highly-levered failing banks. Until recently, there has been no empirical work done on this issue, in part because of a paucity of empirical measures of liquidity creation. Berger and Bouwman (2009) develop measures of liquidity creation and provide empirical evidence on the relationship between bank capital and liquidity creation. They show that higher capital leads to higher liquidity creation in the case of large banks (which create over 80% of the
liquidity in the US economy), and lower liquidity creation in the case of small banks. Since capital requirements also affect the asset portfolios of banks through their lending decisions (see Thakor, 1996) and these requirements may be binding for some banks, this raises issues about the interaction of credit and liquidity risks that need to be explored. Mehran and Thakor (2011) show both theoretically and empirically that bank capital and value are positively related in the cross-section, pointing to the private benefits of higher capital for banks. Admati et al. (2011) similarly argue that the commonly-asserted punitive costs of bank equity do not exist in reality, and stress the virtues of having higher capital.

Complicating this issue further is that the liquidity provision function of banks is also affected by the financial markets. Two observations are germane in this regard. First, access to financial markets weakens the liquidity insurance feature of demand-deposit contracts. To see this, note that the root cause of the fragility in the Diamond and Dybvig (1983) world is the underlying demand-deposit contract. The rationale for this contract—as modeled by Diamond and Dybvig (1983)—is the desire for liquidity insurance on the part of risk-averse depositors with uncertainty about future liquidity needs. However, as shown by von Thadden (1998), the very presence of financial markets allows depositors to withdraw early and invest in the financial market, which puts a limit on the degree of liquidity insurance. In fact, when the market investment opportunity is completely reversible, deposit contracts cannot provide any liquidity insurance. This is related to the earlier work of Jacklin (1987), who shows that deposit contracts have beneficial liquidity insurance features, provided that restricted trading of deposit contracts can be enforced.19 In any case, these arguments suggest that the proliferation of financial markets weakens the liquidity-provision rationale

---

19 Actually, Jacklin (1987) shows that with the “extreme” Diamond-Dybvig preferences, a dividend-paying equity contract can achieve the same allocations without the possibility of bank runs. However, for other preferences, a demand-deposit contract does better, provided that trading opportunities are limited.
for demand deposits, which may help explain the market-based proliferation of close substitutes for deposits.

A second observation has to do with whether the development of financial markets leads to a diminished role for the Central Bank in providing liquidity via its LOLR function. In the Bagehot tradition, one could ask whether the LOLR has a role to play in providing liquidity to liquidity-constrained-yet-solvent institutions when capital markets and interbank markets are well developed. Goodfriend and King (1988) argue that solvent institutions then cannot be illiquid since informed parties in the repo and interbank market would step in to provide the needed liquidity. In this spirit, former European Central Bank (ECB) board member Tommaso Padoa-Schioppa suggested that the classical bank run may only happen in textbooks since the “width and depth of today’s interbank market is such that other institutions would probably replace those which withdraw their funds” (as quoted in Rochet and Vives, 2004).

While these remarks correctly suggest that the development and deepening of financial markets could reduce the need for a LOLR in providing liquidity support, we believe that it would be hasty to conclude that there is no role for a LOLR, particularly when information asymmetries are considered. For example, Rochet and Vives (2004) show that a coordination failure in the interbank market may occur, particularly when fundamentals are weak, and that this may lead to a need for liquidity support by the LOLR for a solvent institution. The 2007–2009 financial crisis gives ample reason to believe that coordination failures in interbank markets are real and that the role of a LOLR is still important.

