COMPETITION AND STABILITY IN BANKING

Xavier Vives
IESE Business School

Banking went from being one of the most regulated sectors in the economy after the crisis in the 1930s, to a more lightly regulated sector with the liberalization process that started in the 1970s in the United States. The previous period was marked by few crises, with much more instability in the second, culminating in the 2007 subprime crisis. In the first period, competition was considered detrimental to stability and in many countries competition policy was not applied fully to this sector until recently, despite its importance within the economy and the costs and inefficiencies induced by financial repression. Indeed, central banks and regulators were often complacent about collusion among banks, preferring to deal with a concentrated sector characterized by soft rivalry.

This changed with deregulation and the idea that competition enhances efficiency, be it productive, allocative, or dynamic (innovation). Competition policy is now taken seriously in the banking sector.¹ However, crisis hit in 2007, starting with subprime mortgages and then becoming systemic after the demise of Lehman Brothers in September 2008. Cumulative banking losses are estimated at 1.1 trillion euros (through November 2009) and massive...
bailouts (state aid with commitments involving public intervention in the European Union and United States) of up to 30 percent of GDP have overridden competition policy concerns. Indeed, public help programs have distorted competition and created an uneven playing field in terms of the cost of capital and perception of safety and soundness. Market power concerns about mergers have been also overruled. In the United Kingdom, Lloyds TSB took over the troubled HBOS (merger of Halifax and Bank of Scotland) in a merger opposed by the Office of Fair Trading, thereby creating a large entity, while the same Lloyds TSB had not been allowed to take over Abbey in 2001. The investment banking business has been consolidated in the United States, with the forced takeovers of Bear Stearns by J.P. Morgan, and Merrill Lynch by Bank of America. The result is potentially weak competition among the remaining players. Those events have deepened a current trend toward increased consolidation within countries, across countries, and across business lines (for example, forming financial conglomerates).²

Banking and financial markets display the whole array of classical market failures, due to externalities (fragility due to coordination problems and contagion), asymmetric information (excessive risk taking with agency problems, moral hazard and adverse selection), and potential market power. This has brought in regulation to protect the system, small investors, and market competitiveness. The problem is that the lender of last resort, deposit insurance, and too-big-to-fail policies introduce further distortions and exacerbate excessive risk taking. In fact, the crisis has uncovered massive regulatory failure and potential contradictions between regulatory intervention and competition policy.

This paper takes stock of what we know about the relationship between competition and stability, and suggests how to deal with the interplay of regulation and competition issues in banking in the aftermath of the systemic crisis started in 2007.

Section 1 examines trends in the banking sector and its regulation, taking into account the impact of the crisis. Section 2 explains the uniqueness of banks, why the banking system is fragile, and the role of regulation.³ Section 3 examines the trade-off between competition and stability in banking from a theoretical perspective, from the perspective of both fragility and the potential for excessive risk taking. Section 4 surveys the empirical evidence available. Section

². See, for example, Group of Ten (2001).
³. Sections 2 and 3 are partially based on Vives (2001, 2006).
Competition and Stability in Banking

5 ponders whether we can regulate away the competition-stability trade-off. Section 6 examines the policy response to the crisis, and concluding remarks close the paper.

1. Trends in the Banking Sector

Two periods can be distinguished in the recent history of the financial sector. The first, from the 1940s up to the 1970s, was characterized by tight regulation, intervention, and stability, while the second was marked by liberalization and greater instability (see figure 1). From the 1940s to the 1970s, competition between financial institutions was severely limited by the regulation of rates, activities, and investments; the separation of commercial banking, insurance, and investment banking (through the Glass-Steagall Act of 1933 in the United States); restrictions on the activity of savings banks; and geographical segregation (in the United States). Universal banking remained in some European countries. Deposit insurance was established, and the central bank acted as lender of last resort to the financial system.

---

4. The Glass-Steagall Act prohibited any one institution from acting as any combination of an investment bank, a commercial bank, and/or an insurance company.
The stability of this earlier period contrasts with a sizeable increase in the number of failures and crises in the later period, in which the sector was liberalized and competition introduced.\textsuperscript{5} International evidence points to liberalization as one of the factors behind banking crises, together with inadequate macro policies, adverse macro shocks, and vulnerability of the foreign sector. That is, liberalization, even controlling for a wide range of factors, increases banking fragility. There are also strong indications that the institutional environment (for example, in terms of the rule of law and contract enforcement) and inappropriate regulation that accompanies liberalization reinforce the development of crises.\textsuperscript{6} This is consistent with banking crises in diverse places, among them the United States (the savings and loan crisis), Japan, Scandinavia, and Spain. In all these cases, regulatory failure seems to have played an important role in the crisis.\textsuperscript{7} Despite these crisis episodes, financial liberalization has generally contributed to financial development, and therefore output growth.

Liberalization involved the lifting of controls on rates and banking investment activities, of geographical restrictions (for example, the Riegle-Neal Act of 1994 in the United States), of compulsory investment coefficients, and a convergence among the activities of different types of institutions (for example, savings and ordinary banks, commercial banking and investment banking, and, with the final repeal of the Glass-Steagall Act, through the 1999 Financial Services Modernization Act, between banking and insurance, at least to a point).\textsuperscript{8} Behind the process of liberalization and deregulation we

\textsuperscript{5} See Reinhart and Rogoff (2008a, 2008b).

\textsuperscript{6} See, for example, Demirgüç-Kunt and Detragiache (1998, 2001).

\textsuperscript{7} See Dewatripont and Tirole (1994) for the U.S. case and Hoshi and Kashyap (2000) for Japan. In both cases increased competitive pressures on financial institutions (for example, competition from non-bank intermediaries allowed by deregulation) leads to overexpansion in risky lines of business (for example, real estate), which are not checked, because of lax supervision and regulatory forbearance combined with implicit protection of entities in trouble. In Scandinavia, the roots of the early 1990s crisis lie in a set of factors following the financial liberalization of the 1980s: lax enforcement of capital requirements, poor supervision, lack of internal risk control methods, together with mistakes in fiscal and monetary policy in the context of an asset price bubble (see, for example, Honkapohja, 2009). In Spain, financial liberalization started in the 1970s and the banking crisis of the first half of the 1980s is explained by the large impact of the economic crisis derived from the oil shocks, the close links of banks with industrial firms, lack of diversification of banks’ industrial portfolios, bad management and inadequate supervision (see Caminal, Gual, and Vives, 1990).

\textsuperscript{8} For example, Citicorp (a commercial bank holding company) merged with Travelers Group (an insurance company) in 1998 to form the conglomerate Citigroup.
find advances in information technology, in transaction processing (automatic teller machines, telephone and electronic banking), computer capacity, management techniques, and risk coverage (for example, the use of derivative instruments and securitization techniques). The liberalization of international capital movements and the general reduction in transport costs and barriers to trade—that is, financial globalization—were an integral part of the process.

Liberalization has resulted in an increase in competition, both within and from outside the banking industry, with banks facing direct competition from financial markets and the development of disintermediation and financial innovation. Market integration in Europe and elsewhere has contributed decisively to steeper competition in wholesale and investment banking. Interestingly, the share of assets held by banks relative to non-bank financial intermediaries is declining in developed economies (for example, in the United States through 2007), although bank assets are not declining relative to total financial assets because the share of non-bank intermediaries grows at the expense of directly held assets (see figure 2).9

Figure 2. Distribution of U.S. Financial Assets by Type of Financial Intermediaries

Source: Barth and others (1997), and updated data from Flow of Fund Accounts, Board of Governors of the Federal Reserve System.

The liberalization process has also resulted in a tremendous expansion of financial intermediation, with financial assets of intermediaries increasing sharply, when expressed as a percentage of GDP.¹⁰

**Figure 3. Relative Size of the U.S. Financial Sector and the Banking Industry**

This effective expansion of the financial market has implied that even banking, in spite of the advance of disintermediation, has grown in real terms (see figure 4). Before the 2007 crisis, banking was evolving from the traditional business of taking deposits and granting (and monitoring) loans to the provision of services to investors (investment funds/asset management, advice, and insurance) and firms (consulting, insurance, mergers and acquisitions, underwriting of equity and debt issues, securitization, risk management), and proprietary trading. In a financial conglomerate, we can distinguish a retail bank, an investment or corporate bank, asset management, proprietary trading, and insurance. The infamous model of “originate-and-distribute” banking, where banks try to get rid of credit risk by originating mortgage loans and quickly securitizing them, leaving the monitoring of mortgages in a limbo, is a good example of the evolving banking process.

¹⁰ In the United States, for example, this has risen from less than 100 percent in 1950 to a peak of more than 300 percent in 2007, with assets in banks rising from 50 percent to 100 percent in the same period. See figure 3 for details.
At the same time, even if banks created off-balance-sheet vehicles such as structured investment vehicles and asset-backed commercial paper conduits, in the end they were insuring them with liquidity lines. In any case, the financial margin made way for fee and commission revenue and there was a switch from investment in bricks and mortar (the branches) to investment in communication networks, information technology, and highly specialized human capital. Post-crisis, the financial margin has regained importance (if nothing else, because of the very low or zero interest rate policy of central banks) and the share of banks’ assets in financial assets is up from pre-crisis levels (see figure 2). The return to traditional banking is apparent in figure 4, which reveals that recently, commercial banking has grown more than total financial intermediaries, in terms of real assets.

