The Role of Finance in Economic Development:

Benefits, Risks, and Politics

Thorsten Beck*

Abstract:
Theoretical and empirical research has shown that a sound and effective financial system is critical for economic development and growth. The financial system, however, is also subject to boom and bust cycles and fragility, with negative repercussions for the real economy. Further, the political structure of societies, often pre-determined by historic experience, is critical for the structure and development of the financial system. This paper is a critical survey of three related strands of literature – the finance and growth literature, the literature on financial fragility, and the politics and finance literature.

Keywords: Financial Development; Financial Fragility; Finance and Politics; Economic Development; Access to Finance; Banking Crisis

JEL codes: G0, G1, G2, O1, O4,
1. Introduction

What role does finance have in the development of market-based economies? Early on, politicians and economists alike emphasized the importance of the financial system for the rise of capitalism, industrialization, and economic development. Smith (1776) pointed to the role of money in lowering transaction costs, thus permitting greater specialization, and fostering technological innovation. Alexander Hamilton (1781), one of the founding fathers of the United States, argued that “banks were the happiest engines that ever were invented” for spurring economic growth. Joseph Schumpeter argued in 1911 that financial intermediaries play a pivotal role in economic development because they choose which firms get to use society’s savings.¹ On the other hand, development economists for many decades have ignored the financial system and focused on other policy areas. Lucas (1988) described the role of finance in the growth process as overstated, and Robinson (1952) argued that financial development primarily follows economic growth. Following the seminal works by Goldsmith (1969), McKinnon (1973), and Shaw (1973), however, a large and still active theoretical and empirical literature has related financial development to the economic growth process. Empirical studies have found a positive impact of financial deepening on economic growth, statistically and economically significant. More recently, studies have related the development of the financial sector to other real sector outcomes, including the pattern of countries’ trade balance and changes in income distribution and poverty levels.

The same mechanisms that underpin the positive role of finance, however, are also a source of risk and fragility. The history of finance is full of boom-and-bust cycles, bank failures, and systemic bank and currency crises (Reinhart and Rogoff, 2009). Just as there is a comprehensive literature on the impact of finance on growth, there is an equally important

¹ See Schumpeter (1934). The author is grateful to Ross Levine for sharing many of the historic references.
literature that has explored the causes and socioeconomic costs of financial fragility, including systemic banking crises. Historic analyses and case studies have given way to more systemic cross-country explorations of idiosyncratic and systemic banking distress and their determinants.

Given the importance of finance for growth, its inherent risks, and the large socioeconomic costs of banking crises, it is not surprising that the financial sector is often at the top of the policy agenda. However, the importance of access to credit as entry barrier into the real sector and the relative ease with which owners and creditors of financial institutions can be expropriated also makes financial sector policies an important tool in the political process. Subsidized credit programs and credit guarantees are often an easy and cheap tool of fiscal policy as they create contingent rather than real liabilities. The dependence of most real sector enterprises on access to external finance makes the financial sector critical in the attempt of ruling elites to entrench their socioeconomic dominance and prevent entry of competitors. The reliance of financial institutions and markets on contractual institutions makes them dependent on the political sphere.

The recent crisis has brought these issues to the forefront of the academic and also political debate. The crisis has also shed doubt on the previous findings of a positive impact of finance on growth. How much finance is good for growth? Are financial crises too high a price for having a thriving financial system? Are credit boom-and-bust cycles behind economic cycles? What is the politics behind financial development and fragility?

This paper explores the role of the financial sector for economic growth, the causes and consequences of financial fragility, and the politics behind financial deepening and fragility. In doing so, I identify the critical role of the financial sector within capitalist economies, a role with bright and dark sides. Specifically, I survey the large theoretical and
empirical literature that links a sound financial system to the process of economic development. I discuss the theoretical and empirical literature on bank fragility and banking crises and survey the literature on the political economy of financial deepening. Importantly, I relate these three strands of the literature to each other and to the current crisis.

Financial institutions and markets depend critically on contractual institutions, and this survey is thus closely related to the institutions and development literature (Acemgolu, Johnson, and Robinson, 2005). Specifically, given the intertemporal nature of financial contracts, the financial system is one of the most institution-sensitive sectors of the economy. The financial sector depends as much as contractual institutions on property rights protection and thus the political structures of societies.

This paper is related to other recent surveys. Levine (2005) surveys the theoretical and empirical literature on finance and growth, and Beck (2009) surveys the econometric methodologies behind the empirical finance and growth literature. Demirgüç-Kunt and Detragiache (2005) discuss the literature on banking crises, and Haber and Perotti (2008) offer a critical survey of the finance and politics literature. This survey is also related to recent surveys on the political economy of the financial and legal system (Beck and Levine, 2005). Compared to these previous surveys, this paper tries to bring these three literatures together and relate them to the first global financial crisis of the twenty-first century.

The remainder of this paper is organized as follows. The next section surveys the theoretical and empirical finance and growth literature. Then I discuss the theoretical and empirical literature on financial fragility. Following that is a survey of the political economy literature of finance, and a final section brings these three literatures together and concludes.
2. Finance and Economic Development

What role does the financial sector play in market economies? How important is the financial sector in the growth process of countries? Over the past thirty years, a flourishing theoretical literature has explained the endogenous emergence of financial institutions and markets and has explored their impact on real sector outcomes, including economic growth and income inequality. Over the past twenty years, a still growing empirical literature has explored the effect of financial systems and their different segments on economic growth and other real sector variables. This literature has also explored the relative importance of banks and markets and the impact of financial sector development on other real sector outcomes. More recently, this literature has explored the distributional repercussions of financial deepening and the effect of broadening. I discuss each in turn.

2.1. Finance and Growth: Theory

The theoretical literature on financial intermediation has focused on two important dimensions. Why do financial markets and institutions exist? And what is their impact on savings, investment, and economic growth? I discuss each question in turn.

At the core of the existence of financial institutions and markets are market frictions, which financial institutions and markets can help alleviate, such as asymmetric information between contractual partners resulting in agency problems and risks of illiquidity and default. Building on the insights by Stiglitz and Weiss (1983) on the importance of agency problems, several articles have shown how financial institutions and markets can economize on screening and monitoring costs of many individual lenders and, by diversifying risk across many different projects, improve on a world without them.2 By pooling savings across a large

---

number of savers with differently timed liquidity needs, financial institutions can help overcome liquidity risks and ultimately provide savers with a higher return. Similarly, more liquid financial markets increase incentives for investors to relinquish control over their savings, as they are able to access them through financial markets on an immediate basis, while at the same time earning higher returns. The emergence of financial institutions and markets can thus be explained by the gains for economic agents, a theoretical argument that is consistent with the historical observance that financial institutions and markets have arisen at an early stage of human history and especially as exchange of goods and services across larger geographical distances and within larger societies or between societies has become more prominent.3

The endogenous emergence of financial institutions and markets does not in itself imply a positive impact on economic growth. A large theoretical literature, however, has explored several channels through which financial systems can help increase economic growth rates, both through improved capital accumulation as through higher productivity growth. On a broader level, these theories have shown how financial markets can help overcome the market frictions of indivisible projects and inability to diversify risks that have held back development in many developing economies (Acemoglu and Zilibotti, 1997). I discuss these different channels in turn.

First, and on a very basic level, financial systems can support the efficient exchange of goods and services by providing payment services and thus reducing transaction costs. Financial services can foster specialization by enabling more transactions, thus fostering productivity growth.4

---

3 For a discussion of the role of the shareholding company in Rome, see Malmendier (2009). For a discussion on the role of different financial contracts to finance trade across the Mediterranean, see, for example, Williamson (2010) and the literature quoted therein.

4 See, for example, Greenwood and Smith (1997).
Second, by pooling savings from many individual savers, financial institutions and markets can help overcome investment indivisibilities and allow exploiting scale economies. This does not necessarily have to be national financial institutions but can be local coalitions of investors, as was the case in the early days of the Industrial Revolution for infrastructure projects.

Third, by economizing on screening and monitoring costs and thus allowing more investment projects to be financed and, ex ante, increasing the aggregate success probability, financial institutions and markets can ultimately have a positive impact on investment and resource allocation. Similarly, by identifying the entrepreneurs with the most promising technologies, financial intermediaries can also boost the rate of technological innovation and ultimately growth. A similar argument holds for financial markets: in larger and more liquid markets, agents have greater incentives to invest in research on enterprises and projects, which produces information that can be turned into trading gains, ultimately improving resource allocation.

Fourth, both financial institutions and markets can help monitor enterprises and reduce agency problems within firms between management and majority and minority shareholders, again improving resource allocation. Debt instruments can reduce the amount of free cash available to firms and thus managerial slack (Aghion, Dewatripont, and Rey, 1999), while liquid stock exchanges can allow investors to monitor and discipline enterprises through the threat of takeovers and subsequent dismissal of management. Linking stock performance to manager compensation can help align the interests of managers with those of

---

5 See, for example, McKinnon (1973), Sirri and Tufano (1995), and Acemoglu and Zilibotti (1997).
6 See, for example, Trew (2010).
7 See, for example, King and Levine (1993b) and Blackburn and Huang (1998).
8 See Grossmann and Stiglitz (1980), among others.
owners,\textsuperscript{10} although it is important to define a proper benchmark. Similarly, as in the case of screening, financial institutions can economize on the costs of monitoring by functioning as “delegated monitor.”\textsuperscript{11} By building long-term relationships, financial institutions can further reduce monitoring costs. Both financial markets and institutions can thus improve resource allocation and productivity growth. By reducing control problems of investors vis-à-vis owners and managers of enterprises, improved corporate governance can also increase savings and capital accumulation.