This discussion suggests two somewhat tentative conclusions. First, the development

---

20 Recent evidence provided by Berger, Black, Bouwman and Dlugosz (2017) shows that when the Federal Reserve increased banks’ access to the discount window through its Term Auction Facility, lending by (treated) banks went up. Another line of research studies the impact of liquidity on asset pricing (e.g., Acharya and Pedersen, 2005) and the possible role of asset price bubbles as a source of fragility and contagion (see De Bandt and Hartmann, 2002; and Allen, 2005, for surveys on contagion).
of financial markets (including interbank markets) has improved the risk-sharing opportunities available to banks and has probably decreased the likelihood of a run on an individual bank. Whether the total insolvency risk of individual institutions has declined depends on the actual risk-taking and capitalization. Second, because these improved risk-sharing opportunities have arisen from a greater degree of integration between banks and markets, they may also have contributed to an increase in systemic risk. In particular, financial market linkages (and focus) may have induced herding behavior (Boot, 2014).

Adrian and Shin (2010) point at the effect of favorable financial market conditions on leverage (increasing) and funding (becoming more fragile and short-term). These effects cause stress in the financial system at large when market conditions deteriorate. In other words, while the likelihood of an individual bank failing due to an idiosyncratic shock may have declined, there may be a concomitant increase in the probability that liquidity and solvency problems may propagate quickly through the financial system as a whole, leading to higher systemic risk. This raises thorny regulatory issues, which we turn to next.

3.6.2 Regulatory Implications

The preceding discussion has focused the spotlight on one fact: banks and markets are becoming increasingly integrated. This is happening in part because greater competition is inducing banks to follow their borrowers to the capital market and offer products that combine features of bank-based and market-based financing. It is also happening because banks themselves are using the financial market increasingly for their own risk management purposes. And the availability of market participants as purchasers of new bank products encourages financial innovations by banks. But, as Thakor (2012) shows, this can also
increase the likelihood of financial crises. There is thus a multitude of factors that have contributed to an astonishingly rapid melding process.\textsuperscript{21}

An important implication of this integration is that it is becoming more and more difficult to isolate banking risks from financial market risks. A financial market crisis inevitably cascades through the banking system, and what happens in the banking system does not take long to reverberate through the financial market. So, if the main task of bank regulators is the safety and soundness of the banking system, they must now also worry about the financial market whose participants are outside the bank regulator’s domain. Explicit recognition that these sorts of effects have created the specter of “endogenous systemic risk” has led to the creation of the Financial Stability Oversight Council (FSOC) in the US and the European Systemic Risk Board (ESRB) in the EU as parts of the post-subprime-crisis regulatory landscape.

Moreover, even though the explicit insurance guarantee applies only to bank deposits, the temptation for government regulators to bail out various uninsured participants—including investment banks and financial market investors—in the event of a crisis in the capital market seems difficult to resist on ex post efficiency grounds, particularly because of the implications for bank safety and systemic stability.\textsuperscript{22} It will be interesting to examine the connotations of this for ex ante incentives and the magnitude of the implicit “soft” safety net provided by the government. What seems safe to conjecture is that a perception of a greater regulatory concern with ex post efficiency—and hence a greater desire to intervene—has

\begin{footnotesize}
\begin{itemize}
\item[21] Interestingly, the fact that this integration can increase the risk to which the bank’s depositors are exposed can induce banks to slow down the integration to enhance the value of the bank’s services to its depository customers, as shown by Merton and Thakor (forthcoming). The key is that many bank services benefit from being remote from the credit risk originating from the bank itself. This point also relates to the value of creating risk-free claims in the economy (see also Dang et al, 2015).
\item[22] The guarantee provided in 2008 to a collapsing Bear Stearns by the government to facilitate its sale to JPMorgan Chase is an example, as are the general measures to let investment banks qualify for a commercial banking license (and in doing so allow them access to deposits and let them qualify for deposit insurance). Bailouts were common in the financial crisis, also in Europe. Following the crisis, attempts have been made to make, so-called, bail-ins possible; meaning that upon the rescue of a financial institution unsecured financiers would lose their money.
\end{itemize}
\end{footnotesize}
elevated the importance of moral hazard. And this has happened in an environment in which regulatory issues are becoming increasingly international, both due to the cross-border proliferation of financial institutions and the increasing integration of banks with financial markets, which are typically international in scope. 23