Restructuring is taking the form of consolidation, with the number of banks declining from 1997 to 2007 in both the United States (down 22 percent) and Europe (E.U.-15 down 29 percent). In Europe, domestic and, more recently, cross-border, and in the United States, interstate mergers have prevailed. One result is that despite an increase in national concentration (United States) in the past 20 years (see figure 5

11. For example, Hypobank-Vereinsbank in Germany, UBS-SBC in Switzerland, BNP-Paribas in France, IMI-San Paolo and Crédito Italiano-Unicrédito in Italy, Santander-BCH to form BSCH and BBV-Argentaria to form BBVA in Spain. Exceptions are some cross-border deals in the Benelux and Scandinavia. Some cross-border mergers have failed because of political interference of national authorities. See Danthine and others (1999).
for assets), local concentration (measured by deposits in MSA and non-MSA counties) has, if anything, tended to decline (Berger, Demsetz, and Strahan, 1999; White, 2009, table 7).  

**Figure 5. U.S. CR5 Ratio**

![Graph showing U.S. CR5 Ratio](image)

Sources: Federal Deposit Insurance Corporation (FDIC) and Federal Reserve.
a. Share of the five largest depository institutions expressed as a percentage of total assets. The merger of Wells Fargo and Wachovia is accounted for in 2008.

In Europe, the prevalence of domestic mergers tends to increase local concentration (for example, in 19 of 27 E.U. markets, the CR5 ratio for assets was over 50 percent in 2007). Figure 6 presents the CR5 concentration ratio for the E.U.-15.

In the United States, the CR5 ratio for assets rose from 23 percent in 2001 to 36 percent in 2008 (with several post-crisis operations, including J.P. Morgan-Washington Mutual and Wells Fargo-Wachovia). This contrasted with a gentler shift in the E.U.-15, from 52 percent to 54.5 percent (unweighted average) and from 37.6 percent to 44 percent (weighted average) in the same period.

In short, liberalization has come with an increase in competition amongst financial intermediaries, but bank assets over total financial assets have held steady, and the incidence of crises has risen. Meanwhile, banking has shifted significantly towards service

---

12. Metropolitan Statistical Areas (MSAs) are used as proxies for urban local markets, while non-MSAs are used as proxies for rural local markets.
14. The CR10 ratio for deposits rose from 36 percent in 2000 to almost 51.5 percent in 2008.
The introduction of competition in banking has come with checks on risk taking with capital requirements, allowing banks to rely on their own internal models to assess and control risk, and disclosure requirements for financial institutions, to improve transparency. The crisis marks a return to traditional banking and tends to exacerbate the consolidation trend.

Figure 6. Share of CR5 as a Percentage of Total Assets

Source: European Central Bank.

Figure 7. Herfindahl Index Based on Total Assets

Source: European Central Bank.
and foster market discipline. Flexible views of capital requirements, supervision, and market discipline have become the pillars of the Basel II framework. The rationale behind these reforms was to make capital requirements more risk sensitive. Supervisors would assess how well banks were matching their capital to risks assumed and banks would disclose information on their capital structure, accounting practices, risk exposures, and capital adequacy. In short, capital requirements plus appropriate supervision and market discipline were considered the main ingredients of a sound banking system. All this is under revision because of the crisis.

2. The Role of Banks, Fragility, and Regulation

Banks provide transaction and payment system services, insurance, and risk sharing (transforming illiquid assets into liquid liabilities). A central function of banks is to finance and monitor entrepreneurial projects that are illiquid and opaque due to asymmetric information problems, such as adverse selection and moral hazard. A lender needs relationship-specific skills to collect those loans that are illiquid, because the financed projects are opaque. Indeed, a main function of the banking and financial system is to overcome problems associated with asymmetric information in an economy.

2.1 Fragility and the Uniqueness of Banks

The essence of banks is that they create liquidity, but this leaves them vulnerable to runs. Banks protect entrepreneurs from the liquidity needs of depositors/investors. There are different versions of the story, but this is the cornerstone of modern banking theory (Diamond and Dybvig, 1983; Holmstrom and Tirole, 1997, 1998; Diamond and Rajan, 2001). The demand deposit contract, redeemable at par, creates a coordination problem for investors, which allows bankers to not extort rents on their abilities to collect illiquid loans (Diamond and Rajan, 2001) or disciplines bank managers subject to a moral hazard problem (Calomiris and Kahn, 1991; Gale and Vives, 2002). Because of asymmetric information, firms may get no funding because they do not have enough pledgeable income (fraction of their

16. Allowing banks to choose from a menu of approaches (for example, standardized and internal rating) to measure risk (credit, market, and operational).
Competition and Stability in Banking

return that can be committed to be paid to outsiders). Banks come to the rescue, for example, by creating liquidity-holding collateral and committing to make payments (Holmstrom and Tirole, 1997, 1998). In short, the standard deposit contract and loan provision to opaque entrepreneurial projects are complementary and central to a bank’s function.

At the base of the fragility of banking, there is a coordination problem of investors, who may decide to call back their short-term deposits or certificates of deposit and make a sound bank fail. The literature has presented two views of crises: the multiple equilibrium panic view (Bryant, 1980; Diamond and Dybvig, 1983) and the information-based view (Gorton, 1985, 1988; Jacklin and Battacharya, 1988). According to the former, sunspots (events unrelated to fundamentals) trigger runs, while according to the latter, bad news about bank assets triggers runs. Recently, these views have been reconciled with the introduction of asymmetric information and the identification of links between the probability of a run and the strength of fundamentals (Goldstein and Pauzner, 2005; Rochet and Vives, 2004). Thus, a solvent bank may be subject to panic, with depositors withdrawing funds invested and the bank forced to liquidate assets quickly, incurring a penalty. The cause of the problem is banks’ dependence on short-term debt.

Moreover, systemic risk can arise from contagion by a single entity’s failure and this may produce a strong, negative externality, affecting both the financial and real sectors of the economy. For example, through interbank market commitments, the failure of one entity may cause the downfall of others (see Allen and Gale, 2000). Similarly, large shifts in asset prices, such as an abrupt fall in the stock market or the failure of a major intermediary, may generate a domino effect and a systemic crisis, affecting the entire payment system.

Crises, however, may have positive aspects and in some circumstances can be optimal from an ex ante point of view, by making payment to depositors contingent on returns and improving risk.

17. Postlewaite and Vives (1987) provided an early model with a unique equilibrium, where the probability of a crisis is determined by the realization of the liquidity needs of depositors, which involves private information.

18. In section 3.1, I will examine the impact of competitive pressure on instability in this context.

sharing (Alonso, 1996; Allen and Gale, 1998) or by helping control the incentives of the banker (Calomiris and Kahn, 1991; Gale and Vives, 2002). For example, in the presence of moral hazard, incentive efficiency requires that the expected utility of investors/depositors be maximized, subject to the constraint that the bank manager exerts effort. This can be accomplished by liquidating the project, when observable interim returns are lower than a certain threshold (this is the minimal threshold that induces the manager to exert effort, a higher threshold would just increase the costs associated with liquidation). The threat of liquidation disciplines banks managers, but typically there is excessive liquidation and fragility (that is, “excessive” coordination failure; see Rochet and Vives, 2004). The reason is that competitive banking will typically reward investors over and above the optimal liquidation threshold. The challenge of regulation and supervision is to permit sufficient crises to keep the right incentives for bankers, while taking into account the degree of competition in the market. We will return to this question in section 6.

In short, banks are unique because of their particular mix of features: high (short-term) leverage, dispersed debtholders (implying a low level of monitoring), and opaque bank assets of long maturity, which exacerbate moral hazard, fragility and a high social cost of failure, and vulnerability to contagion (via interbank commitments or indirect market-based balance sheet linkages). All of these factors add up to enormous potential for systemic impact. At the same time, banks are central—indeed, essential—to the economic system. When banks stop functioning, so does a modern monetary economy.

The situation may be even worse in an emerging economy, where the role of banks is relatively more important, since asymmetric information problems are more acute and financial markets less developed. Banks and their monitoring capacity are, therefore, central to economic development, and any potential fragility may dramatically worsen downturns.

2.2 Market Failures and Regulation

Financial markets involve the whole range of major market failures: externalities, asymmetric information, and market power. The banking system’s inherent fragility leads to the failure of institutions, panics, and systemic crises that potentially have a major impact because of economy-wide externalities. The great depression of the 1930s and the subprime crisis are good examples, as were the financial crises
in the United States, Scandinavia, Mexico, East Asia, and Russia, all of which remind us of the potential for severe economic disruption. A bank’s failure hurts non-financial firms precisely because individual bank-firm relationships are valuable (Petersen and Rajan, 1994). In fact, even a contraction of bank capital may induce a credit crunch, with severe disruption to the private sector. A major market failure is, therefore, the lack of internalization by financial intermediaries of the social cost of bankruptcy and potential systemic risk. Contagion may occur because of network effects in the payment system, interbank market, or derivatives markets. Market liquidity and funding liquidity may interact, causing downward spirals.

Asymmetric information is in fact the raison d'être of financial intermediaries. However, as we will see in section 3.2, in a bank the agency problem leads to excessive risk taking, because of moral hazard and risk-shifting incentives. Adverse selection in credit and financial markets may lead to the failure of competition and even market breakdown. An unregulated market leaves small investors unprotected.

