Competition and stability in banking

Xavier Vives
IESE Business School and CEPR

Introduction

The recent history of the financial sector can be divided into two periods. The first, from the 1940s up to the 1970s, was characterised by tight regulation, intervention, and stability, while the second was marked by liberalisation and greater instability (see Figure 1).

Such was the distaste for competition in banking that until recently competition policy in many countries was not applied fully to the sector, despite the economic costs and inefficiencies brought about by financial repression. All this began to change with deregulation and the idea that competition enhances efficiency, be it productive, allocative, or dynamic.

This later period of liberalisation reached crisis point in 2007. Starting with a disturbance in the subprime mortgage market, the instability spilled over to a full-blown global crisis following the demise of Lehman Brothers in September 2008. As of November 2009, the cumulative banking losses are estimated at €1.1 trillion.

The subsequent public bailouts of up to 30% of GDP have overridden the concern for competition policy, creating an uneven playing field in its place. In the UK for example, Lloyds TSB took over the troubled HBOS in 2008 – a merger opposed by the Office of Fair Trade – while the same Lloyds TSB had not been allowed to take over Abbey in 2001. The investment banking business has been consolidated in the US, with the forced takeovers of Bear Stearns, by JP Morgan, and Merrill Lynch, by Bank of America. The result is potentially weak competition among the remaining players. These events have deepened the trend towards increased consolidation within countries, across countries, and across business lines – often resulting in large financial conglomerates (Group of Ten 2001).

Regulation of the banking sector is nothing new. Indeed, banking and financial markets display the full array of classical market failures. Externalities arise from coordination problems and contagion, asymmetric information often leads to excessive risk taking, and extreme market power is common. In response, policymakers have introduced 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. The latest crisis has uncovered such regulatory failure on a massive scale, as well as potential contradictions between regulatory intervention and competition policy.

In the UK, Lloyds TSB took over the troubled HBOS in 2008 – a merger opposed by the Office of Fair Trade – while the same Lloyds TSB had not been allowed to take over Abbey in 2001.

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 global financial crisis.

Section 1 examines trends in the banking sector and its regulation. Section 2 explains the uniqueness of banks, why the banking system is fragile and the role of regulation. Section 3 examines the trade-off between competition and stability in banking from a theoretical perspective. Section 4 surveys the empirical evidence.

---

1 This paper was prepared for the Thirteenth Annual Conference of the Central Bank of Chile “Monetary Policy under Financial Turbulence,” Santiago, Chile, November 2009, and is to be published in the proceedings of the conference. I am grateful to the discussant of the paper, Todd Keister, to an anonymous referee for useful comments, and to Jorge Paz for his helpful research assistance. The research leading to these results received funding from the European Research Council under the European Advanced Grants scheme, Grant Agreement no. 230254. I also thank the support of the Abertis Chair of Regulation, Competition and Public Policy, project ECO2008-05155 of the Spanish Ministry of Education and Science at the Public-Private Sector Research Center at IESE.

2 Sections 2 and 3 are partially based on Vives (2001, 2006).
coefficients, and a convergence among the activities of different types of institutions. Behind the process of liberalisation and deregulation we 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 securitisation techniques).

Liberalisation has resulted in an increase in competition, both from within and 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.

The liberalisation 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. In the US for example, this has risen from less than 100% in 1950 to a peak of more than 300% in 2007, with assets in banks rising from 50% to 100% in the same period (see Figure 2).

Figure 2. Relative size of the US financial sector and the banking industry (Financial assets/GDP).

This effective expansion of the financial market has implied that even banking, in spite of disintermediation, has grown in real terms (see Figure 3).

Before the 2007 crisis, banking was evolving from the traditional business of taking deposits and granting loans, to the provision of services to investors (such as asset management) and firms (such as underwriting of equity and debt issues and securitisation), and proprietary trading. The model of “originate-and-distribute” banking offers an infamous example of the evolving banking process.