The decentralized, mainly national structure of regulatory and supervisory arrangements in a financial world that operates across borders may give rise to potential conflicts of interest between the national authorities and “outsiders.” For example, national authorities might be prone to “too-big-to-fail” (TBTF) rescues, and this worsens the moral hazard on the part of large institutions. Yet one could argue that the moral hazard engendered by TBTF policies could be attenuated somewhat by attaching to TBTF rescues specific provisions that would involve replacing management, wiping out the claims of shareholders and uninsured debtholders, etc. This is true in theory but does not appear to happen often in practice. One reason might be the possibility of capture of local regulators and supervisors due to the closeness of their relationships to the “national flagship” institutions (Boot and Thakor, 1993). There are also issues of “too many to fail” (see Acharya and Yorulmazer, 2007) or “too interconnected to fail” (Herring, 2008), which could also induce regulatory leniency toward these institutions. Alternatively, national authorities may not sufficiently internalize the disruptive consequences that a domestic bank failure could have in other countries. Efficiency might be hampered in other ways as well. For example, the national scope of supervision may help encourage the emergence of “national champions” among regulators, who may then seek to protect institutions in their countries. More fundamentally, 

23 The importance of international coordination was already on the radar screen far before the financial crisis. The Basel (BIS) capital accords (with agreements on minimum capital requirements) could be seen as a first success story of international coordination. Typically progress was a response to crises; e.g. the first step came with the creation of the Basel Committee in 1974 following the Herstatt failure (a relatively small bank that via its international linkages nevertheless had a big effect on international financial markets). Its first major (advisory) document followed in 1975 on host and home country supervisory responsibilities (the Basel Concordat). Several revisions followed, including the 1992 revision that followed the 1991 failure of the Luxembourg based Bank Credit and Commerce International (BCCI) that pointed at an ill-defined home country definition. In 1988 the first Basel capital accord came about (Basel I). For more, see Alessi (2012).
the decentralized structure could give rise to an uneven playing field, regulatory arbitrage possibilities, and coordination failures in the resolution of financial distress in cross-border operating institutions.

Casual observation would seem to suggest that integration and further coordination (if not centralization of authority) of both regulation and supervision might yield substantial efficiency gains not only for the supervisory authorities but also, and perhaps more importantly, for the supervised financial institutions themselves.

3.6.3 Cross-Border Coordination in Regulation and Supervision: The European Union Example

Bank regulation and supervision in the EU was historically under the authority of domestic regulators. That meant that the EU had 35 supervisory authorities responsible for prudential supervision, and a typical large financial institution might had to report to more than 20 supervisors (Pearson, 2003). The global financial crisis showed that national regulation and supervision of banks was inadequate in the highly interconnected European banking market, especially in the euro area where the single currency (the Euro) led to a fast integration of the financial system. The regulatory overhaul in the EU has led to the establishment of a common regulatory and supervisory framework in the EU, the so called Banking Union. The Banking union is built on a single rulebook for financial institutions in the EU and comprises of the Single Supervisory Mechanism (SSM) and the Single Resolution Mechanism (SRM); a pan-European system of joint deposit insurance is envisioned at a later stage.

---

24 This section follows in part chapters 15 and 16 in Greenbaum, Thakor and Boot (2016).
25 Several policy responses created problems during the financial crisis, most notably the non-coordinated actions surrounding deposit insurance. Some countries chose to offer blanket guarantees overnight (e.g., Ireland) and in doing so imposed severe externalities on other countries and also foreign banks in their own markets that were not covered. These foreign countries and banks faced an immediate erosion of their deposit base.
26 The sustainability of the euro as common currency in the Eurozone was the key driving force behind the banking union. It was felt that weak domestic banks undermine their local governments, and via that channel the sustainability of the euro. Alternatively, governments might use their domestic banks as a source of financing, which might encourage irresponsible fiscal policies that are not compatible with having a common currency.
The single rulebook contains the legal framework of regulatory rules that all EU financial institutions must comply with. It legislates capital regulation, synchronizes rules for deposit insurance, and sets up the rules for bank regulation and prevention of bank failures.