Fifth, banks can also help reduce liquidity risk and thus enable long-term investment, as shown by Diamond and Dybvig (1983). By pooling savings of patient and impatient agents, financial institutions can transform short-term liabilities into long-term assets, enabling long-term investment and ultimately economic growth. Similarly, liquid markets can enable investment in long-term investment projects while at the same time allowing investors to have access to their savings at short-term notice (Levine, 1991). Financial institutions can also ease liquidity needs of enterprises, enabling long-term investment and R&D activities.\textsuperscript{12}

Sixth, financial institutions and markets allow cross-sectional diversification across projects, allowing risky innovative activity while guaranteeing an ex ante contracted interest rate to savers (King and Levine, 1993b). Furthermore, aggregate risk that cannot be diversified away at a specific point in time can be diversified by long-living financial intermediaries over time (Allen and Gale, 1997).

Beyond theoretical models, economists have explained the take-off of the Industrial Revolution in some countries earlier than others to the availability of finance. Hicks (1969) argued that the Industrial Revolution in the United Kingdom was possible due to the developing British financial system. Although many inventions were made before the

\textsuperscript{10} Diamond and Verrecchia (1982) and Jensen and Murphy (1990).
\textsuperscript{11} Diamond (1984) and Bencivenga and Smith (1993).
\textsuperscript{12} See Holmström and Tirole (1998) and Aghion et al. (2010).
Industrial Revolution, liquid capital markets enabled investment into long-term projects that could use these inventions. Similarly, the Netherlands and the United States experienced financial deepening before their economic and political rise in the seventeenth and twentieth centuries, respectively.

The relationship between finance and growth is not a one-way street; rather, higher growth induced by financial deepening increases demand for financial services, ultimately resulting in a virtuous circle of economic and financial development (Greenwood and Jovanovic, 1990). This theoretical insight has also resulted in additional challenges for the empirical finance and growth literature, as a positive relationship between financial and economic development might reflect causation from economic to financial development.

Theory, however, is not unambiguous in predicting a positive impact of financial deepening on economic growth. Better resource allocation may depress saving rates enough such that overall growth rates actually drop with enhanced financial development.\(^\text{13}\) This can happen if the income effect of higher interest rates is larger than the substitution effect. The financial sector might also attract too many resources relative to the real sector, with negative repercussions for growth.\(^\text{14}\) Critically, the impact of finance on growth might vary across different levels of income per capita, with the positive relationship being strongest among low- and middle-income countries that are catching up to high-income countries in their productivity levels and fading away as countries approach the global productivity frontier (Aghion, Howitt, and Mayer-Foulkes, 2005). A somewhat separate but related discussion is that about the importance of financial sector development compared to other sectors and

---

\(^\text{13}\) See, for example, Bencivenga and Smith (1991) and King and Levine (1993b).

\(^\text{14}\) Philippon (2010) models the trade-off between the financial sector helping overcome agency problems, while at the same time competing for human resources with the real sector. In a situation where the social value of entrepreneurship is larger than the private value, the financial sector can be too large compared to the entrepreneurial sector. Similarly, Bolton, Santos and Scheinkman (2011) model how individuals can choose to work in the real sector or as dealers in the financial sector. While dealers can provide entrepreneurs incentives to originate good assets, they might extract excessively high informational rents and thus attract too much young talent towards the financial industry, thus leading to lower GDP per capita growth.
policies in explaining growth. This ambiguity has motivated a large empirical literature over the past twenty years, to which I turn now.

2.2. Finance and Growth: From Correlation to Causality

The empirical literature on finance and growth has made enormous progress over the past two decades. To the same extent that the challenge of causality has been addressed, researchers have moved from aggregate macro-level data to more micro-level data, in the process also disentangling the mechanisms and channels through which financial development is associated with economic growth.

Goldsmith (1969) was the first to show empirically the positive correlation between financial development and GDP per capita, using data on the assets of financial intermediaries relative to GNP and data on the sum of net issues of bonds and securities plus changes in loans relative to GNP for thirty-five countries over the period 1860 to 1963. Such a correlation, however, does not control for other factors associated with economic growth and might thus be driven by other country characteristics correlated with both finance and growth. Second, such a correlation does not provide any information on the direction of causality between finance and growth. The early finance and growth literature has therefore used standard cross-country ordinary least squares (OLS) regressions to control for other country characteristics associated with growth differences across countries (King and Levine, 1993a, 1993b). This literature has shown that both banking sector as well as equity market development are robust predictors of GDP per capita growth (Levine and Zervos, 1998).

In a second step, researchers have addressed the issue of reverse causation and omitted variables using instrumental variable techniques. Using external instruments, such as historic country characteristics that can explain cross-country variation in financial sector
development, or internal instruments such as lagged values of financial sector indicators, several papers have shown that the relationship between financial sector development and economic growth is not due to reverse causation or omitted variable bias. Specifically, Levine, Loayza, and Beck (2000) and Beck, Levine, and Loayza (2000) show that instrumenting financial development with the legal traditions of countries and applying dynamic panel techniques with lagged values as instruments confirm the positive relationship between finance and growth.15

An alternative approach has been to explore the relationship between financial development and GDP per capita for a specific country over time. Compared to cross-country research, the time-series approach relies on higher frequency data (mostly yearly) to gain econometric power. The causality approach of the time-series approach, however, is different; specifically, the Granger causality tests are tests of forecast capacity—that is, to what extent does one series contain information about the other series? Unlike the cross-country panel regressions discussed earlier, this concept therefore does not control for omitted variable bias by directly including other variables or by controlling with instrumental variables. Rather, by including a rich lag structure, which is lacking in the cross-sectional approach, the time-series approach hopes to capture omitted variables. Numerous papers have found evidence for Granger causality from finance to economic development, though the evidence has been not unambiguous.16 There is also evidence for bidirectional causality, consistent with theory.17

An alternative to the instrumental variable and time-series approaches is to explore the mechanisms or channels through which financial development affects economic growth,

15 A related literature has explored the relationship between financial liberalization and economic growth. See, for example, Bekaert, Harvey and Lundblad (2005) and Henry (2003).
16 See, for example, Rousseau and Sylla (2005) for the United States, Bell and Rousseau (2001) for India, and Neusser and Kugler (1998) and Xu (2000) for cross-country samples.
17 See, for example, Demetriades and Hussein (1996).
which can also be seen as “smoking gun” approach. This implies testing for a differential impact of financial development on different sectors or industries. In a seminal paper, Rajan and Zingales (1998) show that industries that depend more on external financing grow faster in countries with higher levels of financial development. It is important to note that this is a relative effect, because it is gauged by differences-in-differences—the difference between a high-dependence and low-dependence industry in a well-developed financial system compared to a less developed financial system. Critical for their methodology is that their measure of external dependence captures purely demand-side effects; the authors claim to achieve this by focusing on a sample of large listed U.S. enterprises that should face a perfectly price-elastic supply curve. Following Rajan and Zingales, this differences-in-differences technique has been used widely in the literature, showing that financial development is conducive to the growth of industries with larger growth opportunities, more dependent on intangible assets, and with a larger share of small enterprises.18

An alternative differences-in-differences approach is similar to an event study that focuses on a financial sector policy change. Most prominent in this context is the branch deregulation episode in the United States in the 1970s and ’80s when states liberalized intra-and interstate branching. Using this almost identical policy reform implemented at different points in time, Jayaratne and Strahan (1996) show that deregulation led to lower loan losses and higher economic growth. Subsequent work has shown that this deregulation was also associated with an increase in entrepreneurship19 and lower economic volatility.20

The finance and growth literature has also explored the channels through which financial deepening fosters economic growth. There is robust evidence that the impact is

18 For an application of the differences-in-differences estimation, see, for example, Beck and Levine (2002), Beck (2003), Beck et al. (2008), Braun and Larrain (2005), Fisman and Love (2003), and Raddatz (2006).
20 Morgan, Rime, and Strahan (2004), Acharya, Imbs, and Stargess (2011), and Demyanyk, Ostergaard, and Sorensen (2007). Similar work has been undertaken across different regions of Italy; see Guiso, Sapienza, and Zingales (2004).
more through improved resource allocation, accumulation of knowledge, and productivity growth rather than through capital accumulation.21 There is also evidence that financial deepening affects the corporate structure of the private sector; firms are more likely to incorporate in countries with better developed financial and legal systems (Demirgüç-Kunt, Love, and Maksimovic, 2006).

One of the critical functions of the financial system, as already described, is maturity transformation. By enabling long-term investment projects, finance can help foster economic growth. Through this channel, financial systems can also help reduce volatility. Financial systems can alleviate firms’ liquidity constraints and facilitate long-term investment, which ultimately reduces the volatility of both investment and growth.22 Similarly, well-developed financial markets and institutions can help dampen the negative impact that exchange rate volatility has on firms’ liquidity and thus investment capacity.23

Recent publications have tested for the cross-country heterogeneity of the finance and growth relationship. There is evidence that the effect of financial development is strongest among middle-income countries, whereas other work finds a declining effect of finance and growth as countries grow richer, explaining this effect with finance helping countries catch up to the productivity frontier but not having any effect beyond this.24 More recently, Arcand, Berkes, and Panizza (2011) find that the finance and growth relationship turns negative for high-income countries, identifying a value of 110 percent private credit to GDP as approximate turning point, with the negative relationship between finance and growth turning significant at around 150 percent private credit to GDP, levels reached by some high-income countries in the 2000s.