Yet even if banks created off-balance sheet vehicles, 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 specialised human capital. Post-crisis however, the financial margin has regained importance – if for no other reason than because of the very low or zero interest rate policy of central banks. Indeed Figure 3 reveals that commercial banking has recently grown more than total financial intermediaries, in terms of real assets, suggesting a return to traditional banking.

**Figure 3.** Growth Rates in U.S. Real Financial Assets (three year moving average: 1950-2008).

Restructuring is taking the form of consolidation, with the number of banks declining from 1997 to 2007 in both the US (down 22%) and Europe (EU-15 down 29%). In Europe, domestic and, more recently, cross-border mergers have been commonplace. In the US, interstate mergers have prevailed. One result is that despite an increase in national concentration (US) in the past 20 years (see Figure 4 for assets), local concentration (measured by deposits, in MSA and non-MSA counties) has, if anything, tended to decline (Berger et al. 1999 and Table 7 in White 2009). In Europe, the prevalence of domestic mergers has tended to increase local concentration (see Figure 5).

In the US, the CR-5 ratio for assets rose from 23% in 2001 to 36% in 2008 (with several post-crisis operations, including JP Morgan-Washington Mutual and Wells Fargo-Wachovia). 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 et al. 1999).

In Europe, the prevalence of domestic mergers has tended to increase local concentration (see Figure 5). The introduction on competition in banking has come with checks of risk taking with capital requirements, allowing banks to rely on their own internal models to assess and control risk. Another check has been disclosure requirements for financial institutions, to improve transparency and foster market discipline. A flexible view of capital requirements, supervision, and market discipline has become the foundation of the Basel II framework. The rationale behind these reforms was to make capital requirements more risk sensitive – supervisors would assess how well

This contrasted with a gentler shift in the EU-15, from 52% to 54.5% (unweighted average) and from 37.6% to 44% (weighted average) in the same period.7

**Figure 4.** US CR5 ratio. Share of the five largest depositary institutions expressed as a % of total assets.

Liberalisation has brought with it an increase in competition among financial intermediaries, but the incidence of crisis has risen. Meanwhile, banking has shifted significantly towards service provision, while restructuring has tended to increase aggregate concentration (although the consequences may have varied in relevant local retail markets in the US and Europe). The crisis marks a return to traditional banking and has tended to exacerbate the consolidation trend.

**Figure 4.** US CR5 ratio. Share of the five largest depositary institutions expressed as a % of total assets.

The introduction on competition in banking has come with checks of risk taking with capital requirements, allowing banks to rely on their own internal models to assess and control risk. Another check has been disclosure requirements for financial institutions, to improve transparency and foster market discipline. A flexible view of capital requirements, supervision, and market discipline has become the foundation of the Basel II framework. The rationale behind these reforms was to make capital requirements more risk sensitive – supervisors would assess how well

---

4 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 et al. 1999).

5 Metropolitan Statistical Areas (MSAs) are used as proxies for urban local markets, while non-MSAs are used as proxies for rural local markets.

6 The CR-10 ratio for deposits rose from 36% (2000) to almost 51.5% (2008).

7 See Schilsbach (2009), for the US.

8 The merger of Wells Fargo and Wachovia is accounted for in 2008.

9 Allowing banks to choose from a menu of approaches (for example, standardized and internal rating) to measure risk (credit, market and operational).
banks are 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. With the arrival of the global financial crisis, all this is now under review.

Figure 5. Share of CR5 as a % of total assets.


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, because of 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 building stone of modern banking theory (Diamond and Dybvig 1983, Holmstrom and Tirole 1997 and 1998, and 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 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, and 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.

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.

2.2 Market failures and regulation

Financial markets involve the whole range of major market failures: externalities, asymmetric information, and market power. But regulation is not without its costs either.

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 painful examples. A bank’s failure hurts non-financial firms precisely because individual

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.
bank-firm relationships are valuable (Petersen and Rajan 1994). One major market failure is the lack of internalisation 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 (see Brunnermeier and Pedersen 2009).