The SSM gives the ECB the authority to supervise the largest banks in the euro area. The national supervisors maintain supervision over the remaining – smaller – national banks but subjected to control from the ECB. The ECB monitors whether banks comply with the single rulebook (in particular, whether they are adequately capitalized) and is responsible to trigger timely intervention into failing banks.

The SRM establishes a new, independent EU Agency called the Single Resolution Board to deal with failing banks within the EU. It also establishes the Single Bank Resolution Fund built up by contributions from banks in the EU. The Single Resolution Board will use resolution tools granted by the SRM and the funds from the Single Bank Resolution Fund to effectively deal with failures of national banks and of cross border banks.

The Banking Union follows an earlier attempt to integrate the European financial sector. In 2008 the EU established the European System of Financial Supervision (ESFS) consisting of three new European Supervisory Authorities (ESAs) – for banking, insurance and financial markets – and the European Systemic Risk Board (ESRB). The ESRB is responsible for macroprudential surveillance of the financial system in the EU, somewhat like the Financial Stability Oversight Council in the U.S. Its tasks include prevention of systemic risk, safeguarding financial stability, and the smooth operation of the financial sector in support of the real economy.

The three European supervisory authorities – the European Banking Authority (EBA), the European Insurance and Occupational Pensions Authority (EIOPA), and the European Securities and Markets Authority (ESMA) – are three independent regulatory agencies, each

---

27 The ESAs and ESRB came into existence as recommendations of the so-called Larosièye Committee that had been installed by the EU immediately at the onset of the global financial crisis in 2008.
responsible for supervision in their respective areas. But contrary to the ECB-linked Banking Union, each ESA primarily has “soft” powers, meaning that its role is mainly to provide coordination. The EBA is responsible for determining the uniform regulatory and supervisory technical standards, guidelines and best practices and their applications across the EU. The EBA can provide opinions to the European Parliament, the Council, and the European Commission. The EBA also acts as a mediator to resolve potential conflicts between national supervisors and acts as a coordinator in emergency situations. How the ESAs and ESRB fit within the Banking Union is still open to debate. Several issues may come up, including the different geographic reach (the ESAs and ESRB are linked to the 28 EU member states; the Banking Union primarily addresses the 19 Euro members), and also turf battles between the EU and the ECB are possible (the Banking Union is linked to the ECB while ESAs and ESRB “belong” to the EU).

Another issue is the effectiveness of the Banking Union. Practical considerations suggest that a full integration of all regulatory and supervisory functions at the European level might not be easy to do in a way that guarantees effectiveness. While it is clear that regulatory and supervisory integration need to keep pace with the development of the size and the cross-border footprint of the covered banks, the heterogeneity of underlying supervisory systems and the implied costs of integration should not be underestimated. An interesting illustration is the evidence reported by Barth, Caprio, and Levine (2004) on the variation across the European Union (EU) countries in supervisory institutions and practices. Their conclusion was that supervisory arrangements within the EU are as diverse as in the rest of the world. Illustrating this point further, the EU countries are current or former standard bearers of all major legal origins. A vast literature now documents how legal origin matters for the shape and functioning of the financial system (see LaPorta et al., 1998). Bank regulation and
supervisory practices also differ considerably between civil and common law countries, typically with a more flexible and responsive approach in the latter.

While common sense suggests that a more integrated regulatory and supervisory structure is desirable, the preceding observations predict considerable challenges.