22 See Aghion et al. (2010).
23 See Aghion et al. (2009).
It is important to note that there have also been a number of empirical studies criticizing the finance and growth relationship. A series of articles have shed doubt on the robustness of the finance and growth relationship. Other authors have focused on a direct relationship between the contractual framework that underpins financial sector development and economic growth. Using the Rajan and Zingales (1998) methodologies of matching industry with country characteristics, Claessens and Laeven (2003) find that industries more reliant on intangible assets, such as patents and trademarks, grow faster in countries with stronger contractual institutions. Evidence for a sample of four East European transition economies shows that trust in property rights, rather than access to credit, encourages entrepreneurs to reinvest their profits (Johnson, McMillan, and Woodruff, 2002), whereas evidence for China shows that both the quality of contractual institutions and access to finance explains profit reinvestment (Cull and Xu, 2005). Acemoglu and Johnson (2005) undertake a horse race on the aggregate level between contractual institutions that underpin financial development and protection of private property rights from expropriation by government and find that the latter (rather than the former) matters for long-term economic development. There is evidence, however, that this finding is due to the selection of the proxies for both contractual institutions and property rights protection and can be reinterpreted as the relative importance of the overall institutional framework (including informal institutions and norms) vis-à-vis formal institutions (such as courts) (Woodruff, 2006).

Overall, the overwhelming empirical evidence so far points to a positive relationship between financial deepening and economic growth beyond a pure correlation, a relationship that might vary over time and country conditions, however. In addition, there are important

---

nonlinearities. This evidence also gives a first hint at the fragility that can arise from rapidly expanding financial systems, a topic I return to later.

2.3. Banks versus Markets: Does Financial Structure Matter?

Until now, I have treated the financial system as a homogeneous sector. However, financial institutions, most prominently banks, and financial markets overcome the agency problem in different ways. Financial institutions create private information, which helps them reduce information asymmetries. Financial markets, on the other hand, create public information, aggregated into prices. Similarly, there are differences in the mechanisms through which financial institutions and markets exercise corporate governance. Banks can help improve corporate governance directly through loan covenants and direct influence on firm policy and indirectly through reducing the amount of free cash flows senior management has available. Financial markets can help improve corporate governance by linking payment of senior management to performance, through voting structures and the threat of takeover if the stock price falls below a value that is seen below fair value. Finally, there are different ways financial institutions and markets help diversify risks. Banks offer better intertemporal risk diversification tools, whereas markets are better in diversifying risk cross-sectionally. Markets are better in offering standardized products, and banks are better in offering customized solutions. However, banks and markets can also be complementary through instruments such as securitization, allowing exit strategies for venture capitalists, and by providing competition to each other.26

However, there are also important arguments of why banks are better than markets and vice versa. In liquid markets, investors can inexpensively and quickly sell their shares

26 See Stulz (2001) for an overview.
and consequently have fewer incentives to expend resources monitoring managers. Bank-based systems mitigate this problem because banks reveal less information in public markets. Also, efficient markets can reduce the effectiveness of takeovers as a disciplining tool. Atomistic shareholders have incentives to capture the benefits from a takeover by holding their shares instead of selling them, thus making takeover attempts less profitable (Grossman and Hart, 1980). On the other hand, proponents of the market-based view emphasize that powerful banks frequently stymie innovation by extracting informational rents and protecting established firms (Hellwig, 1991). The banks’ market power then reduces firms’ incentives to undertake profitable projects because banks extract a large share of the profits (Rajan, 1992). Also, banks—as debt issuers—have an inherent bias toward conservative investments, so that bank-based systems might stymie innovation and growth.

Cross-country comparisons have not provided evidence for either view. Evidence on the aggregate cross-country level, on the cross-country cross-industry level, and on the cross-country firm level have not found any evidence that countries, industries, or firms grow faster in countries with either more bank-based or more market-based financial systems. Rather, the overall level of financial development, not structure, explains cross-country variation in economic growth. This is consistent with the financial services view, which focuses on the delivery of financial services and less on who delivers them. However, it is also consistent with the view that the optimal financial structure changes as financial systems develop, consistent with theoretical models to this effect (Boyd and Smith, 1998). It is also consistent with findings on different income elasticities of different segments of the financial system.

The development of contractual savings institutions and capital markets is much more

---

28 See Boot, Greenbaum, and Thakor (1993).
30 See Levine (2002), Beck and Levine (2002), and Demirgüç-Kunt and Maksimovic (2002), respectively.
income-elastic than the development of the banking system (Beck et al., 2008). This finding is consistent with the observation that most low-income countries have more bank-based financial systems. As more detailed data on different segments of the financial system and on the users of financial services, including firms and households, become available, more research can be undertaken in this area.

2.4. International Dimension: Finance and Trade Patterns

The efficiency with which the financial system intermediates society’s savings has a significant effect on resource allocation and thus economic structure. It is therefore not surprising that financial sector development also has a significant effect on the structure of a country’s trade balance. Because financial development can steer resource allocation toward specific sectors and industries, it can also turn into a comparative advantage in certain sectors and industries.

Theoretical models have shown that financial development turns into a comparative advantage for countries in sectors and industries with higher needs for external finance.31 Ju and Wei (2005) show that if the external finance constraint is binding in the economy, then further financial deepening will increase the output of the industry more dependent on external finance. On a more micro-level, Manova (2010) shows that productivity cutoffs for enterprises to become exporters vary across sectors with different needs for external finance and decreases with financial development. In addition to traditional endowments, such as land, labor, and human and physical capital, the degree to which financial systems can ease financing constraints of enterprises can thus also turn into a comparative advantage.

There is quite a bit of empirical support for these theoretical models. Countries with higher levels of financial development have higher shares of manufacturing exports and

31 See Kletzer and Bardhan (1987), Beck (2002), and Matsuyama (2005).
higher export shares in industries with higher financing needs (Beck, 2002, 2003). Countries with higher levels of financial development have higher export shares and trade balance in industries with more intangible assets (Hur, Raj, and Riyanto, 2006). Equity market liberalization increases exports disproportionately more in financially vulnerable sectors that require more outside finance or employ fewer collateralizable assets (Manova, 2008). In addition, total exports in financially more developed countries are more sensitive to exchange rate movements than in countries at lower levels of financial development (Becker and Greenberg, 2007). These cross-country findings are further confirmed by more disaggregate data. Using state-level data for the United States, Michalski and Örs (2011) find that interstate branch deregulation led to a significant increase in exports relative to domestic shipments. On the firm level, several papers have found a significant relationship between credit constraints and the decision to become an exporter.32

2.5. Access to Financial Services

Until now, I have discussed the relationship between financial development and aggregate economic welfare or aggregate real sector outcomes. However, financial development can have distributional effects because it benefits different groups of households or firms to a different extent. Transaction costs and risk profiles vary across the firm and household population and can be binding constraints for certain groups, especially small enterprises and the poor, when trying to access financial services. Small firms consistently report higher financing obstacles than medium and large enterprises, and they are also more adversely affected in their operation and growth by these obstacles.33 This can have an impact on firm size distribution across economies. For example, survey analysis has shown that smaller firms

32 See, for example, Berman and Héricourt (2008) and Muûls (2008).
33 See Beck et al. (2006) and Beck, Demirgüç-Kunt, and Maksimovic (2005), respectively.
grow relatively faster in Germany than in Côte d’Ivoire, whereas the opposite holds for large firms (Sleuwaegen and Goedhuys, 2002).

Financial sector development can help reduce information asymmetries for small enterprises. The growth-constraining effect of financing obstacles has been found to be smaller in countries with better developed financial systems, and industries that have naturally more small enterprises grow faster in countries with higher levels of financial development. The positive effect of financial and institutional development can also be observed in the use of external finance. Better protection of property rights increases external financing of small firms significantly more than it does for large firms, particularly due to the differential impact it has on bank and supplier finance (Beck, Demirgüç-Kunt, and Maksimovic, 2008). Easier physical access to banking outlets is also associated with lower financing obstacles (Beck, Demirgüç-Kunt, and Martinez Peria, 2007). Finally, evidence for a sample of European countries shows that financial development enhances new firm entry in sectors that depend more heavily on external finance and that the smallest size firms benefit the most from higher financial development in terms of higher entry rates. The same analysis also shows that financial development promotes the postentry growth of firms in sectors that depend more on external finance (Aghion, Fally and Scarpetta, 2007).

Quasi-natural experimental evidence confirms the importance of credit constraints for firm growth. Banerjee and Duflo (2004) analyzed detailed loan information on 253 Indian small and medium enterprises before and after they became eligible for a directed subsidized lending program and found that the additional credit resulted in a proportional increase in sales rather than a substitution for other nonsubsidized credit, indicating that these firms were credit-constrained before receiving subsidized credit. Similarly, Zia (2008) finds that small

---

34 See, for example, Beck, Demirgüç-Kunt, and Maksimovic (2005) and Beck, Demirgüç-Kunt, and Maksimovic (2008).
nonlisted and nongroup firms in Pakistan reduce their sales after they become ineligible for subsidized export credit, indicating the existence of credit constraints; in contrast, large, listed, and group firms do not reduce their sales after losing access to subsidized credit.

The evidence of access to financial services at the household level is more nuanced. Access to credit is not unambiguously associated with higher welfare; different estimation methods and different samples provide contradictory evidence. More recent evidence, however, has shown a differential impact of improved access to financial services on different household groups (Banerjee et al., 2010), with households that are inclined to become entrepreneurs more likely to do so with improved access to credit or savings services, while others spend more on consumption.