While asymmetric information is in fact the raison d'être of financial intermediaries, as we will see in Section 3.2 the agency problem can lead to excessive risk taking, because of moral hazard and risk-shifting incentives. At the same time, adverse selection in credit and financial markets may lead to the failure of competition and even market breakdown.

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.

Imperfect competition is the norm and not the exception in banking. 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 have produced a race to the bottom.

Generally speaking, competitive banking will be excessively fragile, requiring policies such as lender of last resort 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, and to protect the small investor.

Financial regulation is not without its side effects however. Chief among these is potential moral hazard induced by protection and bailouts extended to failing institutions. The lender of last resort and deposit insurance are two basic instruments on which the stability of the banking system rests. Yet such blanket insurance 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. A central bank with a “tough” reputation can alleviate the time-inconsistency problem, while 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 only 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). Yet such 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 primarily through the liability or asset side of a financial intermediary’s balance sheet. Competition, in particular, may increase instability by:

1. exacerbating depositors/investors’ coordination problem on the liability side, and fostering runs and/or panics, which may affect the system overall; and
2. 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 (1), 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 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 (see Smith 1984).

Imperfect competition is the norm and not the exception in banking. Indeed, asymmetric information creates barriers to entry or results in competition not delivering efficient outcomes.

Chang and Velasco (2001) present a model of financial crisis in emerging markets in the Diamond and Dybvig (1983) tradition. They find that financial liberalisation increases the expected welfare of depositors, but may also increase fragility. Liberalisation is modelled 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 maximises depositors’ ex ante utility, taking into account the probability of a run, associated with an exogenous sunspot, and autarchy is a feasible allocation.11

Let us now consider a stylised banking crisis model, based on Rochet and Vives (2004) and Vives (2010a). The model has three dates: \(t=0,1,2\). On date \(t=0\), the bank has equity \(E\) (or, more in general, stable funds including insured deposits) and collects uninsured certificates of deposit or short-term uninsured debt worth \(D_s=1\). 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 certificates of deposit are repaid at their face value \(D_s\) and the bank’s equity holders obtain the residual (if any). A continuum of fund managers makes investment decisions in the interbank market. At \(t=1\) each fund manager, after the observation of a (conditionally independent) private signal about the future realisation 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 \(\theta \sim \mathcal{N}(\bar{\theta}, \tau^2)\) and the private signal for investor \(i\) is \(s_i = \theta + \varepsilon_i\), with i.i.d. distributed noise \(\varepsilon \sim \mathcal{N}(0, r, \mu^2)\) orthogonal to \(\theta\).

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

Let \(m \equiv M/D\) be the liquidity ratio; \(\theta^* = (D-M)/I\), the solvency threshold of the bank; \(\lambda > 0\) the fire sales premium on early sales of bank assets; and \(\theta_0 = (1+\lambda)\theta^*\) the “supersolvency” threshold, such that a bank does not fail, even if no fund manager renews the certificates of deposit. Under these conditions the bank fails if \(\theta < \theta^*_L\) or when \(\theta \geq \theta^*_L\) but

\[
\hat{y} \geq m + \frac{1 - m}{\lambda} \left( \frac{\theta^*_L - 1}{\lambda} \right)
\]

When taking into account the balance sheet constraint at \(t=0\), \(E+D_s=I+M\) we have \(\theta^*_L = (1-m)/(I+\lambda-m)\), where \(I=D/E\) is the short-term leverage ratio and \(d=D/D_s\) 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.12 If the state of the world is known, then if \(\theta < \theta^*_L\), the dominant strategy will involve withdrawal; if \(\theta \geq \theta^*_L\), then the dominant strategy will be to remain (not to withdraw); and for \(\theta \sim (\theta^*_L, \theta^*_U)\) both equilibria coexist. We can show that with incomplete information, an equilibrium is characterised by two thresholds \((s^*, \theta^*)\) with \(s^*\) yielding the signal threshold below which an investor withdraws and \(\theta^* \sim [\theta^*_L, \theta^*_U]\) 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, \(m \epsilon (0,1)\), such that \(\theta^*=\theta^*_L\) for \(m\geq m\), and for \(m < m\) we have \(\theta^* > \theta^*_L\). In this case, the equilibrium is unique if \(r, \lambda \leq 2 \pi (1+\lambda-m)/\lambda^2\).13 The reason is as follows. Let \(R(s)\) be a player’s best reply threshold to the (common) signal threshold \(s\) used by other players. The game then involves strategic complementarities, with \(R^* > 0\). A higher