As a final observation, the status quo on the role of the central bank has been challenged following the 2007–2009 financial crisis. The crisis effectively created a situation in which central banks got a heavier role in supervision. While central banks always had a role in safeguarding the stability of the financial system, during the 2007–2009 crisis we have seen that both the Federal Reserve and the ECB became directly involved in rescuing depository as well as non-depository financial institutions. In case of the ECB, this expanded role has been formalized in the Banking Union. More specifically, decisions have been made such that the responsibility for the supervision of the largest banks will migrate to the ECB (the Single Supervisory Mechanism—SSM). This reflects a significant change in thinking. Prior to the crisis, the consensus appeared to be that caution was in order when it came to expanding the mandate of central banks, because an expanded mandate could compromise the pivotal function of central banks in conducting monetary policy.\(^{28}\)

### 3.6.4 Other Reform Suggestions

The struggle for better cross-border coordination in regulation and supervision should go hand in hand with more fundamental reforms in the regulatory structure. The first is that the scope of regulation and supervision needs to be clearly identified and, if possible, contained. Effective supervision and regulation—given the mushrooming cross-sector and cross-border

---

\(^{28}\) An important distinction needs to be made between business conduct regulation and prudential regulation. We have focused on the latter. The former is closer to the functioning of financial markets and would typically lend itself more readily for centralization at the European level. The progress with the Banking Union has actually led to a more noticeable integration push along the prudential dimension (obviously, all this was crisis induced). As follow up on the ‘soft-power’ Larosière report, the EU now has proposals in progress to promote a Capital Market Union that would seek further harmonization in financial markets.
footprint of financial institutions—requires a better delineation of safety and systemic risk concerns. The cross-sector integration of financial institutions and the increasingly more seamless integration of financial markets and institutions have considerably broadened the scope of regulation and the potential sources of systemic risk.

Another relevant question is whether market discipline could help in containing systemic risks, or whether market responses merely amplify such risks (see Flannery, 1998). Here the picture gets a bit murky. Basel II tries to encourage market discipline via its third pillar that is aimed at greater transparency. The idea is that market discipline could help supervisors in safeguarding the well-being of the financial sector. This has merit on the face of it and has support in the literature as well. The literature has viewed market discipline working in three ways: (1) by providing regulators with market-based signals of bank risk-taking through the yields on subordinated debt issued by banks; (2) by providing banks with disincentives to take excessive risk through the upward adjustments in sub-debt yields in response to greater bank risk; and (3) by choking off the supply of sub-debt when sufficiently high risk-taking by the bank is detected by the market, thereby providing additional encouragement to the bank to temper its risk-taking. Nonetheless, it has been shown both theoretically and empirically that market discipline can be effective only if the claims of uninsured investors (sub-debt and equity) are not protected via de facto ex post insurance in a government-sponsored rescue of a failing institution. For a theoretical treatment of these issues, see Decamps, Rochet, and Roger (2004), and for empirical analyses that support the risk-controlling role of market discipline, see Barth, Caprio, and Levine (2004), and Goyal (2005).

However, despite all of the research support for the role of market discipline, our knowledge of whether market discipline facilitates or hinders the regulatory task of maintaining banking stability during a financial crisis is quite limited. In particular, when
the financial sector is severely stressed, as during the 2007–2009 credit crisis, market discipline may induce herding behavior, as everybody “heads simultaneously for the exit,” and this actually could be a source of instability. This suggests that regulation and supervision in “normal times” should perhaps be distinguished from that during crisis episodes. Market discipline, although valuable in normal times, may be very distortive in times of systemic stress. This may be one reason why, during crises, regulators have been inclined to provide more or less blanket guarantees to distressed institutions, ostensibly to counter the potentially adverse effects of market discipline.