Related to the debate on access to finance by different groups is the question on enterprise versus household credit. Although the theoretical and empirical literature has clearly shown the positive impact of enterprise credit for firm and aggregate growth, theory has made ambiguous predictions on the role of household credit. Although Jappelli and Pagano (1994) argue that alleviating credit constraints on households reduces the savings rate, with negative repercussions for economic growth, Galor and Zeira (1993) and De Gregorio (1996) argue that household credit can foster economic development if it increases human capital accumulation. Tentative cross-country evidence has shown that the positive effect of financial deepening comes mostly through enterprise credit, and there is no significant relationship between the importance of household credit and economic growth (Beck et al., 2009). This finding, together with the observation of an increasing share of household credit in total bank lending in many developed economies over the past decades,

35 See, for example, Pitt and Khandker (1998), Morduch (1998), Coleman (1999), Karlan and Zinman (2010); see Karlan and Morduch (2010) for a recent overview.
36 See also Dupas and Robinson (2009) for an assessment of improved access to savings services in Kenya.
mostly for mortgages, can go some way toward explaining the diminishing growth benefits from financial deepening in high-income countries.

2.6. The Distributional Effects of Financial Development: Theory and Evidence

Given that financial sector development helps reduce access problems, the question arises about distributional effects of financial sector deepening. Although the cross-country literature has focused mainly on average income growth, researchers have recently turned their attention to distributional implications of financial sector deepening.

Theory makes ambiguous predictions about the distributional repercussions of finance. On the one hand, financial imperfections, such as information and transactions costs, are especially binding on the poor, who lack collateral and credit histories, as already discussed. Thus, any relaxation of these credit constraints will disproportionately benefit the poor. Furthermore, these credit constraints reduce the efficiency of capital allocation and intensify income inequality by impeding the flow of capital to poor individuals with high expected return investments.\textsuperscript{37} From this perspective, financial development helps the poor both by improving the efficiency of capital allocation, which accelerates aggregate growth, and by relaxing credit constraints that more extensively restrain the poor, which reduces income inequality.

In contrast, some theories predict that financial development primarily helps the rich. According to this view, the poor rely on informal, family connections for capital, so that improvements in the formal financial sector inordinately benefit the rich. The model by Greenwood and Jovanovic (1990), previously discussed, predicts a nonlinear relationship between financial development, income inequality, and economic development. At all stages of economic development, financial development improves capital allocation, boosts

\textsuperscript{37} See Galor and Zeira (1993), Aghion and Bolton (1997), and Galor and Moav (2004).
aggregate growth, and helps the poor through this channel. However, the distributional effect of financial development, and hence the net impact on the poor, depends on the level of economic development. At early stages of development, only the rich can afford to access and directly profit from better financial markets. At higher levels of economic development, many people access financial markets so that financial development directly helps a larger proportion of society.

First empirical cross-country evidence points to a propoor effect of financial sector deepening. Beck, Demirgüç-Kunt, and Levine (2007) show that countries with higher levels of financial development experience faster reductions in income inequality and poverty levels. Clarke, Xu, and Zou (2006) show a negative relationship between financial sector development and the level of poverty. This suggests that financial sector development is not only progrowth but also propoor. Unlike other policy areas, which might have opposing effects on growth and equity, financial sector development does not present such concerns.

The theoretical models discussed here also give insights into the possible channels through which financial development can help reduce income inequality and poverty. On the one hand, providing access to credit to the poor might help them overcome financing constraints and allow them to invest in microenterprises and human capital accumulation.38

On the other hand, there might be indirect effects through enterprise credit. By expanding credit to existing and new enterprises and allocating society’s savings more efficiently, financial systems can expand the formal economy and pull larger segments of the population into the formal labor market. First explorations of the channels through which finance affects income inequality and poverty levels point to an important role of such indirect effects.

Specifically, evidence from both the United States and Thailand suggests that an important effect of financial sector deepening on income inequality and poverty is indirect. By changing

---

the structure of the economy and allowing more entry into the labor market of previously un-
or underemployed segments of the population, finance helps reduce income inequality and poverty, but not by giving access to credit to everyone. This is also consistent with cross-
country evidence that financial deepening is positively associated with employment growth in developing countries (Pagano and Pica, 2012). It is important to stress that this is preliminary evidence to be confirmed or refuted by future research, but it has focused the debate on an important question: should policy makers focus on deepening or broadening financial sectors? It has also helped broaden the debate on financial services for the poor beyond microcredit to other financial services, such as savings services, payment services (especially in the context of receiving remittances from family members that emigrated to other parts of the country or the world), and insurance services.

2.7. Finance and Economic Development: Conclusions and Looking Forward

There is strong historical, theoretical, and empirical evidence for a positive role of financial deepening in the economic development process. Evidence for cross-country heterogeneity and nonlinearity in this relationship, however, has posed new challenges for researchers and establishes a direct link to the theme of the next section—financial fragility resulting from rapid financial deepening. There have been attempts to reconcile the long-term positive effects of finance with the negative short-term effects of rapid credit growth (Loayza and Rancière, 2006). More research along these lines is certainly needed. Furthermore, recent evidence that financial sector deepening might actually have a negative effect on growth beyond a certain threshold has raised additional questions on the optimal size and resource allocation to the financial sector.

40 For a more in-depth discussion of these issues and the relevant literature, see World Bank (2007).
The increasing availability of micro-data has broadened the research agenda to exploring the effect of broadening access to financial services by enterprises and households, which will give additional insights into the channels through which finance fosters growth and helps reduce poverty. The literature on finance, income inequality and poverty, is still in its early years; more research can be expected in this area. This will also help bring empirical work closer to theoretical explorations of the finance and growth link.

3. Financial Fragility: Causes and Policies

The same mechanism that makes finance growth-enhancing also contains the seed of destruction, as illustrated by the Diamond and Dybvig (1983) model. By transforming short-term liabilities into long-term assets, banks can foster economic growth but can also become susceptible to bank runs, be they informed or uninformed. Agency problems between banks and their depositors and creditors can lead to excessive risk taking and fragility.

In the following, I focus mostly on fragility of banks, although boom-and-bust cycles are very common to financial markets in general and are often related to banking distress. However, due to the maturity-mismatch and promissory intensity, the contagion risk is highest in the banking sector. After covering the sources of banking fragility, I discuss the empirical literature on idiosyncratic and systemic bank fragility, before turning to regulatory issues.

3.1. Financial Fragility: Bank Runs and Moral Hazard

Theoretical models focus on two different sources of fragility related to liability and asset risk. As already discussed, liability risk arises from the maturity mismatch between assets and liabilities. When some depositors withdraw their funds prematurely and unexpectedly, this
can lead to bank runs and collapses. Such runs can be either based on fundamentals, and thus be information-based, or irrational. If the return on banks’ long-term assets is stochastic, new information about future negative shocks on these investments can lead to the expectation by depositors that banks will not be able to meet future commitments and will therefore lead to runs (Jacklin and Bhattacharya, 1988). Irrational bank runs, on the other hand, arise from the simple fear of some depositors that others might withdraw before them. Irrational bank runs might also be based on the inability of uninformed depositors to distinguish between liquidity and solvency shocks of banks, that is, the inability to distinguish between regular withdrawal behavior of depositors and the reaction of informed depositors to negative information about the future solvency of the bank (Chari and Jagannathan, 1988).

A bank run is not only disruptive for the bank in question, as it might imply costly divestment of assets and/or liquidation of assets, it can also have contagion effects throughout the banking and financial system through the domino effect. Such effects can happen either through bank runs on other banks or through the payment system or the interbank market. Observation by depositors of runs on one bank can lead to panic runs on other banks to not be the last one to withdraw money or due to information updates about the underlying solvency position of other banks. Domino effects can also happen through linkages in interbank market, with failure of one bank to satisfy commitments leading to negative solvency shocks at other banks. Similarly, a payment system based on netting out of positions between banks can lead to contagion effects throughout the banking system, unlike the real-time gross payment system, where each transaction is settled separately and immediately.41

A second important risk is on the asset side, related to the principal agent problem between the bank and depositors and other creditors. Previously I discussed the principal

---

41 For other work on contagion through the payment system and interbank market, see Rochet and Tirole (1996), Allen and Gale (2000b), and Freixas, Parigi, and Rochet (2000).
agent problems between banks and borrowers, but there is a similar agency problem, based on asymmetric information, between banks and their depositors and creditors. As can be easily shown, banks’ incentives to properly screen and monitor borrowers and thus ensure repayment decrease in their leverage ratio. This has been also described as the put option character of banking: given limited liability, bank owners have the option to sell the bank with the strike price being the value of liabilities. Whereas depositors bear only the downside risk of banks’ risk decisions, owners and managers (acting in the interest of owners) participate in both the up and down side of these risk decisions. Although this is a common problem throughout corporate finance, the situation is exacerbated in the case of banks by the fact that debtholders are disperse (there are many depositors with small deposits), often uninformed, and always have the incentive to free-ride on the efforts of others, and the opacity of banks’ assets, most of which are not market-priced or priceable.