Imperfect competition is the norm and not the exception in banking. Very important frictions prevent banking from being perfectly competitive. Indeed, asymmetric information creates barriers to entry or results in competition not delivering efficient outcomes. For example, in the case of credit rating agencies, conflicts of interest due to the issuer-pays model, entry restrictions, and a failure of the reputation mechanism seem to produce a race to the bottom. Other sources of friction are switching costs, network effects (in retail banking, credit cards, or markets in general), and the ubiquity of two-sided competition in the banking sector.

Generally speaking, competitive banking will be excessively fragile, requiring policies such as lender of last resort (LOLR) facilities, deposit insurance, too-big-to-fail approaches and prudential regulation to rush to the rescue. These measures protect the system against negative economy-wide externalities. Regulation, meanwhile, aims to make banking and financial systems more stable, to avoid the negative effects associated with failing institutions and systemic crises. Regulation also aims to protect the small investor. Other policies promote a competitive system and maintain competitive markets (competition policy).

20. See Brunnermeier and Pedersen (2009) for a model of a liquidity spiral combining market and funding liquidity.
Financial regulation has side effects though. The most important one is potential moral hazard induced by protection and bailouts extended to failing institutions.\textsuperscript{21} The LOLR and deposit insurance are two basic instruments on which the stability of the banking system rests. Blanket insurance, however, is often offered to banks and depositors, according to a too-big-to-fail (TBTF) policy, which is usually justified by the potentially system-wide consequences of a large institution’s failure, but this help may reflect a problem of time-inconsistency. In the presence of moral hazard in the banking sector (for example, the banker’s level of effort in monitoring projects), a well-intentioned regulator will find it optimal to help ex post if this salvages the value of projects. Bankers, anticipating the help, will tend to exert suboptimal effort (see Gale and Vives, 2002). This is an example of the time-inconsistency problem facing a central bank. After the fact, costly liquidation of projects will not be optimal, so the central bank may be soft. The commitment problem is compounded by the bank manager’s interest in the bank continuing. Investing a central bank with a “tough” reputation can alleviate the time-inconsistency problem. Similarly, suspension of convertibility may remove incentives encouraging depositors to run (Diamond and Dybvig, 1983), but if the banking authority cannot pre-commit to such a deposit freeze and uses an ex post efficient (softer) intervention, this will encourage runs (Ennis and Keister, 2009).

In emerging markets, asymmetric information problems are more acute and reliance on the banking system to overcome them is more important. Moreover, these economies face a more severe policy commitment problem, which leads to excessive bailouts and potential devaluation of claims from foreign investors. This exacerbates moral hazard and provides a reason for importing external discipline (for example, acquiring foreign short-term debt). However, external discipline may come at the cost of excessive liquidation of entrepreneurial projects (the tradeoffs involved are examined in Vives, 2006).

3. Competition and Stability

Competition may influence stability, basically through the liability or asset side of a financial intermediary’s balance sheet. Competition,

\textsuperscript{21} We will come back to this issue in sections 3.2 and 5, where we examine some regulatory pitfalls.
in particular, may increase instability by (i) exacerbating depositors/investors’ coordination problem on the liability side, and fostering runs and/or panics, which may affect the system overall; and (ii) increasing the incentives to take risk (on either the liability or asset sides), thus increasing the probability of failure. I will examine each of these possibilities in turn. For (i), I will sketch a model, since it is not yet well understood.

3.1 Competition, Runs, and Fragility

The first thing to note is that competition is not responsible for fragility. Indeed, vulnerability to runs may emerge independently of market structure. This conclusion is based on work by Matutes and Vives (1996) with a model that combines Diamond’s banking model (Diamond, 1984) with a differentiated duopolistic structure à la Hotelling. In this model, depositors’ expectations determine the probability of a bank failing endogenously. These expectations are self-fulfilling, due to diversification-based scale economies: a bank that is perceived to be safer commands a larger margin and attracts a higher market share, allowing a better diversification. The model admits multiple equilibria, with corner solutions where only one bank is active or an equilibrium where no bank is active, for example during a system-wide crisis of confidence. This arises due to the coordination problem between depositors (as noted in the network externalities literature) and its presence does not depend on market structure. A monopoly bank may suffer a run. However, an increase in rivalry does increase the probability of failure in an interior equilibrium of the depositor’s game, where banks have positive market shares.22

Chang and Velasco (2001) present a model of financial crisis in emerging markets in the Diamond and Dybvig (1983) tradition. They find that financial liberalization increases the expected welfare of depositors, but may also increase fragility. Liberalization is modeled as moving away from a monopoly toward an increasingly competitive situation. A monopolist bank holds depositors to their reserve level, which implies that they remain indifferent to an autarchic system with no financial intermediation. The monopolist bank does so by reducing payments to depositors and therefore

---

its short-run liabilities. Profits act as a buffer against unexpected withdrawals. Consequently, the bank is less likely to fall within the range where a self-fulfilling crisis occurs than in a competitive situation. Furthermore, monopoly banking has to deliver a lower level of welfare, since a competitive bank maximizes depositors’ ex ante utility, taking into account the probability of a run, associated with an exogenous sunspot, and autarchy is a feasible allocation.\(^{23}\)

The recent reconciliation of the self-fulfilling theory of crises with the information- and fundamentals-driven views offers some insight into the competition-stability relationship, without having to resort to sunspot variables to explain how investors coordinate in equilibrium.

Traditional bank runs typically resulted from massive withdrawals by individual depositors. Modern bank runs are typically the outcome of non-renewal of short-term credit in the interbank market, as in the Northern Rock case or the 2007 run on structured investment vehicles (SIVs).

Let us consider a stylized banking crisis model based on Rochet and Vives (2004) and Vives (2010c). The model has three dates: \(t = 0, 1, 2\). On date \(t = 0\), the bank has equity \(E\) (or, more generally, stable funds including insured deposits) and collects uninsured certificates of deposit (CDs) or short-term uninsured debt worth \(D_0 \equiv 1\). These funds are used to finance risky investment \(I\) and cash reserves \(M\). The returns \(\theta I\) on these assets are collected on date \(t = 2\). If the bank can meet its obligations, the CDs are repaid at their face value, \(D\), and the bank’s equity holders obtain the residual (if any). A continuum of fund managers make investment decisions in the interbank market. At \(t = 1\) each fund manager, after the observation of a (conditionally independent) private signal about the future realization of \(\theta\) decides whether to cancel \((y_i = 1)\) or renew his or her CD \((y_i = 0)\). It is assumed that all random variables follow a Gaussian distribution with \(\theta \sim N(\bar{\theta}, \tau_\theta^{-1})\) and the private signal for investor \(i\) is \(s_i = \theta + \varepsilon_i\) with independent and identically distributed noise \(\varepsilon_i \sim N(0, \tau_{\varepsilon}^{-1})\) orthogonal to \(\theta\).

Let \(\bar{y}\) be the amount of withdrawals. If \(\bar{y} \geq M\), then the bank has to sell some assets to meet payments. A fund manager or investor adopts a behavioral rule of the type: cancel the investment if and only if the probability that the bank fails is above threshold \(\gamma \in (0, 1)\). This will occur, for example, if the fund manager is rewarded for

\(^{23}\) Todd Keister raised a similar point when discussing the paper.
taking the right decision (that is, withdrawing if and only if the bank fails).

Let $m \equiv M/D$ be the liquidity ratio; $\theta_L \equiv (D - M)/I$, the solvency threshold of the bank; $\lambda > 0$ the fire sale premium on early sales of bank assets; and $\theta_H \equiv (1 - \lambda)\theta_L$ the “supersolvency” threshold, such that a bank does not fail, even if no fund manager renews their CDs. Under these conditions the bank fails if $\theta < \theta_L$, or when $\theta \geq \theta_L$ but

$$\hat{\gamma} \geq m + \frac{1 - m}{\lambda} \left( \frac{\theta}{\theta_L} - 1 \right).$$

When taking into account the balance sheet constraint at $t = 0$, $E + D_0 = I + M$, we have $\theta_L \equiv (1 - m)/(\ell^{-1} + d^{-1} - m)$, where $\ell = D/E$ is the short-term leverage ratio and $d = D/D_0$ the return on short-term debt. An increase in the face value of debt $D$ may result from the bank facing a more competitive environment.

The model can be reinterpreted, replacing banks with countries and the short-term debt with foreign-denominated, short-term debt.