---

11 Todd Keister raised a similar point when discussing the paper.

12 Related examples can be found in Morris and Shin (1998, 2004) and Corsetti et al. (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).

13 All the results presented in this section are to be found in Vives (2010).
threshold \( \tilde{s} \) applied by others induces a player to also a higher threshold too. We can show that if \( \tau_s \tilde{s}^{-1/2} \leq \sqrt{2\pi (l^1 + d^1 - m)} \lambda^{-1} \) then \( R'(\tilde{s}) \leq 1 \). This ensures that \( R(\cdot) \) crosses the 45\(^{\circ} \) line only once and that the equilibrium is unique. In Figure 7, the uniqueness case is illustrated by the flatter best reply curve and the three equilibria case by the steeper best reply curves.

**Figure 7.** Best response of a player to threshold strategy \( \tilde{s} \), used by rivals.

![Figure 7](image_url)

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 \((\bar{\tau}_c + \tilde{\tau}_c)/(\bar{\tau}_c + (1 + d^1 - m)\lambda^{-1} \sqrt{2\pi} \tilde{\tau})\). 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 with small noise in the signals in relation to the prior \((\tilde{\tau}/\sqrt{\tilde{\tau}})\). With small noise in the signal, a player faces greater uncertainty about others’ behaviour, 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 \( Pr(\theta < \theta') \). A crisis occurs due to low values of fundamentals. In contrast, the complete information model contains multiple self-fulfilling equilibria in the range \((\theta^*, \bar{\theta}_c)\). Consequently, the model bridges between the self-fulfilling theory of crisis (e.g., Diamond and Dybvig 1983) and theory linking crisis to fundamentals (e.g., Gorton 1985).

In the range \([\theta^*, \bar{\theta}_c]\), coordination fails, from the perspective of the institution attacked. Thus, the bank is solvent but illiquid, that is, the bank would have no problem if only investors would renew their certificates of deposit, but it has fallen into a range where they do not, and the bank becomes illiquid. Thus, the risk of illiquidity is represented by \( Pr(\theta < \theta^*) \) and the risk of insolvency by \( Pr(\theta < \bar{\theta}_c) = \Phi(\sqrt{2\pi} \tilde{\theta} / \tilde{\theta}) \), where \( \tilde{\theta} \) 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 and an increase in \( d \) or \( \lambda \) boosts both \( \theta' \) and \( s' \), the probability of crisis \( Pr(\theta < \theta') \), and the range of fundamentals \((\theta^*, \theta')\) for which there is coordination failure (Vives 2010a).\(^{14}\) Any rise in bank vulnerability, whether it affects the face value of bank deposits or the fire-sales premium for early liquidation, increases fragility, by increasing the degree of strategic complementarity. Furthermore, if released, public signal \( \theta \) 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 \( R(s'; \tilde{\theta}) + s' = 0 \). From this, it follows that:

\[
\frac{ds'}{d\tilde{\theta}} = \frac{\partial R / \partial \tilde{\theta}}{1-R'},
\]

whenever \( R' < 1 \) is met, since \( R' > 0 \). As a result, a rise in \( \tilde{\theta} \) will affect the equilibrium threshold \( s' \) more than the direct impact on the best response of a player \( \partial R / \partial \tilde{\theta} \). 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 emphasised by Morris and Shin 2002). Every investor knows that an increase in \( \theta \) 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 (See Madies 2006 and Garratt and Keister 2009).