The fragility of banking and the negative repercussions of bank failure for the financial system and the economy at large has made it one of the most regulated sectors in human history, with the exception of few episodes in modern history, such as free banking in Scotland between 1695 and 1845 and the experience of some U.S. states in the nineteenth century. Deposit insurance has been considered a policy to reduce the likelihood of bank runs because it insures depositors for their savings (Diamond and Dybvig, 1983). Alternatively, liquidity support by a lender-of-last-resort can help address banks’ short-term liquidity problems (Bagehot, 1873). Capital requirements and lending restrictions have been advocated to address the potential for asset fragility, as well as strong supervisors that can replace weak or missing monitoring and discipline from depositors. I return to the regulatory approach toward banking and the financial sector in general later, but first discuss the link between idiosyncratic bank fragility and systemic banking distress.
3.2. From Idiosyncratic to Systemic Distress

Financial history is full of bank failures and financial boom-and-bust cycles, linked to a variety of factors, often with similar features (Reinhart and Rogoff, 2009). To the same extent that well-developed financial systems can foster economic growth, banking crises are often associated with deep economic recessions and long-term negative growth repercussions. Recent comparisons of economic crises have shown that economic recessions related to banking distress tend to be deeper and longer than other recessions.\textsuperscript{42} Specifically, output losses of recessions with credit crunches are two or three times as high as in other recessions. Many of these banking crises are related to the failure of several (rather than single) financial institutions or even systemic distress throughout the banking system. In the following, I discuss the link between idiosyncratic bank failures, due to incentive misalignments and inherent fragility and systemic banking distress.

Due to the agency problems between lenders and borrowers and lenders and depositors, described above, bank credit is inherently cyclical and typically more volatile than the economic cycle. As agency costs on both sides of banks’ balance sheets fluctuate with the business cycle, so do cost of credit and credit flows. Credit booms typically feed on themselves and are often linked with asset price booms. As asset prices rise and thus collateral values, more credit is granted. The reverse can happen during a downturn, where asset prices fall, borrowers’ balance sheets deteriorate, and bank lending is typically reduced at a faster pace than GDP and can in turn dampen real sector activity further. Credit and asset price movements feed on each other, which can lead to upward and downward spirals.\textsuperscript{43}

Beyond credit cycles related to agency costs, financial systems are subject to recurrent systemic fragility, often related to financial liberalization. In the 1970s and ’80s, the search

\textsuperscript{42} See Claessens, Kose, and Terrones (2008).
\textsuperscript{43} See Bernanke and Gertler (1989), Kiyotaki and Moore (1997), and Fisher (1933).
for growth benefits led many countries to liberalize financial systems, privatize government-owned banks, and open capital accounts. These liberalizations often led to credit and asset booms and aggressive risk taking by banks. Though fiscal profligacy, exchange rate policy, and external shocks have also contributed to crises, problems in the financial systems were often at the core of the fragility, and these systemic banking crises put in doubt the overall positive contribution of financial development to economic development. Careful analysis of these crises, however, has shown that often the absence of the necessary regulatory reforms that should accompany liberalization can explain the fragility (Demirgüç-Kunt and Detragiache, 1999). Systemic banking distress is often also related to currency crises. Rapid real exchange rate movements can undermine banks’ solvency position, while the need to support failing banks can undermine exchange rate stability. Common causes might drive both, such as macroeconomic policies. Theoretical and empirical work has confirmed the close interlinkages of banking distress and currency crises.44

Banking crises, however, have not been limited to developing and emerging economies. Even before the current crisis, the 1980s and 90s saw the savings and loan crisis in the United States, the Japanese banking crisis, and several banking crises in Scandinavia. Many of the systemic banking crises in developing and developed countries involved large amounts of nonperforming assets and bank insolvencies, the need for nationalization and recapitalization, bank holidays, and government guarantees for deposits and assets. In many cases, frameworks for systematic work-out of nonperforming assets were created, either by banks on a decentralized basis or by creation of asset management companies.

Statistical analyses of systemic banking crises have pointed to several macroeconomic signals, including real exchange rate appreciation (often linked to rapid capital inflows), low

44 See Chang and Velasco (2001) and Kaminsky and Reinhart (1999)
growth, high real interest rates and inflation, as well as rapid credit growth. Other studies have linked the exchange rate regime, the degree of dollarization, and banking market structure to the likelihood of systemic banking distress.

The costs of systemic banking distress can be substantial, as reported by Laeven and Valencia (2008), reaching over 50 percent of GDP in some cases in fiscal costs and over 100 percent in output loss. Fiscal costs arise typically from recapitalization of failing banks or deposit insurance losses, whereas output costs can arise through several channels, most importantly through firms losing access to external finance. Several papers have shown how borrowers have suffered after their institution failed. Several studies have shown the negative economic repercussions of bank failures in the 1920s and ’30s in the United States and the consequent loss of lending relationships, and documented the decline in lending and local GDP following the closure of a large (solvent) affiliate in a regional bank holding company in Texas in the 1990s (Ashcraft, 2005). Other studies have shown the importance of lending relationships across a sample of Korean firms that worked with either failed or surviving banks after the crisis and the negative effect of bank insolvency announcement during the East Asian crisis on market values of the banks’ borrowers. On a more aggregate level, cross-country comparisons have shown that during banking crises, industries that depend more on external finance are hurt disproportionately more, an effect that is stronger in countries with better developed financial systems.

The trade-off between the positive growth effects of financial deepening following liberalization and the costs of systemic banking distress that also often comes after liberalization has raised the question of relative benefits and costs of liberalization. A cross-

46 See Demirgüç-Kunt and Detragiache (2005) for a survey.
47 Bernanke (1983), Calomiris and Mason (2003), and Kupiec and Ramirez (2009).
48 Ferri, Kang, and Kim (2001), and Djankov, Jindra, and Klapper (2005), respectively.
country comparison of the growth benefits of liberalization and growth costs of subsequent crises, however, has shown that the benefits outweigh the costs significantly, that is, the positive growth effect of financial liberalization is larger than the negative growth effect from a crisis that follows liberalization (Rancière, Tornell, and Westermann, 2006).

Directly related to the transmission of idiosyncratic shocks to systemic distress is the debate on the relationship between competition and banking distress. Some models predict that more concentrated and less competitive banking systems are more stable, as profits provide a buffer against fragility and provide incentives against excessive risk taking. This “charter value” view of banking sees banks as choosing the risk of their asset portfolio.50 In more competitive environment with more pressures on profits, banks have higher incentives to take more excessive risks, resulting in higher fragility. In systems with restricted entry and therefore limited competition, on the other hand, banks have better profit opportunities, capital cushions, and therefore fewer incentives to take aggressive risks, with positive repercussions for financial stability. In addition, in more competitive environment, banks earn fewer informational rents from their relationship with borrowers, reducing their incentives to properly screen borrowers, again increasing the risk of fragility.51 These models thus predict that deregulation resulting in more entry and competition, such as in the United States in the 1970s and 1980s and in many emerging markets, would lead to more fragility.

An opposing view is that a more concentrated banking structure results in more bank fragility. Boyd and De Nicoló (2005) argue lower interest following from higher competition incentivizes borrowers to choose less risky investment projects. Thus, in many

---

50 For theoretical models, see Marcus (1984), Chan, Greenbaum, and Thakor (1986), and Keeley (1990).
parameterizations of the model, Boyd and De Nicoló (2005) find a positive relationship between concentration and bank fragility and thus the probability of systemic distress.\textsuperscript{52}

The empirical evidence on the relationship between competition and stability has not been conclusive, partly due to measurement challenges on both competition and stability. Empirical studies for specific countries—many if not most for the United States—have not come to conclusive evidence for an either stability-enhancing or stability-undermining role of competition. However, two conclusions can be drawn. First, a higher degree of market concentration does not necessarily imply less competition. Specifically, testing for the relationship between market structure and stability and for the relationship between competition and stability does not necessarily yield the same results. Second, as predicted by several theoretical studies, there is an important interaction effect between the regulatory and supervisory framework, on the one hand, and market structure and competitiveness, on the other hand, in their effect on banking system stability, a topic I return to later.

The cross-country literature has found that more concentrated banking systems are less likely to suffer a systemic banking crisis as are more competitive banking systems.\textsuperscript{53} There also seems to be evidence that banks in more competitive banking systems hold more capital, thus compensating for the potentially higher risk they are taking.\textsuperscript{54} This debate is certainly still ongoing and has received additional impetus with the recent crisis.

3.3. Market Discipline or Regulation?

The functions and structure of financial institutions make their failures especially damaging for other financial institutions and the economy at large. The provision of payment services is

\textsuperscript{52} Martinez-Miera and Repullo (2010), however, show that higher interest rates also imply higher interest revenues for banks, which might result in a U-shaped relationship between competition and bank fragility. Similarly, Caminal and Matutes (2002) show that less competition can lead to less credit rationing, larger loans, and higher probability of failure if loans are subject to multiplicative uncertainty.

\textsuperscript{53} Beck, Demirgüç-Kunt, and Levine (2006a) and Schaeck, Cihák, and Wolfe (2009).

\textsuperscript{54} Schaeck and Cihák (2007) and Berger, Klapper, and Turk-Ariss (2009).
only feasible if banks belong to a network, the maturity transformation results in the risk of maturity mismatch and liquidity shortages in the case of shocks such as bank runs, and the screening and monitoring function of financial institutions implies the creation of private information. The consequence of these functions is that the failure of a financial institution results in negative externalities beyond the private costs of failure; it imposes external costs on other financial institutions and the economy at large. As already discussed, these external costs materialize especially in systemic banking crises, though also with idiosyncratic bank failures, increasing in the size and importance of financial institutions.