Investors, therefore, engage in a symmetric binary action game of strategic complementarities. If the state of the world is known, then if $\theta > \theta_L$, the dominant strategy will involve withdrawal; if $\theta > \theta_H$, then the dominant strategy will be to remain (not to withdraw); and for $\theta \in (\theta_L, \theta_H)$ both equilibria coexist. We can show that with incomplete information, an equilibrium is characterized by two thresholds $(s^*, \theta^*)$ with $s^*$ yielding the signal threshold below which an investor withdraws and $\theta^* \in [\theta_L, \theta_H]$ the state-of-the-world critical threshold, below which the acting mass of investors makes the bank fail. There are at most three equilibria. There is a critical liquidity ratio, $m \in (0, 1)$, such that $\theta^* = \theta_L$ for $m \geq \overline{m}$, and for $m < \overline{m}$ we have $\theta^* > \theta_L$. In this case, the equilibrium is unique if $\tau \tau \pi \varepsilon \theta \leq \sqrt{2\pi((\ell^{-1} + d^{-1} - m)\lambda)^{-1}}$. The reason is as follows. Let $R(\hat{s})$ be a player’s best reply threshold to the (common) signal threshold $\hat{s}$ used by other players. The game then involves strategic complementarities, with $R' \geq 0$. A higher threshold $\hat{s}$ applied by others induces a player to a higher threshold as well. We can show that if $\tau \tau \pi \varepsilon \theta \leq \sqrt{2\pi((\ell^{-1} + d^{-1} - m)\lambda)^{-1}}$, then

24. Related examples can be found in Morris and Shin (1998, 2004) and Corsetti and others (2006). In a game of strategic complementarities, the marginal return on a player’s action increases in the level of the actions of rivals. Best replies, then, are monotone increasing. See Vives (2005).

25. All the results presented in this section are to be found in Vives (2010c).
$R'(\hat{s}) \leq 1$. This ensures that $R(\cdot)$ crosses the 45º line only once and that the equilibrium is unique. In figure 8, the uniqueness case is illustrated by the flatter best reply curve and the three-equilibria case by the steeper best reply curve.

**Figure 8: Best Response of a Player to Threshold Strategy, $\hat{s}$, Used by Rivals**

![Diagram](image)

Multiple equilibria come about when strategic complementarity is strong enough (the steeper best response in figure 8). This is a function of the slope of the best response. The maximal value of the slope is $\left(\tau_\theta + \tau_\varepsilon\right) / \left[\tau_\varepsilon + (\ell^{-1} + d^{-1} - m)\lambda^{-1}\sqrt{2\pi\tau_\varepsilon}\right]$. Strategic complementarity will be larger in a more competitive situation ($d$ larger) and when the fire sales penalty $\lambda$ is higher. It will tend to be smaller when noise in the signals is small in relation to the prior, that is, $\tau_\theta / \sqrt{\tau_\varepsilon}$ small. With small noise in the signal, a player faces greater uncertainty about others' behavior, reducing complementarity.

At equilibrium with threshold $\theta^*$, when $\theta < \theta^*$, the acting mass of withdrawing investors make the bank fail and therefore the probability of a “crisis” occurring is $\Pr(\theta \leq \theta^*)$. A crisis occurs due to low values of fundamentals. In contrast, the complete information model contains multiple self-fulfilling equilibria in the range $({\theta_L}, {\theta_H})$. Consequently, the model bridges between the self-fulfilling theory of crisis (for example, Diamond and Dybvig, 1983) and theory linking crisis to fundamentals (for example, Gorton, 1985).

In the range $({\theta_L}, {\theta_L^*})$, coordination fails from the perspective of the institution being attacked. Thus, the bank is solvent but illiquid—that is, the bank would have no problem if only investors would renew their CDs, but it has fallen into a range where they do not, and thus becomes
illiquid. Thus, the risk of illiquidity is represented by $\Pr(\theta_L \leq \theta < \theta^*)$ and the risk of insolvency by $\Pr(\theta < \theta_L) = \Phi\left[\sqrt{\tau_0 (\theta_L - \bar{\theta})}\right]$, where $\bar{\theta}$ is the prior mean or public signal and $\Phi$ is the cumulative normal distribution $N(0,1)$.

Whenever $m < \bar{m}$ and there is a unique equilibrium, an increase in $d$ or $\lambda$ boosts both $\theta^*$ and $s^*$, the probability of crisis $\Pr(\theta < \theta^*)$, and the range of fundamentals $[\theta_L, \theta^*]$ for which there is coordination failure (Vives, 2010c). Any rise in bank vulnerability, whether it affects the face value of bank deposits or the fire sale premium for early liquidation, increases fragility by increasing the degree of strategic complementarity. Furthermore, if released, the public signal $\bar{\theta}$ has a negative multiplier effect on equilibrium thresholds, which is enhanced if $d$ or $\lambda$ are higher. Indeed, the equilibrium signal threshold is determined by $R(s^*; \bar{\theta}) - s^* = 0$. From this, it follows that:

$$\frac{ds^*}{d\bar{\theta}} = \frac{\partial R/\partial \bar{\theta}}{1 - R'} > \frac{\partial R}{\partial \bar{\theta}}$$

whenever $R' < 1$ is met, since $R' > 0$. As a result, a rise in $\bar{\theta}$ will affect the equilibrium threshold $s^*$ more than the direct impact on the best response of a player $\partial R/\partial \bar{\theta}$. This multiplier effect is largest when $R'$ approaches 1, that is, when strategic complementarities are strong, and we approach the region of multiplicity of equilibria. This is so when $d$ or $\lambda$ are large. Public information has a coordinating potential beyond its strict information content (as emphasized by Morris and Shin, 2002). Every investor knows that an increase in $\bar{\theta}$ will shift the best replies of other investors downward and everyone will be more cautious about withdrawals.

Consistent with this result, experimental evidence reveals that bank runs occur less frequently when banks face less stress, in the sense of a larger number of withdrawals being necessary to induce insolvency. The presence of market power in the interbank market may either facilitate liquidity provision—because liquidity is a public good, so sound banks may have an incentive to provide liquidity to

26. Goldstein and Pauzner (2005) also show how increasing the deposit rate increases the probability of a run of depositors in a model of the global games type.
a bank in trouble to avoid contagion (Allen and Gale, 2004; Sáez and Shi, 2004)—or may impede its provision because banks with surplus funds underprovide lending strategically to induce fire sales of the bank-specific assets belonging to needy intermediaries (Acharya and others, 2010).

The comparative statics results hold even if there are multiple equilibria for the extremal (stable) equilibria. It can be shown that extremal equilibrium thresholds \((\theta^*, s^*)\) decrease with \(\bar{\theta}\) and with decreases in stress indicator \(d\) or \(\lambda\). Considering out-of-equilibrium adjustment in the form of best-reply dynamics where, at any stage after the perturbation from equilibrium, a new state of the world \(\theta\) is drawn independently and a player responds to the strategy threshold used by other players at the previous stage, a similar result holds, since the middle “unstable” equilibrium becomes irrelevant. Interestingly, the region of potential multiplicity \(\tau \tau \pi \lambda \varepsilon \theta \leq \sqrt{2\pi} (\ell^{-1} + d^{-1} - m) \lambda^{-1}\) is enlarged with an increase in stress indicator \(d\) or \(\lambda\) and/or an increase in the precision of the public signal in relation to the private ones \(\tau_\theta / \sqrt{\tau_c}\).

Regulation in the form of solvency and/or liquidity requirements may help to control the probabilities of insolvency and illiquidity (Vives, 2010c). Indeed, the probability of insolvency \(\Pr(\theta < \theta_L)\) is decreasing in \(m = M / D\) (assuming that \(1 - \ell^{-1} - d^{-1} < 0\), as is usual in a commercial bank), the solvency ratio \(\ell^{-1} = E/D\), and \(d^{-1}\), since \(\theta_L \equiv (1 - m)/(\ell^{-1} + d^{-1} - m)\). The probability of a crisis \(\Pr(\theta < \theta^*)\), including the probability of illiquidity, is decreasing in \(m\), \(\ell^{-1}\) (and also in \(d^{-1}\) and \(\lambda^{-1}\)) since \(\theta^*\) is as well.

From this it follows that both solvency and liquidity requirements required to control the probability of insolvency and illiquidity may have to become tighter in a more competitive environment, where \(d\) is higher. Furthermore, the liquidity requirement may have to become tighter when \(\lambda\) is higher.\(^{28}\) However, note that there is a partial substitutability between \(m\) and \(\ell^{-1}\), since they both contribute to reducing \(\theta_L\) and \(\theta^*\). In the limit case of almost perfect signals, \(\tau_c \rightarrow \infty\), which allows for a closed-form solution, we can check that in a more competitive environment (with higher return on short-term debt \(d\)), the solvency requirement (but not the liquidity ratio) should be strengthened, while in an environment where the fire sales penalty \(\lambda\) increases, the liquidity requirement must be strengthened, whereas the solvency requirement can be relaxed.

\(^{28}\) See Vives (2010a).
Consistent with these results, there is evidence that banks that relied less on wholesale funding and had higher capital cushions and liquidity ratios, fared better during the crisis.\textsuperscript{29}

In short, runs can happen independently of competition levels, but rising competitive pressure worsens investors’/depositors’ coordination problem, and increases the following: (i) potential instability (enlarging the multiplicity of equilibria region); (ii) the probability of a crisis; (iii) the range of fundamentals for which there is coordination failure of investors (and the institution is solvent but illiquid); (iv) the impact of bad news on fundamentals; and (v) the solvency requirement. Generally speaking, the socially optimal probability of crisis is positive because of its disciplining effect. These results, then, do not mean that competitive pressure should be minimized.

3.2 Competition and Risk Taking

Banks will have excessive incentives to take risk in the presence of limited liability (for shareholders and managers) and moral hazard (non-observable risk on the asset side). This is exacerbated by deposit insurance with flat premiums. The problem is particularly acute for banks close to insolvency/bankruptcy. Indeed, limited liability means that banks will take excessive risk on the asset side, unless the bank’s risk position can be assessed (for example, by large holders of CDs). A bank, then, cannot increase its market share and profits by taking more risk, because investors will discount it. However, introducing flat premium deposit insurance (or bailouts) destroys the market’s disciplinary effect, by eliminating investor concerns about potential bank failure.