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 a bank in trouble to avoid contagion (Allen and Gale 2004, Sáez and Shi 2004) or may impede its provision (as banks with surplus funds under provide lending strategically, to induce fire-sales of the bank-specific assets belonging to needy intermediaries (Acharya et al. 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 \( \tilde{\theta} \) 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. The region of

---

\(^{14}\) 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.
potential multiplicity $\tau_\theta \theta \tau_\theta \theta \theta \leq \sqrt{2 \pi (l^1 + d^{-1} - m)\lambda}$ 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 $\tau_\theta / \sqrt{\tau}$.

Regulation in the form of solvency and/or liquidity requirements may help to control the probabilities of insolvency and illiquidity (Vives 2010a). Indeed, the probability of insolvency $Pr(\theta < \theta_1)$ is decreasing in $m = M/D$ (assuming that $1 - l^1 - d^{-1} < 0$ as is usual in a commercial bank), the solvency ratio $l^1 = E/D$, and $d^{-1}$, since $\theta_1 = (1 - m)/(l^1 + d^{-1} - m)$. The probability of a crisis $Pr(\theta < \theta^*)$, including the probability of illiquidity, is decreasing in $m$, $l^1$ (and also in $d^{-1}$ and $\lambda$) since $\theta^*$ is as well.

From this it follows that both solvency and liquidity requirements needed 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 (see Vives 2009). However, note that there is a partial substitutability between $m$ and $l^1$, since they both contribute to reducing $\theta_1$ and $\theta^*$. In the limit case of almost perfect signals, $\tau_\theta \rightarrow \infty$, 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 one can be relaxed.

In general the socially optimal probability of a crisis is positive, because of its disciplining effect. 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 (See Ratnovski and Huang (2009) with evidence from the 72 largest commercial banks in OECD countries).

In short, runs can happen independently of competition levels, but rising competitive pressure worsens investors’/depositors’ coordination problem, and increases:

1. Potential instability (enlarging the multiplicity of equilibria region).
2. The probability of a crisis.
3. The range of fundamentals for which there is coordination failure of investors (and the institution is solvent but illiquid).
4. The impact of bad news on fundamentals.
5. The solvency requirement.

It is worth pointing out that, 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 minimised.

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 flat deposit insurance. 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 certificates of deposit). A bank, then, cannot increase its market share and profits by taking more risk, because investors will discount it. Introducing flat premium deposit insurance (or bailouts) however, destroys the market’s disciplinary effect market by eliminating investor concerns about potential bank failure.

Intense competition may worsen the problem of excessive risk taking as high profits provide a buffer.

Intense competition may worsen the problem of excessive risk taking as 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 regard to reputational benefits, both of which may be eroded by more competition. 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

15 A better reputation reduces the cost of outside finance to the bank.
deposit market and choosing 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 (see also Hellman et al. 2000 and Cordella and Yeyati 2002).

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, 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, and Gehrig 1998).16

**All in all, an increase in the level of competition will tend to increase risk taking incentives and the probability of bank failure.**

On the other hand, 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 Nicoló 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 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 (Besanko and Thakor 1993, Petersen and Rajan 1994 and 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 (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.

### 4. Evidence

Increased competition after liberalisation and deregulation in the US in the 1980s led banks to take more risks (Keeley 1990, Edwards and Mishkin 1995, Demsetz et al. 1996, Galloway et al. 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 US bank holding companies for the period 1971-1986. Furthermore, he finds that interest rates on large certificates of deposit 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, too big to fail 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, and Dick 2007, respectively). Saurina et al. (2007) claim that non-performing loans in Spanish banks fell as the loan market’s Lerner index rose.

16 Note also that endogenous fixed costs due to information gathered via lending may induce a natural oligopoly in banking (Dell’Ariccia et al. 1999, Dell’Ariccia 2001).

17 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.
competition and eroded banks’ market power (measured again by Tobin’s q), banks with lower charter values tended to have lower equity–asset ratios (lower solvency), and to experience higher credit risk (loan losses over total loans).