The external costs that bank failure imposes on the rest of the financial system at large has made banking one of the most regulated sectors. The question, however, is what kind of regulations reduce the risk of bank runs and contagion and of excessive risk taking. Diamond and Rajan (2001) show that the combination of short-term liabilities and long-term assets provides for the necessary incentives of depositors to monitor and discipline banks and thus reduces principal agent problems. Regulations that reduce incentives of depositors to monitor and discipline banks, on the other hand, can exacerbate the principal agent problem and are thus counterproductive. Combining the analysis of fragility risks on asset and liability sides thus provides different insights than when focusing only on the liability side.

This matches with empirical evidence that has pointed to the risks of regulations and mechanisms protecting depositors and reducing possibilities and incentives for them to monitor and discipline banks. Specifically, while deposit insurance schemes have been conceived to protect depositors (especially those with no means to monitor and discipline banks) and to avoid contagion effects, explicit deposit insurance schemes also have perverse incentive effects, as they send the signal that authorities stand ready to bail out banks, and
thus provide incentives for banks to take excessive risks. This has been confirmed by empirical evidence, at least for developing countries.\footnote{See Demirgüç-Kunt and Kane (2002) for an overview.}

Hand in hand with the skepticism about the role of deposit insurance goes an emphasis on mechanisms that encourage and enable depositors and creditors to monitor and discipline banks. This private monitoring view thus focuses on transparency of financial statements, liability of auditors and senior management for financial statements, and the availability of marketable securities, such as subordinated debt, whose holders have an incentive to closely monitor and discipline banks. This view does not minimize the role of supervisors, but emphasizes that supervisors have an important role in enabling the private sector to play an appropriate role in the supervision process.\footnote{For a broader discussion, see Barth, Caprio, and Levine (2006).}

This view stands in contrast to the official supervision view that takes the failure of markets, including depositors, to appropriately monitor and discipline banks as a starting point and therefore posits a strong and active if not interventionist role for supervisors. This implies the right to intervene in good and bad times, screen and replace management, and restrict activities. The two approaches are not exclusive, a country like the United States has both powerful supervisors and the necessary conditions for market discipline, with an important caveat, as will be discussed. In many developing countries, the emphasis has been on building up powerful supervisory authorities, again on the premises that market discipline cannot work in small markets with few sophisticated investors.

The recent crisis seems prima facie a rejection of the private monitoring view, as the market did not price risk accurately. One might also argue that discipline cannot be provided by markets that provide incentives for financial institutions to engage in herding (Boot, 2011) On the other hand, the period up to the crisis can be seen as a period where authorities across
the developed world systematically undermined market discipline by signaling ex ante that systemically important financial institutions would be bailed out.\(^57\) The repercussions to the Lehman Brothers insolvency—seen by authorities as a signal to reestablish market discipline—can be seen as the nail in the coffin to market discipline, as authorities in the future will avoid the messy failure of an institution so central to the global financial system and consequent freeze of global financial markets.

Even where supervisors were powerful—as in the United States—they failed to detain excessive risk taking in time. Levine (2010) argues convincingly that regulatory policies created incentives to aim for short-term profit, while at the same time allowing increased long-term fragility. Regulatory capital arbitrage was allowed whereby risky assets were shifted off the balance sheet, securitized in the form of special investment vehicles, and then put back on the balance sheet in the form of triple A–rated securities that did not need any capital. The crisis thus “represents the unwillingness of the policy apparatus to adapt to a dynamic, innovating financial system” (Levine, 2011).\(^58\) The assessment that the inability of supervisors to detect and reduce fragility at early stages with traditional means contributed to the crisis is also consistent with evidence that the quality of bank supervisory standards, as measured by the Basel Core Principles is not significantly associated with bank stability, with the notable exception of transparency standards (Demirgüç-Kunt, Detragiache, and Tressel, 2008).

The skepticism against supervisory and regulatory standards goes hand in hand with a differenced view on capital regulation (Barth, Caprio and Levine, 2006). While designed as cushion against unexpected losses and to reduce incentives to “bet-the-bank”, bank governance and ownership structure critically influence the impact of capital regulations on

\(^{57}\) On a macroeconomic level, this is often referred to as the Bernanke-Greenspan put, in reference to the two chairmen of Federal Reserve in office during the boom period of the early 2000s.

\(^{58}\) For a more in-depth discussion, see also Barth, Caprio and Levine (2012).
risk-taking, as shown by several empirical studies. Specifically, Laeven and Levine (2009) find that banks with more powerful owners, as measured by the size of their shareholdings, tend to take greater risks, and even more so in countries with more stringent capital requirements, while capital regulations have the opposite effects on widely-held banks dominated by managers.

Although their effect in preventing systemic banking distress is doubtful, powerful supervisors can have a negative effect on the efficiency and fairness of lending, as shown by Beck, Demirgüç-Kunt, and Levine (2006b). Specifically, self-reported obstacles on corruption in loan officers’ lending decisions are higher in countries with more powerful supervisors, and they are lower in countries where supervisory entities rely more on market power.

There is similar and consistent evidence for financial markets (La Porta, Lopez-de-Silanes, and Shleifer, 2006). Specifically, there is limited evidence for a positive role of public enforcement benefiting stock market development, while disclosure and liability rules enabling private monitoring foster larger and more liquid stock markets.

3.4. Financial Fragility: A New Research Agenda

The first global financial crisis of the twenty-first century has provided new challenges and new opportunities for exploring the causes of financial fragility and policy implications. It has reignited interest in the competition-stability debate. It has shed doubts on risk measures. It has shown that runs can not only happen on the retail level; new forms of contagion, such as through common asset exposure, have occurred. The crisis has also shown the urgent need for bridging the gap between macro- and financial economists, where the former focus on macroeconomic models without a properly modeled financial system and the latter ignore

macroeconomic repercussions of financial sector fragility. The interest in bridging the gap has become obvious in recent work incorporating the two literatures and also in the debate on macroprudential regulation that looks beyond idiosyncratic risks toward the interaction between financial institutions and different segments of the financial system and feedback loops between banking, financial markets, and the real economy. This also includes recent papers that empirically explore the relationship between monetary policy, securitization through financial markets and banks’ risk taking.⁶０

The recent crisis has also reignited the debate on the appropriate regulation of the financial system. Stronger capital requirements and more powerful supervisors are called for. One important area, often ignored, that calls for intensive reform is the resolution of failing banks. A similar trade-off as for deposit insurance exists: on the one hand, imminent failure of a large, too-important-to-fail bank provokes an immediate bail-out decision to protect depositors, other financial institutions, and the financial system at large as well as the real economy that—as shown—will suffer from bank failure. On the other hand, the expectation of a bail-out creates perverse incentives on the side of banks to take excessive risks, knowing that the downside part of these risks will be covered by taxpayers. Enforcing market discipline, however, by forcing financial institutions into regular insolvency proceedings, as with any regular corporation, can lead to the effects already described and observed in the case of Lehman Brothers. A solution that minimizes these external effects of bank insolvency while at the same time enforcing market discipline is thus necessary, especially for systemically important financial institutions. Recent suggestions include ex ante planned winding-down plans (living wills) and the issue of debt instruments that are automatically converted into equity claims when the value of existing equity nears zero.

⁶₀ See for example, Jimenez et al. 2012 and Mian and Sufi, 2009.
Financial crises offer opportunities and challenges for researchers; they force them to rethink existing paradigms, develop new models, and reassess empirical relations. Similarly to the Great Depression of the twentieth century, the first global crisis of the twenty-first century will have an important impact on how researchers, analysts, and policy makers view financial fragility and the regulatory framework necessary to reduce this fragility and its impact.

4. The Politics of Financial Development

Given the important of the financial system for economic growth and fragility, it is not surprising that financial sector policies often top the agenda of policy makers, though more so during crisis periods. As I discuss, however, there are different views on where this interest stems from—to maximize the growth benefits for the economy or to protect the interests of the incumbent elite. 61

4.1. Finance and Politics

There are different hypotheses of why financial institutions and markets are so high on policy makers’ list. The public interest view argues that policy makers act in the best interest of society, ultimately maximizing the social planner’s problem, though possibly with less information available. This view also argues that the market failures inherent in financial markets and already discussed require a strong government involvement in the financial system beyond regulation and supervision. The private interest view, on the other hand, argues that policy makers, including regulators, act in their own interest, maximizing private

61 For a recent more complete survey, see Haber and Perotti (2008).
rather than public welfare. Politicians thus do not intervene into the financial system to further public welfare but to divert the flow of credit to politically connected firms.62

The importance of access to external finance for entrepreneurs makes it an important tool in the struggle for real market shares. Access to finance can be used as barrier to entry into the real economy; empirical analysis has shown that exit rates during banking crises—especially for finance-dependent young firms—are abnormally high in countries with more corrupt political institutions (Feijen and Perrotti, 2005). Access to external finance can also be used by incumbent political and economic elites to protect rents and entrench their dominant position. The use of financial as opposed to other resources is facilitated by the contingent nature of the liabilities as well as the authority of governments to create money.

The U.S. financial history is plenty of examples for political influence on the banking system. Benmelech and Moskowitz (2010) show a strong relationship between voter suffrage and financial regulation. In times of low suffrage, the ruling elite used interest rate ceilings and entry barriers into banking to prevent competitors from gaining strength. Granting banking licenses was used to gain access to preferable loans for states and for the ruling elites. As voting suffrage expanded, there was a trend toward free banking where licensing was no longer controlled by the state legislature.