Intense competition may worsen the excessive risk taking problem, because high profits provide a buffer and increase the bank’s “charter value.” In a dynamic setting, market power enhances the bank’s charter value, making it more conservative. Indeed, a bank with more market power enjoys higher profits and has more to lose if it takes more risk, fails and its charter is revoked. If future profits weigh enough, the bank will moderate its risk taking. Besanko and Thakor (1993) make this point with reference to the value created through relationship banking, and Boot and Greenbaum (1993) with

\textsuperscript{29} See Ratnovski and Huang (2009) with evidence from the 72 largest commercial banks in OECD countries.
regard to reputational benefits, both of which may be eroded by more competition.\textsuperscript{30} Matutes and Vives (2000) consider an imperfect competition model where banks are differentiated, have limited liability, and failure involves social costs (which could include a systemic component). The authors show that deposit rates are too high when competition is intense and the social cost of failure high. If the risk assumed by bank investments is not observable, then the incentives to take risk become maximal. Flat premium deposit insurance tends to make banks more aggressive, by increasing the elasticity of the residual supply of deposits available to the bank (this is also the result in Matutes and Vives, 1996). Furthermore, with risk-insensitive insurance, deposit rates will be too high amidst intense competition, even with no social cost of failure and no discipline on the asset risk taken. Allen and Gale (2004) consider banks competing à la Cournot in the deposit market and choose a risk level on the asset side. With insured depositors they show that as the number of banks grows, banks have maximal incentives to take risk on the asset side.\textsuperscript{31}

With heterogeneous borrowers, tougher competition may lead to a riskier bank portfolio and higher probability of failure. This is because more rivalry may reduce incentives to screen borrowers (the bank has fewer informational rents; see Allen and Gale, 2004). A larger number of banks may also increase the chance that bad borrowers get credit by reducing each bank’s screening ability, due to the adverse selection/winner’s curse problem (Broecker, 1990; Riordan, 1993; Gehrig, 1998).\textsuperscript{32}

However, competition tends to push down the rates that firms pay for loans and may, therefore, improve the average quality of loan applicants and/or reduce the need to ration credit. For example, better terms for entrepreneurs mean that they earn more profits and become more cautious, thus reducing the likelihood of the bank failing (Caminal and Matutes, 2002; Boyd and De Nicolò, 2005). Martinez-Miera and Repullo (2008), however, show that this argument does not consider the fact that lower rates also reduce the banks’ revenues from non-defaulting loans. When this is accounted for, there is a U-shaped relationship between competition and the risk of bank

\textsuperscript{30} A better reputation reduces the cost of outside finance to the bank.

\textsuperscript{31} See also Hellman and others (2000) and Cordella and Yeyati (2002).

\textsuperscript{32} Note also that endogenous fixed costs due to information gathered via lending may induce a natural oligopoly in banking (Dell’Ariccia and others, 1999; Dell’Ariccia, 2001).
failure (in particular, when the number of banks is sufficiently large, the risk-shifting effect is always dominated by the margin effect). In summary, when both banks and firms have to monitor their investments, there is a potentially ambiguous relationship between market structure and risk taking.

A bank faces both adverse selection and moral hazard problems when lending to firms. A higher rate set by the bank will tend to draw riskier applicants (adverse selection) and/or induce the borrower firms, which also have limited liability, to choose riskier projects (moral hazard). We know that banks may then prefer to ration credit rather than raising the interest rate. A bank with market power has more incentive to alleviate this asymmetric information problem by monitoring the firms’ projects and establishing long-term relations with customers.\footnote{Besanko and Thakor (1993), Petersen and Rajan (1994, 1995).} This effect tends to increase firms’ access to credit. As usual, market power also increases the lending rate and therefore the tendency to ration credit to avoid an increase in the average risk for a pool of applicants. Even if we forget about the possibility of banking failure for a moment, market power presents a welfare trade-off, since more bank market clout reduces the bank’s moral hazard, but aggravates the problem for the entrepreneur. The result is that some market power tends to be good, unless monitoring is very costly. If banking failure is a possibility, then the analysis becomes more complex. Higher lending rates due to market power tend to depress investment and, under plausible assumptions with multiplicative uncertainty, decrease the bank’s overall portfolio risk. More rivalry should, therefore, increase the probability of bank failure. However, more competition may also destroy incentives to monitor, and thereby reduce lending. If the latter effect is strong enough, a monopolistic bank may end up more exposed to aggregate uncertainty (because it tends to ration credit less) and therefore more likely to fail.\footnote{Caminal and Matutes (2002).}

All in all, despite the complexity of the relationship between competition and risk taking, it seems plausible to expect that, once a certain threshold is reached, an increase in the level of competition will tend to increase risk taking incentives and the probability of bank failure. This tendency may be checked by reputational concerns, by the presence of private costs of managerial failure, or by appropriate regulation and supervision.

34. Caminal and Matutes (2002).
4. Evidence

Increased competition after liberalization and deregulation in the United States in the 1980s led banks to take more risks (Keeley, 1990; Edwards and Mishkin, 1995; Demsetz and others, 1996; Galloway and others, 1997). Keeley finds that a higher Tobin’s q (as a measure of charter value) was positively associated with high capital-to-asset ratios in U.S. bank holding companies for the period 1971–86. Furthermore, he finds that interest rates on large CDs for large bank holding companies between 1984 and 1986 were negatively related to q. It also seems that the increase in risk was held by large, TBTF banks in particular (Boyd and Gertler, 1993). There is controversy, however, over whether this increase in competition led to lower or higher loan losses (see Jayaratne and Strahan, 1998; Dick, 2007, respectively). Saurina and others (2007) claim that non-performing loans in Spanish banks fell as the loan market’s Lerner index rose. Salas and Saurina (2003) found that 31 years of liberalization measures in Spain increased competition and eroded banks’ market power (measured again by Tobin’s q), banks with lower charter values tended to have lower equity-to-asset ratios (lower solvency), and to experience higher credit risk (loan losses over total loans).

Liberalization in a weak institutional environment and/or with inadequate regulation shifts risk to the taxpayer and increases the likelihood of systemic crisis (Demirgüç-Kunt and Detragiache, 1998, 2001). A similar situation seems to have arisen in the wake of the subprime crisis, with declining lending standards associated with securitization (Dell’Ariccia and others, 2008).

The relationship between concentration and stability is complex. On the one hand, a concentrated banking system with a few large banks may be easier to monitor and banks are potentially more diversified. On the other hand, large banks may be TBTF, receive larger subsidies, and have incentives to take more risk. Furthermore, large banks tend to be more complex, harder to monitor, and more interdependent (increasing systemic risk). The evidence also points to a complex relationship between concentration and stability.

Several studies have attempted to provide cross-country evidence on the effects of liberalization and increasing competition on both individual and systemic bank failures. In a cross-country study of

35. However, a problem with their approach is that the risk premium in the Lerner index is a function of loan losses ratio, which is a measure of non-performing loans.
23 developed nations, Berger and others (2009) show that market power (as measured by the Lerner index and Herfindahl index for deposits or loans at the national level) increases banks’ loan portfolio risk but decreases overall risk, because banks with market power hold more equity capital. In a cross-country study of 69 nations (1980–97), Beck and others (2006) show that systemic crises are less likely in concentrated banking systems (measured by the three-firm concentration ratio on total assets, controlling for macro, financial, regulatory, institutional, and cultural characteristics) and that fewer regulatory restrictions (on entry, activities, facility for competition) are associated with less systemic fragility. This suggests that concentration is no proxy for competition and questions whether market power is really a stabilizing influence. The pertinent connection, however, is between concentration in relevant markets (which need not be directly linked to aggregate asset concentration) and competition. Furthermore, concentration is endogenous and more competition may increase concentration in a free entry world, since there is less room for entrants.\footnote{See Vives (2000).} In this sense, it should come as no surprise to find that both concentration and competition are positively associated to stability. Concentrated systems tend to have larger and better-diversified banks (controlling for the size of the domestic economy eliminates the relationship between concentration and crises), but no connection is found with the ease of monitoring banks. The message of Beck and others (2006) seems to be that “more competitive banking systems are associated with less fragility, when controlling for concentration.” Schaek and others (2009) reach a similar conclusion, using the Panzar-Rosse H-statistic as a proxy for competition, with data from 45 countries (1980–2005). These authors, however, find that concentration itself is associated with a higher probability of crisis.

In a cross-country study using individual bank data, Boyd, De Nicolò, and Loukoianova (2009) apply a model-based definition of stress or crisis to find that more concentration leads to a higher probability of a systemic shock, but no greater probability of government intervention. The authors claim that in the literature, indicators of banking crises are in fact indicators of government response to the crisis, and that these are predicted by base indicators such as sharp reductions in profits, loans, and deposits. These authors interpret the results in Beck and others (2006) as an indication that
more concentration leads to less intervention (more forbearance by regulators) and more systemic crises, and that fewer entry barriers lead to less intervention and fewer crises. In a cross-country study with individual bank data (in emerging economies and U.S. banks), Boyd, De Nicolò, and Jalal (2009) also find that more concentration increases the probability of bank failure and that competition fosters more willingness to lend. Using cross-country data (1973–2002), Shehzad and De Haan (2009) find that certain aspects of liberalization reduce the likelihood of systemic crises, provided there is adequate supervision.