To complicate matters even further, it would be dangerous to conclude that market discipline, say via the use of market value accounting and other mechanisms, is something that can be relied upon in good times and eschewed in bad times. The key is to figure out the appropriate regulatory actions in good times—when banks have the flexibility to comply without compromising their viability—that would enable banks to be more capable of withstanding the stresses of market discipline during bad times. In such good times risk management by banks tends to be corrupted at the same time that market discipline is the weakest and risk in “underpriced” (see Thakor (2015, 2016) and Boot (2014), for example). Market discipline may thus not work in good times, or even be counterproductive due to the “underpriced” risk inviting excessive leverage and risk taking. It will also be important to remember that banks cannot be completely insured from the effects of market stress during bad times (e.g., through the use of blanket guarantees for all claimants), or else the ex ante effectiveness of market discipline is lost entirely (e.g., Decamps, Rochet, and Roger, 2004).

This brings up the issue of introducing firewalls in the financial sector. For example, does a subsidiary structure reduce systemic risk concerns? We do not think that an answer is readily available. More generally, what type of constraints, if any, should be put on the corporate structure of financial institutions? Until the 2007–2009 financial crisis, the general
belief was that deregulation in the financial sector would continue further, possibly leading to even bigger and broader financial institutions. But now it is far from clear what the future will bring. Some have suggested reintroducing the Glass–Steagall Act to insulate local banking from the risks and fads that periodically afflict financial markets. Proposals that echo the Glass–Steagall Act include the Dodd–Frank Act in the US, the Vickers Report in the UK (Vickers, 2011), and the Liikanen Report (Liikanen, 2012) in the EU. To what extent these are effective, and not overly costly, is open to debate.\footnote{All these proposals seek to protect core banking functions against risks originating from financial markets. In the case of the Dodd-Frank Act, restrictions particularly aim at containing risks coming from private equity, hedge fund investments and derivatives. Vickers and Liikanen focus on \textit{internally} separating banking operations. See BIS (2013, chapter 5) for a discussion and comparison of the various proposals. Farhi and Tirole (2017) develop a theoretical framework for such containment.} In any case, changes in the industrial structure of the financial sector might be of paramount importance for the design and effectiveness of regulation and supervision.\footnote{Also important might be the ownership structure of financial institutions (see Berger et al., 2008). The concentration in the credit rating business and the importance of ratings for structured finance (securitization) is another issue. Structural changes could be desirable here as well.}

A second issue has to do with the evolution of capital regulation. The introduction of Basel II rules meant that banks could fine-tune their required capital ratios based on their (certified) internal models. There are questions about whether these models induce procyclicality, and whether such model-dependency induces systemic risk by itself (e.g., institutions using the same models, and thus potentially being subject to the same shortcomings). There have also been concerns about the potential adverse consequences of the discretion that Basel II provides.\footnote{This concern stems from the observation that individual banks are unlikely to sufficiently internalize the systemic-risk externalities of their actions. Consequently, the latitude that Basel II grants in having banks use their own internal risk assessment models to determine appropriate capital levels might be troublesome. Banks may tweak these models in order to generate prescriptions to keep low levels of capital (see Behn, Haselman and Vig (2014) for evidence on this for German banks). The follow-up with Basel III seeks to provide remedies.} Perhaps similar concerns led the FDIC to impose a minimum leverage ratio on banks in the Basel II environment; an element that the post-crisis Basel III amendments have introduced in the Basel framework as well. The FDIC has argued that requiring a minimum level of capital—regardless of risk—is essential for timely

---

\footnotetext[29]{All these proposals seek to protect core banking functions against risks originating from financial markets. In the case of the Dodd-Frank Act, restrictions particularly aim at containing risks coming from private equity, hedge fund investments and derivatives. Vickers and Liikanen focus on \textit{internally} separating banking operations. See BIS (2013, chapter 5) for a discussion and comparison of the various proposals. Farhi and Tirole (2017) develop a theoretical framework for such containment.}

\footnotetext[30]{Also important might be the ownership structure of financial institutions (see Berger et al., 2008). The concentration in the credit rating business and the importance of ratings for structured finance (securitization) is another issue. Structural changes could be desirable here as well.}