One of the striking elements of the U.S. banking history is the predominance of branching restrictions throughout long periods, results of a compromise between populists and local bankers who wanted to protect their rents. Not until the 1970s and under the impact of technology were these restrictions loosened. As shown by Kroszner and Strahan (1999), the decision to deregulate branching on the state level was a function of both political structure and the lobbying strength of the banking sector.

Recent theoretical and empirical work has modeled and estimated the relative power and coalitions between labor and firm insiders, including management and majority shareholders, against minority shareholders to explain cross-country differences in contractual institutions underpinning financial institutions and markets.\textsuperscript{63} An alternative explanation is the experience of the middle class losing their financial assets due to inflation in the period between the wars and subsequent political resistance against vibrant financial markets. This went hand in hand with a decision in these countries toward state-funded and -managed pension funds, with lower need and demand for private pension funds and thus lower demand for protection of individual investor rights (Perotti and Schwienbacher, 2009).

4.2. Finance and Politics: From Government Banking to Activist Policies

Government intervention in the financial sector has been strong throughout history. Some countries initially only allowed government-owned banks or at a minimum made bank licensing subject to parliamentary approval. Many developing countries nationalized their banking system after independence to gain direct control over this critical part of the economy. Across the globe, government ownership has been widespread throughout the twentieth century and today is still dominant in some developing countries.\textsuperscript{64} In this context, it is important to distinguish between government ownership as long-term policy and as crisis resolution tool, as applied after the 2008 global financial crisis in many European countries. Government-owned banks have often been seen as critical in helping overcome market failures and funnel domestic savings to strategically important projects (Gerschenkron, 1962). On the other hand, government ownership can lead to inefficient allocation of scarce resources and political capture. These different views of government ownership relate

\textsuperscript{63} Pagano and Volpin (2005) and Perrotti and von Thadden (2006).
\textsuperscript{64} See La Porta, Lopez-de-Silanes, and Shleifer (2002).
directly to the public and private interest view on government’s involvement in the financial sector, as already discussed.

A large empirical literature has shown that government ownership is inefficient. Government-owned and -managed banks run the risk of being captured by the ruling elite or special interest and have – on average - a miserable record in expanding access. Firms with political connections have easier access to state banks and receive larger loans but are less likely to repay.\textsuperscript{65} On an aggregate level, government ownership of banks is associated with lower levels of financial development and lower rates of economic growth (La Porta, Lopez-de-Silanes, and Shleifer, 2002).

Privatization, however, is not a panacea, especially if undertaken in an institutionally weak environment, and can lead to capture by socioeconomic elites linked to political elites. Poorly designed and executed privatization processes can lead to fragility and banking crises, as numerous examples have shown over the past thirty years. Studies of privatization processes have shown the benefits of privatizing government-owned banks but also the pitfalls.\textsuperscript{66} One interesting case is Mexico, where privatization in 1988 was restricted to domestic shareholders, most of whom had no banking experience and borrowed money from their own banks to acquire the banks from the government. A subsequent boom-and-bust period led to new nationalization and recapitalization episode in the mid-1990s, after which most of the banks were sold to multinational banks.

Beyond ownership, government intervention into the financial system can take many forms. Excessive reserve requirements, interest rate ceilings and floors, and credit quotas are some of the policies that collectively are referred to as financial repression (Fry, 1988). As shown mostly by country-level studies, some of them of qualitative rather than quantitative

\textsuperscript{65} Faccio (2006), Khwaja and Mian (2005), and Claessens, Feijen, and Laeven (2008).
\textsuperscript{66} Compare Clarke, Cull, and Shirley (2005) and other articles in the same issue.
nature, most of these policies have benefited the incumbent elite and enterprises connected to it, rather than marginal groups. Credit quotas and interest caps and floors have impeded the efficient allocation of society’s savings to its most productive uses and have especially hurt “smaller” depositors and borrowers. In case of binding ceilings, banks are prevented from charging adequate risk premiums for riskier and more opaque borrowers or from recovering fixed transaction costs through a mark-up on smaller loan amounts. Furthermore, competition between credit institutions and for more deposits is hampered as financial institutions have no incentives to become more efficient or to attract more deposits if they cannot finance more marginal customers. Similarly, given fixed transaction costs in financial intermediation, floors on deposit interest rates make savers with small transaction amounts unattractive for financial institutions. Credit quotas have resulted in fragmentation of credit markets and higher costs for nonpriority sectors. In many cases, financial institutions have found ways around these restrictions, but at high costs and with consequent efficiency losses. Such policies are especially strong in sectors that are considered marginalized or as politically critical, such as rural areas.

The failure of many of these interventionist policies have led to a move toward more market-based financial systems over the past two decades, often associated with the Washington Consensus. Many developing countries have moved toward more stable banking systems, and some have experienced financial deepening. Expectations of more inclusive financial systems and more significant financial deepening have often not been achieved, however. Although this raises questions on the proper sequencing and implementation of reform policies to deepen and broaden financial systems, it has also led to a renewed debate on the role of government in the financial system, beyond institution building and providing the macroeconomic framework. Activist or market-enabling policies that try to overcome
coordination and first mover problems in small financial systems with incomplete markets have been suggested, such as attempts to provide the necessary infrastructure to launch new financial products, such as factoring, or credit guarantee programs.\textsuperscript{67}

One region where market-based financial sector reform has been successful is the transition economies of Central Europe. The challenge for these countries was to make banks independent from government and from their past links with state-owned enterprises, as the continuing relationships between banks and incumbent enterprises and the resulting fragility had severe macroeconomic repercussions. The need for recapitalization of banks resulted in rising fiscal deficits, monetary overhang, and thus inflation. The solution to this continuous cycle of repayment problems, accumulation of nonperforming assets, recapitalization, and inflation was the adoption of a disciplining tool to impose a hard budget constraint on enterprises and banks alike. Credibly committing to monetary stability in turn forced the necessary reforms in the financial sector to avoid future recapitalization. In many countries, banks were therefore not only privatized but sold to foreign banks, which helped sever the links between state-owned enterprises and banks.\textsuperscript{68} What essentially was needed was a straightjacket that tied policy makers’ hands and prevented them from bailing out financial and nonfinancial institutions.

4.3. Political Structure as Basic Factor for Financial Development

How do some countries develop the necessary legal and regulatory structures to support market-based financial systems and others do not? Why do some countries have political structures that are conducive to financial development and property rights protection, more

\textsuperscript{67} For a more detailed discussion, see De la Torre, Gozzi, and Schmukler (2007).

\textsuperscript{68} See Giannetti and Ongena (2009).
broadly, and others do not? Here the literature on the historic roots of financial sector
development overlaps with the literature on institutions and growth.

The government’s position as party to (financial) contracts and arbiter of the same
contracts creates a conflict of interest and make the political Coase theorem that distribution
and efficiency can be determined independently break down (Acemoglu, 2003). A basic
condition for thriving financial markets is thus constraints on political power and the
protection of individual property rights from expropriation through both other private parties
and the government. It is important to stress in that context that property rights must be for
everyone, not just the elite. Autocratic regimes dislike independent and competitive financial
systems because they are afraid they might finance opponents. On the other hand, the
financial system can serve as source of rents for the ruling elite, as especially the case of
resource-based economies has shown. Evidence from broad cross-country samples shows that
countries with autocratic political regimes are more likely to have restrictive regulation and
entry barriers into the banking systems (Barth, Caprio, and Levine, 2006), whereas the most
robust predictor of long-term financial development (as opposed to short-term boosts) is
political accountability (Quintyn and Verdier, 2010), evidence that is consistent with historic
analysis that shows for a panel of countries over the period 1880–1997 that more restrictions
on political power and stable political regimes are more conducive to financial development
(Bordo and Rousseau, 2006). Countries with captive political institutions also suffer more
from financial instability (Acemoglu et al., 2003). The political environment is also important
for the effectiveness of financial liberalization, which is more likely to produce instability in
countries with captive political institutions (Bekaert, Harvey, and Lundblad, 2006).

Historically, the financial centers of modern history in Europe developed in
independent cities in northern Italy and, later in sixteenth century, the Netherlands, both areas
where government was supported by broad parts of the population and with systems of checks and balances. The Glorious Revolution in seventeenth-century England secured property rights from government expropriation and reinforced the status of an independent judiciary. This enabled the British Crown to borrow at much lower interest rates in the international markets, as these institutions reassured investors of a low default risk (North and Weingast, 1989).