Diversification can be achieved through mergers between financial institutions, but large banks need not be more diversified. Empirical studies in the United States find strong benefits of consolidation (improving profitability and production efficiency, and reducing insolvency risk) when the degree of macroeconomic (geographic) diversification increases (Hughes and others, 1996, 1998). Specifically, these authors find that geographic diversification offsets the tendency of larger banks to risk insolvency more when controlling for diversification. Expanding assets is associated with a less than proportionate increase in expected profit and a more than proportionate increase in risk. An expansion in asset size and the number of branches within the same state is associated with a more than proportionate increase in expected profit and a less than proportionate increase in risk. An expansion in asset size, branches and diversification across states is associated with an improvement in value efficiency and reduction of insolvency risk. Consolidation within the state reduces insolvency risk, but does not improve market value. It has also been claimed that greater consolidation has increased systemic risk in the United States, by looking at the positive trend of stock return correlations for large and complex banking organizations in the period 1988–99 (De Nicolò and Kwast, 2002).

Internationalization is a way to achieve diversification. Furthermore, allowing multinational banks into previously protected markets may increase the range of financial services offered in the domestic market and reduce margins. A side effect may be the erosion in the charter value of domestic banks, inducing them to take on more risk. Some observers have found that both cross-border banking and foreign bank entry have improved financial intermediation, fostered growth, and reduced fragility (see Claessens, 2006; Barth and others, 37. See also Demsetz and Strahan (1997).
Competition and Stability in Banking

This reflects the direct and indirect effects of domestic banks’ competitive reactions. Some evidence, however, points to mixed distributional effects of foreign bank entry. Detragiache and others (2008) find that foreign bank entry in poor countries may reduce private credit growth. Berger and others (2001) find that large foreign-owned institutions concentrate on large-scale projects and may leave out small firms. Still, large, well-capitalized foreign banks may provide stability to the domestic financial system of an emerging economy. Because the brand name and franchise value of the bank are at stake, the headquarters of foreign banks could be expected to help a subsidiary should a problem develop, but this need not hold for systemic problems (for example, the collapse of Argentina’s currency board).\textsuperscript{38} Moreover, even if foreign bank headquarters were willing to help, they may not do so at the optimal social level, since they will not take into account the external effects of their help. For example, the headquarters of foreign banks may want to limit their exposure to a country facing a currency crisis and could therefore tighten liquidity provision to branches or subsidiaries in that country. Finally, the incentives of a foreign lender of last resort and supervisor may not line up with local interests. A foreign supervisor will not consider the consequences (systemic or not) for domestic residents of restructuring a local branch or subsidiary, but only the consequences of a crisis of a subsidiary abroad in terms of systemic stability at home.\textsuperscript{39}

Finally, there is ample evidence that institutions close to insolvency have incentives to gamble for resurrection (for example, the savings and loan crisis).

It is worth noting that the financial crisis seems to have affected banks in countries with different concentration levels and market structures. Although it has been pointed out, for example, that concentrated banking systems like those in Australia and Canada have fared better in the crisis than unconcentrated ones, such as those in the United States or Germany, countries with concentrated systems, such as the Netherlands or the United Kingdom (retail banking), also ran into trouble. Moreover, other factors come into play: in Canada (and to a lesser extent Australia), bank funds come mostly from deposits and not the wholesale market and are subject to strict regulations. Reliance on non-interest income has also proved

\textsuperscript{38} Headquarters have to back the deposits in a branch, but need not do it for a subsidiary.

\textsuperscript{39} See Vives (2006) for further discussion.
to be a source of increased risk and vulnerability. By the same token, it is not evident that certain types of institutions have been more vulnerable than others. Both specialized investment banks (in fact, all the U.S. ones have collapsed or converted to commercial banks), insurance companies like AIG, and universal banks (UBS, Citigroup, or German and U.K. banks) have suffered.

In conclusion, the evidence points to the following:
— Liberalization increases the occurrence of banking crises, while a strong institutional environment and adequate regulation mitigate them.
— There is a positive association between some measures of bank competition (for example, low entry barriers, openness to foreign entry) and stability.
— The association between concentration and stability presents mixed results.
— Larger banks tend to be better diversified, but may also assume more risks.

5. CAN WE REGULATE AWAY THE COMPETITION-STABILITY TRADE-OFF?

We have seen how limited liability means that banks may assume excessive risks on the asset side, unless that risk position is observable and market discipline works. Disclosure requirements may help to uncover the bank’s risk position (or, more realistically, ensure better assessment). This is represented by the top row in table 1. If the bank’s asset risk position is not observable, then incentives to assume more risk increase considerably (second row of table 1), becoming maximal when risk-insensitive insurance is introduced, since it destroys monitoring incentives (third row in table 1). Risk-based deposit insurance moderates risk-taking incentives and undoes the bank’s limited liability charter, but banks may still take too much risk in the presence of a large social cost of failure, which they do not internalize (bottom row in table 1). In the top and bottom rows, an instrument such as capital requirements may effectively control risk taking, but in the middle rows we need to complement capital requirements with asset restrictions.

40. See Baele and others (2007), De Jonghe (2010), Demiğüş-Kunt and Huizinga (2010), and Ratnovski and Huang (2009).
The general trend in banking regulation has been to control risk-taking through capital requirements and appropriate supervision. Both risk-based deposit insurance and disclosure requirements have been proposed to limit risk-taking behavior. Advanced economies have tried to move towards the top and the bottom rows of table 1. This shift came with reforms to the 1988 Basel Accord on capital requirements to better adjust them for risk (Basel II). Capital requirements, supervision and market discipline are the three pillars on which the Basel II regulatory reform was based. Transparency has its limits, though. While introducing disclosure requirements for banks’ market positions is feasible, assessing the risk level of a bank’s illiquid loan portfolio is more difficult. Furthermore, more disclosure may in fact induce information-based runs among investors, generating instability.

The present crisis is a testimony to the failure of the strategy to move towards the top and bottom rows of the table. Disclosure and

<table>
<thead>
<tr>
<th>Banking regime</th>
<th>Risk-taking incentive</th>
<th>Liability (rates)</th>
<th>Asset (investment)</th>
<th>Necessary regulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free banking (observable risk/high disclosure)</td>
<td>Medium-low</td>
<td>Absent</td>
<td>Capital requirements</td>
<td></td>
</tr>
<tr>
<td>Free banking (unobservable risk/low disclosure)</td>
<td>Medium-high</td>
<td>Maximal</td>
<td>Capital requirements and asset restrictions</td>
<td></td>
</tr>
<tr>
<td>Risk-insensitive insurance</td>
<td>High</td>
<td>Maximal</td>
<td>Capital requirements and asset restrictions</td>
<td></td>
</tr>
<tr>
<td>Risk-based insurance</td>
<td>Low</td>
<td>Absent</td>
<td>Capital requirements</td>
<td></td>
</tr>
</tbody>
</table>

Source: Author’s compilation.

42. According to Basel II’s guidelines on capital requirements, banks can choose between a “standardized” approach in which external rating agencies set the risk weight for the different types of loans (say corporate, banks, and sovereign claims) or an internal-rating-based approach in which banks estimate the probability of default and also the loss given default, in an advanced version of the method. The idea is to calibrate the capital requirement so that it covers the value at risk (expected and unexpected) from the loan under some assumptions.
risk assessment have been deficient because of the failure of rating agencies—among other reasons—and market discipline has been ineffective because of the blanket insurance offered by TBTF policies. Furthermore, capital regulation has not taken into account systemic effects (the social cost of failure) and asset restrictions have been lifted under the pressure of investment bank lobbies.\footnote{43}

We are stuck in the “risk-insensitive insurance” row with maximal risk-taking incentives. We need therefore to design appropriate capital requirements and asset restrictions. Optimal regulation would need a combination of risk-based insurance for deposits (which implies that insurance premiums are contingent on the rates offered by banks and their asset risk position, eliminating or exactly offsetting limited liability) and systemic capital charges that internalize the social cost of failure of banks. If banks’ asset risk position is not observable, then insurance cannot be contingent on it and banks will be induced to take maximal risk on the asset side. This will have to be controlled using asset restrictions (for example, separating banking and proprietary trading/investment banking activities). Furthermore, the appropriate level of the systemic capital charge will depend in general (in an increasing way) on the intensity of competition, and will be binding in a low or medium friction environment.

According to Matutes and Vives (2000), the capital requirement level is an increasing function of both the social cost of failure $K$ and the intensity of competition (inverse friction) in the market ($\lambda$, which in the model goes from maximal differentiation $\lambda = 0$ to no differentiation $\lambda = 1$). This is because typically the level of friction is not only a behavioral parameter but one that enters the utility function.\footnote{44} In this case a capital requirement should be set as a function of the level of $\lambda$. This result is consistent with the analysis in section 3.1, which requires that the solvency requirement be tightened in a more competitive environment.

Figure 9 depicts the regions in the space of intensity of competition $\lambda$ (with $\lambda = 0$ for an independent monopolies situation

43. The fact that financial regulation is subject to strong lobby pressure is well known. Kroszner and Strahan (1999), for example, document its role in the abandonment of branching restrictions in the United States.