Liberalisation in a weak institutional environment or with inadequate regulation – or a mixture of the two – shifts risk to the taxpayer and increases the likelihood of systemic crisis (Demirgüç-Kunt and Detragiache 1998 and 2001). A similar situation seems to have arisen in the wake of the subprime crisis, with declining lending standards associated with securitisation (Demirgüç-Kunt and Detragiache 1998 and 2001). Specifically, both concentration and competition are linked to aggregate asset concentration) and increasing competition on both individual and systemic bank failures. In a cross-country study of 23 developed nations, Berger et al. (2009) show that market power (as measured by the Lerner index or the HHI on deposits/loans at 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 (1980-1997), Beck et al. (2006) show that systemic crises are less likely in concentrated banking systems (measured by the three-firm concentration ratio on total assets) 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 stabilising influence. The connection that really matters, however, is that between concentration in relevant markets (which need not be directly linked to aggregate asset concentration) and competition. Furthermore, concentration is, in fact, endogenous and more competition may increase concentration in a free entry world (as there is less room for entrants) (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 et al. (2006) seems to be: “More competitive banking systems are associated with less fragility, when controlling for concentration”. Schaek et al. (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 Nicoló, 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 et al. (2006) as an indication that more concentration leads to less intervention and more systemic crises, and that fewer entry barriers lead to more intervention and fewer crises. In a cross-country study with individual bank data, Boyd et al. (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 liberalisation 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 US find strong benefits of consolidation (improving profitability and production efficiency, and reducing insolvency risk) when the degree of macroeconomic (geographic) diversification increases (Hughes et al. 1996 and 1998 and Demsetz and Strahan 1997). Specifically, these authors find that geographic diversification offsets the tendency of larger banks to risk insolvency more (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 US, by looking at the positive trend of stock return correlations for large and complex banking organisations in the period 1988-1999 (De Nicoló and Kwast 2002).

There is ample evidence that institutions close to insolvency have incentives to gamble for resurrection

Internationalisation is another 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 and Barth et al. 2004). 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 et al. (2008) find that foreign bank entry in poor countries may reduce private credit growth. Berger et al. (2001) find that large foreign-owned institutions concentrate on large-scale projects and may leave out small firms. Still, large, well-capitalised 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 Argentine's currency board).

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 (see Vives 2006).

Finally, there is ample evidence that institutions close to insolvency have incentives to gamble for resurrection (e.g. S&Ls 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 less concentrated ones, such as those in the US or Germany, countries with concentrated systems, such as the Netherlands or the UK (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 to be a source of increased risk and vulnerability (see Baele et al. 2007, De Jonghe 2010, Demigüç-Kunt and Huizinga 2010, and Ratnosvski and Huang 2009). By the same token, it is not evident that certain types of institutions have been more vulnerable than others. Both specialised investment banks (in fact, all the US ones have collapsed or converted to commercial banks), insurance companies like AIG, and universal banks (UBS, Citigroup, or German and UK banks) have suffered.

In conclusion, the evidence points to the following:

1. Liberalisation increases the occurrence of banking crises, while a strong institutional environment with adequate regulation mitigates them.
2. There is a positive association between some measures of bank competition (e.g. low entry barriers, openness to foreign entry) and stability.
3. The association between concentration and stability presents mixed results.
4. 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 may have led banks to take on 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 asset risk position of the bank is not observable then incentives to assume more risk increase considerably (second row of Table 1). These incentives become 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 internalise (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 (See Matutes and Vives 2000, Hellmann et al. 2000, and Repullo 2004).

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 behaviour. 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.

Table 1. Possible banking regimes, the incentives to take risk on the liability and asset sides, and the necessary regulatory instruments, when charter values are low and the social cost of failure is high.

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

5.1 Regulatory failure

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 risk assessment have been deficient (among other things because of the failure of rating agencies), and market discipline has been ineffective because of the blanket insurance offered by too big to fail policies. Furthermore, capital regulation has not taken into account systemic effects (the social cost of failure) and assets restrictions have been lifted, under the pressure of investment bank lobbies.19

5.2 Status quo

At present we are stuck in the “risk-insensitive insurance” row with maximal risk taking incentives. We therefore need to design appropriate capital requirements and asset restrictions. Optimal regulation would require 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 internalise 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-medium friction environment.