\footnotetext[31]{This concern stems from the observation that individual banks are unlikely to sufficiently internalize the systemic-risk externalities of their actions. Consequently, the latitude that Basel II grants in having banks use their own internal risk assessment models to determine appropriate capital levels might be troublesome. Banks may tweak these models in order to generate prescriptions to keep low levels of capital (see Behn, Haselman and Vig (2014) for evidence on this for German banks). The follow-up with Basel III seeks to provide remedies.}
regulatory intervention in the event of problems. Such timely intervention seems particularly
important in cross-border situations, given the complexities created by bank failures when
multiple countries are involved. Timely intervention could help contain conflicts between
local authorities in such cases (see Eisenbeis and Kaufman, 2005). This is one reason why
new rules are proposed—commonly referred to as Basel III—that stipulate higher capital
requirements, and indeed also a leverage ratio, although one could justifiably argue that the
levels of even these higher requirements may be well short of adequate.32

A third issue is deposit insurance. The 2007–2009 financial crisis has made it clear that,
when a real crisis hits, national authorities effectively feel compelled to fully guarantee the
deposit bases of their financial institutions to eliminate the possibility of massive runs. This
heavy dependence on insured deposits is an issue that needs a re-examination. Extant
research (see Bhattacharya, Boot, and Thakor, 1998) has clearly shown the moral hazards
that insured deposits entail. Moreover, Barth, Caprio, and Levine (2004) have shown that
high levels of (de facto or de jure) deposit insurance impede the effectiveness of market
discipline and increase the likelihood of a banking crisis. A question is whether strict
regulatory limits should be put on the risks that institutions can expose these deposits to.
Earlier research had at some point advocated narrow banking, which fully insulates insured
deposits. But are there alternatives? And, more generally, can insured deposits be made less
important as a funding vehicle for financial institutions?

A fourth issue is whether regulation and supervision sufficiently address macro
prudential issues, in particular systemic concerns. Despite many references to systemic risk,
it appears that the majority of regulatory initiatives are focused on the well-being of
individual financial institutions. That is, a micro prudential focus dominates (see

32 Berger and Bouwman (2013) provide empirical evidence that higher capital produces greater benefits for
banks during financial crises, including a higher probability of survival. This is consistent with Thakor’s (2012)
theory that higher capital weakens incentives for banks to introduce financial innovations that are associated
with higher probabilities of financial crises.
Brunnemeier et al., 2009). This should be addressed to better reconcile regulation and supervision with the systemic concerns that are paramount.

The fifth issue is that very little is known about the efficiency and effectiveness of various regulatory and supervisory structures. As Barth et al. (2003) put it, “there is very little empirical evidence on how, or indeed whether, the structure, scope or independence of bank supervision affects the banking industry.” Their own research suggests that the effect is at best marginal, but measurement problems are vexing. They suggest that narrowing the focus on the effect that regulation has on systemic risk may help. But here, too, little is known about the regulatory structures that are most efficient in dealing with systemic risk. We need considerable additional research to sharpen our identification of the costs and benefits of different regulatory and supervisory arrangements. Given the strikingly different national supervisory arrangements that exist today, our lack of knowledge on this issue is a significant barrier to progress toward a harmonized “superior” model.33

Finally, more research is needed on the role that bank culture can play in attenuating problems of excessive risk and financial fragility. Song and Thakor (forthcoming) have recently provided a theory of bank culture in which culture acts as a mediating variable in bank risk taking and may attenuate the propensity of banks to herd on excessive and correlated risk taking. This theory is a start, and it complements discussions within the Federal Reserve System in the U.S, and in the ECB with respect to what can be done to change corporate culture in banking (Dudley, 2014). The evidence in Cohn, Fehr and