44. For example, customers value differentiation, a source of friction and market power. Thus, an increase in differentiation means that banking customers will value the volume offered by the bank more and therefore a more lenient capital requirement becomes appropriate.
and $\lambda = 1$ for perfect competition) and social cost of failure $K$ for which it is optimal to disintermediate (“optimal disintermediation”), and for which deposit rates are too high ($r' > r^0$) or too low ($r' < r^0$), from the welfare point of view. For a given level of competition $\lambda$, if $K$ is very large it is optimal to disintermediate, if $K$ is intermediate then banks are too aggressive, taking too much risk on the liability side, and a binding capital charge (as a function of $\lambda$) should be imposed. If $K$ is low, then banks are not aggressive enough, the capital charge will not be binding, and an increase in competitive pressure would be welfare-enhancing. Competition policy pressure is needed in a high friction environment.

**Figure 9: Comparison of Market and Optimal Deposit Rates, as a Function of the Friction in the Market ($1/\lambda$) and Social Cost of Failure ($K$)**

With precise knowledge of $K$ and $1/\lambda$ the competition-stability trade-off can be regulated away. Just set up the appropriate capital charge and let banks compete (with the usual enforcement of competition policy).

The competition-stability trade-off also applies to emerging economies. An emerging market economy is characterized by high uncertainty, increased likelihood and incidence of financial and currency crises, a predominantly financial role for banks, and a weak supervisory structure. These characteristics make it much more difficult to apply the regulatory strategies applied
by developed countries in an emerging market economy. This reflects several factors. First, moving toward a disclosure strategy is more difficult, because information problems are more acute and producing information is more expensive. Second, risk-based deposit insurance can work only when insurance can be priced according to objective indicators of bank risk, which will be more difficult to obtain in an emerging market economy (and we know that even in a developed economy they can be hard to get). This makes it harder to move toward a risk-based insurance strategy, as the potentially problematic application of Basel II shows. It follows that banking and financial market regulation must be adapted for emerging market economies. These will tend to have higher project liquidation costs and social costs of failure, and a higher level of friction. The first two factors will push policymakers to tighten the regulations, while higher friction may pull them in the opposite direction.

In conclusion, the trade-off between competition and stability is complex, but seems real (at least along some dimensions). Well-designed regulation may alleviate this trade-off, but needs to consider it. Specifically, this means capital requirements that allow for systemic externalities must be adjusted to the level of friction in the market, becoming tighter when competition is more intense. In a world where fine-tuning regulations is difficult (and the experience to date with banking regulation seems to confirm this), it seems unwise to try to completely eliminate market power in banking. This may have implications for the optimal degree of concentration, which is likely to be intermediate. In emerging economies, optimal policy needs to carefully balance the impact of the different levels of friction and social cost of failure. In any case, it is clear that competition should be limited for institutions close to insolvency. This should be done in a framework that permits prompt corrective action, allowing the supervisor to intervene as soon as red flags indicating depleting capital go up.45

45. In fact, according to the U.S. Federal Deposit Insurance Corporation Improvement Act of 1991 (FDICIA), when solvency falls below a certain limit, the bank cannot expand its assets. A further decline in solvency may trigger the need to recapitalize or even the imposition of rate ceilings. The FDICIA seeks to reduce regulatory discretion through rigid intervention rules, which are gradually applied (see, for example, Dewatripont and Tirole, 1994).
6. THE POLICY RESPONSE TO A FINANCIAL CRISIS\textsuperscript{46}

6.1 Interventions and distortions

When a systemic crisis hits, the pressure to stabilize the system is tremendous. In the 2007–08 crisis, we saw an array of interventions: asset purchase and guarantee schemes (including extensions of deposit insurance, and guarantees in the interbank market and in mutual funds), capital injections, outright nationalization, and forced mergers. These interventions represent a large distortionary potential in terms of moral hazard, long-term effects on market structure, protection of inefficient incumbents, and creation of an uneven playing field (among different institutions and different countries). For example, TBTF institutions receiving help may end up with lower capital costs than others (not only in the short term, but also in the long term, because of the implicit guarantee involved). The result is that, ex ante, the incentives are to take excessive risk. This is compounded by subsidy races to help national champions and marketplaces. This effect is particularly apparent in the European Union, posing a threat to the single market. The help provided to the system may foster regulatory forbearance to cover losses. There is indeed evidence that regulatory forbearance is prevalent and that government is less likely to close or take over failing banks when the sector is weak: the cases of the savings and loan crisis in the United States, Japan’s banking crisis, and evidence on 21 emerging countries (Brown and Dinç, 2010). Finally, help to banks spills over into other sectors that demand more help (such as car manufacturing).

The crisis has brought forced mergers backed by government subsidies and/or guarantees.\textsuperscript{47} The upshot is that surviving incumbents enjoy more market power and lower capital costs, because they are

\textsuperscript{46} This section is based on Vives (2010b).

\textsuperscript{47} In the United States, backed by the Federal Reserve, Bear Stearns merged with J.P. Morgan in March 2008. Later that year, J.P. Morgan acquired the banking assets of Washington Mutual from the Federal Deposit Insurance Corporation, and Merrill Lynch merged with Bank of America (thereby exceeding the 10 percent national market share deposit threshold established by the Riegle-Neal Act of 1994, as did Wells Fargo when it acquired Wachovia in 2008). In the United Kingdom, the merger of HBOS and Lloyds TSB was approved against the Office of Fair Trading’s opinion (with partial nationalization), despite the merged entity ending up with a 30 percent market share in current accounts/mortgages and competition problems in small and medium enterprise banking services in Scotland. It is worth noting that Lloyds was not allowed to take over Abbey in 2001.
TBTF (and/or because of the public help). Remember that merger policies affect both competition and dynamic incentives. The takeover of a failed bank may reward an incumbent with temporary, monopoly rents, inducing monopoly inefficiency but prudent behavior. This is optimal only if subsequent entry is facilitated (Perotti and Suarez, 2002). The danger now is that incumbents increase their market power and are protected from new entries. A merger policy must have a long horizon, and even in a crisis situation, must consider the optimal degree of concentration in the industry, dynamic incentives for incumbents to be prudent, and ease of entry.

State intervention and even outright ownership have been necessary to stabilize the system. Indeed, when the taxpayer is footing the bill, the public sector must have a say in how the institution receiving help is run. Government ownership is distortionary, however: government sits on both sides of the regulatory relationship; if not disciplined by competition, it makes the banking system less efficient and encourages inefficiency, less financial stability, higher risk exposure, and more bank losses (Barth, Caprio, and Levine, 2004; Caprio and Martinez Peria, 2002; De Nicolò and Loukoianova, 2007); and political objectives and incentives rule. It also eliminates the market for corporate control, creates an uneven playing field with implicit and explicit guarantees, and leads to less competition and lower financial development.

In a crisis, policy makers must walk a tightrope between the supportive measures necessary to avoid contagion and ensure stability, and the desire to nourish vigorous competition over the long term. Some trade-off between the two objectives, particularly in the short-term, is unavoidable. When a systemic crisis strikes, there is little time to react and support measures must be implemented very quickly. Central banks, regulators, and fiscal authorities provide the support measures and the competition authorities must watch for distortions affecting competition (including the formation of market structures that hamper competition).

Help to a bank typically provides a positive externality to other banks, since it limits the spread of a crisis and protects the system, mostly by avoiding contagion, be it informational or because of

48. The evidence presented by Hau and Thum (2009) on board incompetence in German public banks (there is an enormous difference between private and public sectors, in terms of education, financial and management experience in the 29 largest German banks) is suggestive, linking larger losses borne in the crisis to the lack of professionalism in boards.
interbank exposures. This does not distort competition if it is liquidity help that allows a fundamentally sound bank to avoid contagion and ride out the turbulence. If the bank in distress has a solvency problem, then it should be restructured and help should come with strings attached, so that competition is not distorted by “bad” banks displacing “good” ones in customer business. The counterfactual for evaluating whether help is distortionary has to consider what would have happened if there had been no coordination failure among investors from the perspective of the distressed institution, that is, by removing the panic component from market behavior. This is not easy, particularly when compounded by regulatory failures that induce excessive risk taking.

The main tools of intervention to limit distortions are structural (asset divestitures) and behavioral (pricing, advertising, acquisitions) restrictions. Structural commitments may help reduce the post-crisis overcapacity in the banking sector, accumulated during the asset boom in many countries. Indeed, an added component in the present crisis is the extent of overcapacity in the banking system. The period of expansion at low interest rates has led banking to over-expand via credit, particularly in those countries where there has been a real estate bubble (United States, Ireland, United Kingdom, Spain). This means that branches and personnel must be cut, together with the balance sheet, even if credit is normalized (because it should stabilize below the pre-crisis bubble levels). In general, care must be taken to ensure that any commitments, whether structural or behavioral, leave the restructured bank a viable competitor. This is obvious if the bank is a fundamentally sound one. If it is not, then restructuring should prevent the bank from taking over business from healthy rivals that have not enjoyed help. In either case, the restructured bank has to be a viable competitor. To check moral hazard, it is important to remove the imprudent management of the institution receiving help. In this case, the behavioral restrictions on the helped bank could be relaxed.