5.3 A complex trade-off

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 behavioural parameter but one that enters the utility function. 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 8 depicts the regions in the space of intensity of competition $\lambda_*$ (with $\lambda=0$ for an independent monopolies situation 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 ($r^*>r_*$) or too low ($r^*<r_*$), 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...

18 According to Basel II’s guidelines on capital requirements, banks can choose between a “standardised” 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.

19 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 US.

20 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.
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 8. Comparison of market and optimal deposit rates, as a function of the friction in the market ($\lambda$) and social cost of failure

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 characterised by high uncertainty, increased likelihood and incidence of financial and currency crises, predominant financial role of 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 (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. 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.21

6. The policy response to a financial crisis22

6.1 Interventions and distortions

When a systemic crisis hits, the pressure to stabilise the system is tremendous. In the 2007-2008 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 nationalisation, 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, too big to fail 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 EU, posing a threat to the single market.

---

21 In fact, according to the US Federal Deposit Insurance Corporation Improvement Act (FDICIA, 1991), when solvency falls below a certain limit, the bank cannot expand its assets. A further decline in solvency may trigger the need to recapitalise 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).

22 This section is based on Vives (2010b).
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 S&Ls in the US, Japan’s banking crisis, and evidence on 21 emerging countries (Brown and Dinç 2009). 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. The upshot is that surviving incumbents enjoy more market power and lower capital costs, because they are too big to fail (and/or because of the public help). Recall 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 behaviour. 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.

The trade-off between competition and stability is complex, but seems real. Well-designed regulation may alleviate this trade-off, but needs to consider it.

State intervention and even outright ownership have been necessary to stabilise 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; political objectives and incentives rule (see Hau and Thum 2009); and, if not disciplined by competition, it makes the banking system less efficient and encourages inefficiency leading to less financial stability with higher risk exposure and more bank losses (Barth et al. 2004, Caprio and Martinez Peria 2002, De Nicoló and Loukoianova 2007). 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, policymakers 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).

When the taxpayer is footing the bill, the public sector must have a say in how the institution receiving help is run

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 interbank exposures. This does not distort competition so long as 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 however, 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 behaviour. This is not easy, particularly when compounded by regulatory failures, which induce excessive risk taking.

The main tools of intervention to limit distortions are structural (asset divestitures) and behavioural (pricing, advertising, acquisitions) restrictions. Structural commitments may help reduce the post-crisis over-capacity 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 (such as the US and the UK). This means that branches and personnel must be cut, together with the balance sheet, even if credit is normalised (because it should stabilise below the pre-crisis bubble levels). In general, care must be
taken to ensure that any commitments, whether structural or behavioural, 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, which 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 behavioural restrictions on the helped bank could be relaxed.

6.2 Approaches in the US and in the EU

The role of the competition authority in the US has been different from that of the EU, because the EU competition authority has the unique capability, among competition authorities, to control state aid. Since the crisis, the EU has dealt with many banking aid cases (taking 22 decisions in 2008 alone, 81 decisions as of 17 December 2009). Most of the cases (75) were approved without objection. The EU has explicit conditions for state guarantees (EU Communications October-December 2008), which have been formalised as temporary guidelines on restructuring aid to banks. The conditions imposed on helped institutions are mostly sensible, since they try to minimise the distortions introduced by public help, in particular for fundamentally unsound institutions. The European Commission has toughened or influenced significant balance sheet reductions and behavioural restrictions on helped entities such as ING, Northern Rock, RBS, and Commerzbank. 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 (see the European Commission’s Directorate General for Competition’s State Aid report of 14 December 2009 in which the Commission approves impaired asset relief measure and restructuring plan of RBS).