33 Some theoretical work suggests that competition between regulatory regimes might be helpful; see Kane (1988). This touches on a broader point: diversity in the financial sector – diversity in regulatory approaches, bank business models and ownership structures etc. – can be valuable. Too much homogeneity could invite systemic risk by itself (Boot, 2014; Butzbach, 2016; see also Berger et al., 2008).
Marechal (2014) suggests that the culture in banking may encourage dishonest behavior; see also Cerquetti, Fiordelisi and Rau (2016).\textsuperscript{34}

3.7 CONCLUSIONS

We have reviewed some of the literature on why banks exist, the risks they create, and how interbank competition as well as that from markets affects the economic roles served by banks as well as the attendant risks. One important development is that banks have become increasingly integrated with markets. This integration generates two effects that work in opposite directions. On the one hand, individual banks become better equipped to manage their own risks because it becomes easier and less costly to hedge these risks using the market. This could reduce the risk of an individual bank failing due to an idiosyncratic shock. On the other hand, there is an increase in the probability that a shock to a small subset of banks could generate systemic effects that ripple through the financial market, so that this banks–markets integration may be causing an elevation of systemic risk.

It is easy to see that this substantially complicates the task of prudential regulation of banks and raises the specter of a widening of the “implicit” governmental safety net as ex post efficiency concerns tempt the government to bail out even uninsured players. This is no longer a mere theoretical conjecture, as demonstrated by the bailouts of investment banks and insurance companies in 2008–2009. We believe that these are important issues that deserve greater theoretical and empirical attention. In particular, we need to have a better understanding of what the regulatory intervention should be in a crisis. Governmental initiatives such as those witnessed in the US during the 2007–2009 crisis—massive governmental injections of liquidity and capital into banks and other financial institutions

\textsuperscript{34} We have not focused on the internal incentive structure in banks (which might be related to bank culture). As has become clear in the current crisis, internal risk management showed substantial lapses (see Group of Thirty, 2009). Other issues abstained from include procyclical in Basel II and IFRS accounting standards.
without an adequate corporate control role for the government—are very costly and possibly ineffective due to daunting moral hazard and asymmetric information problems. Some key lessons might be learnt from previous financial crises—for example, the Swedish financial crisis of the 1990s (see Ingves and Lind, 1994; and Aghion, Bolton, and Fries, 1999).

To conclude, we believe the most important, yet only partially answered, research questions raised by our discussion are the following:

- What are the implications of the ever-increasing integration of banks and markets for *systemic* risk and fragility?
- What will fintech developments, including P2P lending, portend for banks?
- What issues should we consider in the optimal design of regulation to respond to the (until recently, at least) growing cross-border footprints of major financial institutions and the increasing integration of banks and financial markets?
- What changes, if any, should be imposed on the structure of the financial services industry, and the banking sector in particular, to contain the “mushrooming” nature of systemic risk concerns (i.e., to contain the scope of regulation and supervision)?
- What role, if any, can market discipline play in helping safeguard the stability of the financial sector?
- How do banks and private equity firms (and other non-banking financial institutions) interact and what implications does this have for the regulation of banks and financial markets?
- What role do credit rating agencies play in financial markets, how does this affect banks, and what implications does this have for systemic risks that *bank* regulators care about?
These questions represent a rich agenda for future research.
REFERENCES


920.

Financial Studies* 10, 693–733.

*Journal of Finance* 55, 679–713.

Enforcement Problems and Banks: Intermediation as Credibility Assurance, *Journal of
Banking & Finance* 15, 605–632.


Economy*, No. 11, International Center for Monetary and Banking Studies, Geneva.


Butzbach, O. (2016). Systemic Risk, Macro-Prudential Regulation and Organizational


European Council (2012). Towards a Genuine Economic and Monetary Union. Report by the President of the European Council, Herman Van Rompuy, June..


US Senate Hearings (2002). Rating the Raters: Enron and the Credit Rating Agencies. Hearings before the Senate Committee on Governmental Affairs, Washington, DC.