6.2 Approaches in the United States and in the European Union

The role of the competition authority in the United States has been different from that of the European Union, because the E.U. competition authority has the unique capability, among competition authorities, to control state aid. Since the crisis, the European Union
has dealt with many banking aid cases, taking 22 decisions in 2008 alone, and 81 decisions in total as of 17 December 2009. Seventy-five of the cases were approved without objection.\textsuperscript{49} The European Union has explicit conditions for state guarantees and recapitalization, which have been formalized as temporary guidelines on restructuring aid to banks. The conditions imposed on helped institutions are mostly sensible, since they try to minimize the distortions introduced by public help, in particular for fundamentally unsound institutions.\textsuperscript{50} The European Commission has toughly imposed or influenced significant balance sheet reductions and behavioral restrictions on helped entities such as ING, Northern Rock, RBS, Commerzbank, or WesLandesbank. Interestingly, in the case of RBS, which has been ordered to sell some retail operations, insurance, and its commodity-trading business, the Commission mentioned concentration concerns, with RBS being the leader in retail and corporate banking for small and medium-sized enterprise segments. It also mentioned the benefits of divestment, in terms of limiting moral hazard in the insurance and commodity-trading business.\textsuperscript{51}

Some measures can be understood as efforts to minimize competitive distortions of the aid, and others in terms of restraining moral hazard in the future. In principle, the role of the competition authority is to preserve competition and not to limit moral hazard, which is the regulator’s role. The important point is that measures focusing solely on competitive distortions will also affect ex ante incentives (and moral hazard), since a bank will know that in the event of trouble, help will come with restrictions. This connects to the TBTF issue. More broadly, the concept of competitive distortion may address the issue of competition for those enjoying the advantages of a TBTF umbrella. In this sense, restrictions on business activities that fall outside regulated, core banking business may make sense, although they go beyond the standard competition concerns and

\textsuperscript{49} Sixty-six more cases have been cleared under a temporary framework to support lending to firms. See the European Commission’s Directorate General for Competition’s State Aid report of 17 December 2009, which includes an overview of national measures adopted as a response to the economic and financial crisis.

\textsuperscript{50} There is a potential exception in the behavioral requirement, which implies a commitment to expand lending. This is contradictory to the restrictive behavior that some want to impose on institutions receiving help and may induce bad practices, since the business of a private bank is to lend and what has to be attacked are the causes behind why the bank is not lending.

\textsuperscript{51} See the European Commission’s Directorate General for Competition’s State Aid report of 14 December 2009, in which the Commission approves an impaired asset relief measure and restructuring plan of RBS.
Competition and Stability in Banking

analysis. The RBS case points to the need for coordination between the competition authority and the regulator.

The activism of the European Commission poses the question of the future competitive balance with U.S. banks that received assistance requiring no divestitures. This may prove important, particularly in those segments in which there is global competition. On the advice of Paul Volcker, the Obama administration is advocating limits on size and scope (mostly proprietary trading) of banks, to avoid the too-big-to-fail problem and control risk-taking. Thus, the United States may accomplish through regulation what the European Commission is trying to accomplish through state aid controls. An important side benefit of state aid control in the European Union is that it limits bankers’ incentives to take excessive risk, under the expectation of a bailout if things go wrong. Thus, it addresses the TBTF issue. The competition authority may internalize the fact that if an institution that fails gets help, competition will be distorted. The option of limiting the size (or better, the systemically corrected size) of an institution that breaks up once it has received public help (something that the European Union seems to be implementing) expands the realm of competition policy. When ordering divestitures, however, the competition authority should not take into account systemic considerations. So far, the United States seems to be following another route, where TBTF is explicitly not an antitrust problem (see White, 2009).

In any case, size and scope restrictions are a blunt instrument to deal with the TBTF issue. Controls on size are problematic, because interconnectedness and business specialization are more relevant to systemic risk. In terms of scope, conflict of interest is what leads to potential market failure and should be the focus of any limitations. Higher capital and insurance charges for systemically important institutions, together with effective resolution procedures, may be a better approach to the problem. This should be coupled with a serious look at conflicts of interest in financial conglomerates. The upshot is that in its role of protecting competition, the competition authority may have a say in the TBTF issue and therefore its actions should be coordinated with the regulator. The potential for competition policy to provide a commitment device to partially address TBTF issues should not be dismissed.

The Obama administration’s move is reminiscent of the

52. See White House (2010) for details.
nineteenth century antitrust tradition of looking at large firms with suspicion because of the excessive concentration of power. Later on, antitrust evolved with size becoming less important as attention shifted to the issue of market power in particular markets. The influence that investment banks have had in the deregulation of financial intermediaries and the huge rise in leverage that ensued and led to the crisis is backfiring. We have entered the territory of political economy and the question is how to best control excessive concentrations of power in a democratic society.

7. SUMMARY AND CONCLUSIONS

Liberalization has come with an increase in the competition facing financial intermediaries and in the appearance of crises. Bank assets have not declined in relation to total financial assets and banking has shifted toward service provision. This restructuring has tended to increase aggregate concentration, although the consequences may have varied in relevant local retail markets in the United States and Europe. The crisis points to a return to traditional banking and may tend to exacerbate the consolidation trend.

Banks are unique because of their particular mix of features, which makes them vulnerable to runs with potentially systemic impact, and very important negative externalities for the economy. The fragility of a competitive banking system is typically excessive. Financial regulation comes to the rescue at the cost of side effects and regulatory failure. The most important one is the potential moral hazard induced by protection and bailouts extended to failing institutions. The present crisis is a testimony to the failure of the three pillars of the Basel II system. Disclosure and risk assessment have been deficient (think of the failure of rating agencies), and market discipline has been ineffective because of the blanket insurance offered by TBTF policies. Capital regulation has not taken into account systemic effects (the social cost of failure) and asset restrictions have been lifted under pressure from investment bank lobbies. Supervision has proved ineffective, since it has allowed a shadow banking system to grow unchecked.

Theory and empirics point to the existence of a trade-off between competition and stability along some dimensions. Indeed, runs happen independently of the level of competition, but more competitive pressure worsens the coordination problem of investors/depositors and increases potential instability, the probability of a
Competition and Stability in Banking

493

crisis, and the impact of bad news on fundamentals. This does not imply that competitive pressure has to be minimized, since in general the socially optimal probability of a crisis is positive because of its disciplining effect. On the asset side, once a certain threshold is reached, an increase in the level of competition will tend to boost risk-taking incentives and the probability of bank failure. This tendency may be checked by appropriate regulation and supervision. The evidence points to liberalization increasing banking crises, while a strong institutional environment and adequate regulation reduces them. At the same time, there is a positive association between some measures of bank competition (for example, low entry barriers, openness to foreign entry) and stability.

Regulation can alleviate the competition-stability trade-off, but the design of optimal regulation has to take into account the intensity of competition. For example, capital charges should reflect the degree of friction and rivalry in the banking environment, with tighter requirements in more competitive situations. Given that fine-tuning of regulation has proved very difficult in practice (this is probably an understatement given the massive regulatory failure that the crisis has uncovered), the trade-off between competition and stability is bound to persist, suggesting that coordinating regulation and competition policy is necessary. Banks’ uniqueness, not only during crises, should be recognized and the appropriate lessons drawn and applied during competition policy implementation.

The competition-stability trade-off also applies to emerging economies. An emerging market economy is characterized by high uncertainty, increased likelihood and incidence of financial and currency crises, the predominant financial role played by banks, and weak supervisory structures. These characteristics make it much more difficult to follow the regulatory strategy typically followed in developed countries. Emerging economies tend to have higher project liquidation costs and social costs of failure, and a higher level of friction. The first two factors tend to push for tighter regulations, while higher friction may pull in the opposite direction. In emerging economies, optimal policy should carefully balance the impact of the different levels of friction and the social cost of failure.

Merger policy in banking should be consistent over time and keep in mind an optimal degree of concentration and dynamic incentives (rewarding prudence and easing entry). How to deal with TBTF institutions remains an open issue. In the United States, TBTF is not an antitrust issue, whereas in the European Union the
competition authority controls distortions of competition which arise out of state aid, and this has implications for TBTF. The credibility of the competition authority to impose conditions once an institution has been helped may provide a commitment device which has been lacking in bank bailouts. Controls on size are problematic because interconnectedness and line of business specialization are more relevant to systemic risk than size. In terms of the scope of any bank's activities, conflict of interest is what leads to potential market failure and should be the focus for any limitations.

All this calls for close collaboration between the regulator (in charge of stability and prudential control) and the competition authority (in charge of keeping competition healthy). First of all, regulatory requirements and competition policy need to be coordinated. Capital charges may have to be fine-tuned to match the intensity of competition in different market segments. Second, a protocol for cooperation between the regulator and the competition authority should be developed. This is particularly important in crises. The competition authority can commit to addressing TBTF problems that lead to competition distortions; the regulator can address the TBTF issue and moral hazard through systemic capital charges, effective resolution procedures, and restrictions on the scope of banking activities that target conflicts of interest. Finally, crisis procedures should be established that define liquidity help from recapitalization and conditions for restructuring to avoid competitive distortions. Entities close to insolvency should be tightly regulated (and activities restricted) in a framework permitting prompt corrective action.

On the political economy of regulation, the debate over whether to let firms, banks in particular, get so large that they significantly influence regulation remains open.


Xavier Vives


Competition and Stability in Banking