There are two sets of theories—not necessarily exclusive—that explain the development of the necessary property rights and contractual institutions for financial development across the world. One set of theories sees historical events in Europe more than 200 years ago as shaping the legal and regulatory frameworks across the globe today through their influence on political structures in these countries. Specifically, the legal origin theory sees political conflicts in England and France in the medieval age and during the Glorious and French Revolutions shaping the role and independence of judiciaries in these countries. Different points on the trade-off between centralized power to avoid civil unrest and freedom to allow economic activity in England and France during medieval times shaped the government’s approach to the judiciary, with France taking a much more centralized approach than England (Glaeser and Shleifer, 2002). Alternatively, one can consider the role of the judiciary during the Glorious Revolution, where the judges sided with the winning Parliament, and the French Revolution, where the judges were on the losing side. This resulted in a strengthening of the judiciary’s independence but also their role in lawmaking in England, while it reduced the judiciary to an executing role in France, with law- and rule-making concentrated in legislature and executive. However, this also resulted in a different degree of flexibility and adaptability of the legal systems in England and France. England’s legal system was more adaptable due to a stronger role for jurisprudence and past decisions
and the ability of judges to base decisions on principles of fairness and justice, whereas France’s legal system was more rigid, based on bright-line rules and little if any role for jurisprudence and previous decisions.69

Through the Napoleonic Wars in the early nineteenth century, the Napoleonic legal tradition was spread throughout continental Europe. Subsequently, legal traditions were spread throughout the rest of the world, mostly in the form of colonization, with the British common law tradition adopted in all British colonies and the Napoleonic civil code tradition transplanted to Belgian, Dutch, Portuguese, and French colonies. The legal structures originating in these different traditions have proven to be very persistent, especially in developing countries. Take the example of the Napoleonic legal tradition. First, while the European nations overcame the rigidities of the Napoleonic code, they exported its antagonism toward jurisprudence and its reliance on judicial formalism to minimize the role of judges. This comes with the tradition of avoiding open disputes about legal interpretation and the aversion against jurisprudence. Second, as the Napoleonic doctrine sees judges as purely executing civil servants, judges frequently “are at the bottom of the scale of prestige among the legal professions in France and in many nations that adopted the French Revolutionary reforms, and the best people in those nations accordingly seek other legal careers” (Merryman, 1996, p. 116). Third, and as a consequence of the previous point, there is a stronger reliance on bright-line laws to limit the role of the courts. Once a country adopts the bright-line approach to lawmaking, this can lead into a trap, as courts will not be challenged to develop legal procedures and methods to deal with new circumstances, thus retarding the development of efficiently adaptive legal systems (Pistor et al., 2002, 2003).

69 Other important groups constitute the German and the Scandinavian legal systems, which are based on similar political structures as the French civil code tradition but have a more flexible and adaptable structure.
Empirical evidence has indeed shown that countries with a Napoleonic legal tradition have less independent judiciaries and less adaptable legal systems.\textsuperscript{70} Important in the context of this paper, countries with a Napoleonic legal tradition have—on average—weaker property rights protection and contractual institutions that are less conducive to external finance, including weaker protection for minority shareholders and secured and unsecured creditors. Enforcement of contracts is more costly and slower in civil code countries as is the registration of property and collateral. This has the overall effect of smaller and less effective financial markets in civil code countries (Beck, Demirgüç-Kunt, and Levine, 2003a).

An alternative explanation refers not to the identity of the colonizing power but the mode of colonization. Distinguishing between settler and extractive colonies, Acemoglu, Johnson, and Robinson (2001, 2002) show that the former developed stronger property rights protection than the latter, given the political and societal structures that natural resource extraction in the latter implied. The initial colonization mode, in turn, was determined by the disease environment that European colonizers encountered as well as the incidence of native population in the colonized areas. Areas with more hostile disease environments and/or large native population concentrations were more likely to be settled in an extractive mode. The political structures developed during the colonization period endured after independence, therefore also making the weak property rights and contract enforcement institutions persistently weak beyond independence.

Empirical evidence shows the importance of the colonization mode for the development of financial markets today (Beck, Demirgüç-Kunt, and Levine, 2003a). Countries that were initially colonized in an extractive mode have less developed financial markets today. This effect is in addition to the effect of the legal tradition already discussed.

\textsuperscript{70} La Porta et al. (2004) and Beck, Demirgüç-Kunt, and Levine (2003b).
Beyond the colonization experience, the legal tradition and endowment views show the importance of political structures and persistence in financial system development. These hypotheses suggest that changes in the legal institutions that underpin thriving financial markets are only possible under outside pressure or exogenous shocks, such as new technologies, diseases, or globalization. Similarly, changes in financial sector policies are more likely under exogenous pressure. I already discussed the example of the Central European transition economies where banking crises and pressure to establish macroeconomic stability forced privatization of banks to foreign banks. Similarly, in Brazil the introduction of the Real Plan in 1994 that terminated the long-running inflationary tradition prevented the government from bailing out banks owned by individual states, as it had done several times earlier, and thus forced a complete restructuring of these institutions (Beck, Crivelli, and Summerhill, 2005). In Argentina, the establishment of a currency board in 1991 started the restructuring process of provincial banks (Clarke and Cull, 2002).

Technological innovation was critical in driving branch deregulation in the United States in the 1970s and 1980s. As shown by Kroszner and Strahan (1999), the invention of automatic teller machines (ATMs), in conjunction with court rulings that ATMs are not bank branches, weakened the geographical bond between customers and banks, and improvements in communications technology lowered the costs of using distant banks. These innovations reduced the monopoly power of local banks, weakening their ability and desire to fight against deregulation, ultimately leading to branch deregulation. The timing of this deregulation across states, in turn, was very much a function of initial conditions, ranging from party politics to the importance and independence of insurance companies.
5. Conclusions

This paper surveyed three related strands of the literature—finance and growth, financial fragility, and finance and politics. The three literatures are closely linked to each other. The growth benefits of financial sector deepening and the fragility of banking are two sides of the same mechanism—maturity transformation. While overcoming agency problems between investors and entrepreneurs is an important growth-enhancing role of financial institutions and markets, agency problems between financial institutions and their depositors are the basis for possible financial fragility. Empirically, long-term financial deepening is related to faster economic growth, and short-term credit booms are related to a higher probability of systemic banking distress. Underlying both growth and fragility are political constraints, often related to historic conditions, that prevent the necessary property rights and contractual institutions to develop and might foster connected and politicized lending.

Throughout the paper I discussed the linkages between the three strands of literature. One example—already referred to—are the transition economies of Central and Eastern Europe, which experienced rapid financial deepening over the past twenty years, in line with the transition to market-based economies. The deepening was stronger where the underlying political institutions allowed for checks and balances, property rights protection, and conducive contractual institutions. However, in some countries financial deepening turned into a household and mortgage credit boom, ultimately increasing fragility and ending in a bust during the recent global financial crisis.

The current crisis, especially the development of financial markets in the United States, also offers an interesting background on which to bring together the three themes of this paper. First, as discussed, financial liberalization in the 1970s and ’80s has helped deepen the financial system, with positive repercussions for growth and smoothing volatility.
Financial liberalization, however, has also created the basis for a boom-and-bust cycle, with banks taking increasing risks. There has also been evidence for a herding effect with financial institutions taking increasingly risks in the same sectoral and geographical portfolios. Behind the rapid growth of the subprime mortgage segment, there was a political focus on home ownership for low-income Americans. As laid out convincingly by Rajan (2010), in the absence of easy solutions to reduce income inequality, there was a political focus on reducing consumption inequality, which included boosting access to credit. Government policies such as the Community Reinvestment Act and guarantees provided by government-sponsored financial institutions, such as Fannie Mae and Freddie Mac, created incentives to look beyond the usual risk-return trade-off. However, there were also important regulatory distortions, as described by Levine (2010, 2011). But maybe one of the decisive incentives was provided by the Bernanke-Greenspan put—the expectation of financial market participants to be bailed out by the U.S. authorities in case of distress, be it through monetary policy or more direct intervention.

What effect does the current crisis have on our thinking about the financial sector and its role in the economy? The picture of the financial system as an unconditionally growth-enhancing sector has been tarnished. Excesses have become clear, as also obvious from wage trends in the financial sector, documented by Phillipon and Reshef (2009). In the United States, wages in the financial sector relative to general wage levels have increased substantially during the recent boom period, to a comparable extent as in the 1920s in the period leading up to the Great Depression. Most critically, the current crisis should return the debate on financial sector development to the benefits of financial services and, more specifically, financial intermediation.
Although academics refer mostly to the intermediation functions of financial systems and thus to a facilitating role of the financial sector, practitioners and policy makers often view financial services as a growth sector in itself. This view toward the financial sector sees it more or less as an export sector, that is, one that seeks to build an often nationally centered financial stronghold by building on relative comparative advantages, such as skill base, favorable regulatory policies, subsidies, and so on. Economic benefits discussed often point at important spin-offs coming from professional services (legal, accounting, consulting, etc.) that typically cluster around a financial center. Reconciling these different views will be a challenge going forward.

For better or worse, the financial sector is an integral part of modern market economies. Well-functioning and sound financial institutions and markets underpin the smooth exchange of services and goods and foster long-term investment and thus growth. Aggressive risk taking and distortions can lead to systemic distress and economic crises. It is important, however, to not throw out the baby with the very dirty bathwater, but to focus on the necessary regulatory and supervisory frameworks for sound and effective financial systems. Similarly, a better understanding of the political economy of financial sector policies is important. An increasing governance focus is important—not only for central banks and regulatory authorities but also for financial institutions and their relationships with political and regulatory authorities.
References


Kletzer, Kenneth, and Pranab Bardhan. “Credit Markets and Patterns of International Trade.” 


Growth, ed. Philippe Aghion and Steven N. Durlauf, 865–934. Amsterdam: Elsevier,
2005.

Levine, Ross. “An Autopsy of the U.S. Financial System: Accident, Suicide or Negligent


Levine, Ross, Norman Loayza, and Thorsten Beck. “Financial Intermediation and Growth:


Loayza, Norman V., and Romain Rancière. “Financial Development, Financial Fragility, and


Malmendier, Ulrike. “Law and Finance ‘at the Origin.’” Journal of Economic Literature 47

Manova, Kalina. “Credit Constraints, Equity Market Liberalizations and International Trade.”

Manova, Kalina. “Credit Constraints and the Adjustment to Trade Reform.” In Trade
Adjustment Costs in Developing Countries: Impacts, Determinants, and Policy


Stiglitz, Joseph E. “Credit Markets and the Control of Capital.” *Journal of Money, Credit and Banking* 17 (May 1985): 133–152.