Some measures can be understood as efforts to minimise competitive distortions of the aid, 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 too big to fail issue. More broadly, the concept of competitive distortion may address the issue of competition for those enjoying the advantages of a too big to fail 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 analysis. The RBS case points to the need for coordination between the competition authority and the regulator.

Size and scope restrictions are a blunt instrument to deal with the too big to fail issue. The activism of the European Commission poses the question of (future) competitive balance with US banks receiving assistance that required 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 US 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 EU is that it limits bankers’ incentives to take excessive risk, under the expectation of a bailout if things go wrong. Thus, it addresses the too big to fail issue. The competition authority may internalise the fact that if an institution fails it 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 EU 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 US seems to be following another route, where too big to fail is explicitly not an antitrust problem (see White 2009).

In any case, size and scope restrictions are a blunt instrument to deal with the too big to fail issue. Controls on size are problematic, because interconnectedness and business specialisation 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 too big to fail issue and therefore its actions should be coordinated with the regulator. The potential for competition policy to provide a commitment device to partially address too big to fail issues should not be dismissed.

Size and scope restrictions are a blunt instrument to deal with the too big to fail issue.

The Obama administration’s move is reminiscent of the 19th 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

Liberalisation 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. The resulting restructuring has tended to increase aggregate concentration. With the onset of the crisis however, banking appears to be making a return to traditional banking – a move that may exacerbate the consolidation trend.

Banks are unique. Their particular mix of features makes them vulnerable to runs with potentially systemic consequences. While the fragility of a competitive banking system is typically excessive, the financial regulation that comes to the rescue brings with it side effects and regulatory failure. The most important of which 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. One only needs to think of the failure of rating agencies to be convinced that disclosure and risk assessment have been deficient, and market discipline has been ineffective because of the blanket insurance offered by too big to fail policies. Capital regulation has not taken into account the social cost of failure and, under pressure from investment bank lobbies, asset restrictions have been lifted. Supervision has proved ineffective – 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 crisis, and the impact of bad news on fundamentals. This does not imply that competitive pressure has to be minimised, 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 liberalisation 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.

While regulation can alleviate the competition-stability trade-off, the design of optimal regulation must 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 extremely difficult in practice, the trade-off between competition and stability is bound to persist. This suggests that coordinating regulation and competition policy is necessary. Banks’ uniqueness, not only during crises, should be recognised and the appropriate lessons drawn and applied during competition policy implementation.

Emerging economies should carefully balance the impact of the different levels of friction and the social cost of failure.

The competition-stability trade-off also applies to emerging economies. An emerging market economy is characterised 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 too big to fail institutions remains an open issue. In the US, too big to fail is not an antitrust issue, whereas in the EU the competition authority controls distortions of competition which arise out of state aid, and this has implications for too big to fail. 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 specialisation 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.

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 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 too big to fail problems that lead to competition distortions. The regulator can address the too big to fail 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 recapitalisation 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 remains open over whether to let firms – in particular banks – get so large that they significantly influence regulation.

References


Ratnovski, L. and R. Huang (2009), “Why are Canadian Banks More Resilient?” IMF working...


Xavier Vives is professor of Economics and Finance and academic director of the Public-Private Sector Research Center at IESE Business School. He holds a Ph.D. in Economics from the University of California at Berkeley. He is Research Fellow of CEPR where he was Program Director for Industrial Organization (1994-1997) and for Applied Microeconomics (1991-1993). Xavier Vives is member of the Economic Advisory Group on Competition Policy of the European Commission; Research Fellow and member of the European Economic Advisory Group of CESifo (Munich); Fellow of the European Economic Association, member of its Executive Board and editor of the Journal of the European Economic Association; Fellow of the Econometric Society and member of its Council; and President of the Spanish Economic Association for 2008. His fields of specialization are industrial organization, the economics of information, banking and financial economics. He has taught at the U. of Pennsylvania, Harvard, NYU, INSEAD, UAB and UPF; published over a hundred articles in scientific outlets, and authored several books. He has received several research awards, such as the King Juan Carlos I Prize in 1988;