Financial System: Shock Absorber or Amplifier?[*]

Franklin Allen
University of Pennsylvania

and

Elena Carletti
Center for Financial Studies

June 10, 2007

Abstract

The banking sector is one of the most highly regulated sectors in the economy. However, in contrast to other regulated sectors there is no wide agreement on the market failures that justify regulation. We suggest that there are two important ones. The first is a coordination problem that arises because of multiple equilibria. If people believe there is going to be a panic then that can be self-fulfilling. If they believe there will be no panic then that can also be self-fulfilling. Policy analysis is difficult in this case because our knowledge of equilibrium selection mechanisms is limited. Global games represent one promising modeling technique but as yet there is very little empirical evidence in support of this approach. The second market failure is that if there are incomplete markets the provision of liquidity is inefficient. In particular there must be significant price volatility in order for the providers of liquidity to earn the opportunity cost of holding liquidity. We argue that financial fragility, contagion, and asset price bubbles are manifestations of poor liquidity provision. In the absence of a market failure the financial system acts as a shock absorber. However, if there is a market failure it can be an amplifier.

[*] Prepared for the Sixth BIS Annual Conference 2007, Financial System and Macroeconomic Resilience, 18-19 June 2007, Brunnen, Switzerland.

1. Introduction

In recent decades there has been significant deregulation in many industries. A sector that remains heavily regulated is banking. Why is banking so heavily regulated? One reason is consumer protection but this is relatively minor. The main reason for banking regulation is to prevent financial crises. However, banking regulation is unusual compared to other types of regulation in that there is not wide agreement on what the market failure is that justifies regulation.

With other types of regulation there typically is agreement. For example, antitrust regulation is necessary to prevent the pernicious effects of monopoly. The market failure is the lack of competition. With environmental regulation, there is a missing market. Polluters do not have to pay a price to compensate the people they harm. If there was a market where they did have to do this there would be an efficient allocation of resources and no need for intervention. But there isn't such a market and it is necessary to regulate instead. In contrast, with banking what is the market failure that justifies so much regulation? The purpose of this paper is to address this question and examine the implications for the role of the financial system as a shock absorber or amplifier.

Many banking regulations in the U.S. were originally introduced as a reaction to the banking crises in the early 1930's and the perception that these were an important contributing factor to the severity of the Great Depression. The experience was so awful that it was widely agreed that this must never be allowed to happen again and extensive banking regulation was introduced as a result. The regulation wasn't guided by theory but instead was a series of piecemeal reforms. In many European countries, such as France and Sweden, the response was much stronger and involved government ownership

of the banking sector. Either through regulation or public ownership the banking sector was highly controlled.

These reforms were very successful in terms of preventing banking crises. From 1945-1971 there was only one banking crisis in the world. That was in Brazil in 1962 when it occurred together with a currency crisis. Apart from that there was not a single banking crisis (see Bordo et al. (2001)). The reason that crises were prevented is that risk taking and competition were controlled so much that the financial system ceased to perform its function of allocating resources efficiently. The financial repression that resulted from excessive regulation and public ownership eventually led to pressures for financial liberalization. Starting in the 1970's, regulations were lifted and in many countries with government ownership banks were privatized.

Financial liberalization not only allowed the financial system to fulfill its role in allocating resources. It also led to the return of banking crises and there have been many in the last three decades. Many have been in emerging countries but many have also been in developed countries such as those in Norway, Sweden, and Finland in the early 1990's. Bordo et al. (2001) find that the frequency of crises in the recent period since 1971 is not that different from what it was before 1914.

There is a large literature on the costs of crises and their resolution (see, e.g., Bordo et al. (2001), Hoggarth et al. (2002), Boyd et al. (2005), and Honohan and Laeven (2005)). Much of the debate has been concerned with how exactly to measure costs. A large part of the early literature focused on the fiscal costs. This is the amount that it costs the government to recapitalize banks, reimburse insured depositors, and possibly other creditors. However, these are mostly transfers rather than true costs. The subsequent

literature has focused more on the lost output relative to a benchmark such as trend growth rate.

There are two important aspects of the costs of crises when measured this way. The first is the high average cost and the second is the large variation in the amount of costs. Boyd et al. (2005) estimate the average discounted present value of losses in a number of different ways. Depending on the method used the mean loss is between 63 percent and 302 percent of real per capita GDP in the year before the crisis starts. The range of losses is very large. In Canada, France, Germany, and the U.S., which experienced mild non-systemic crises, there was not any significant slowdown in growth and costs were insignificant. However, at the other extreme the slowdown and discounted loss in output were extremely high. In Hong Kong the discounted PV of losses was 1,041 percent of real output the year before the crisis. The variation in costs underlies how important is the issue of whether the financial system is a shock absorber or amplifier.

It is the large average costs and the very high tail costs of crises that make policymakers so averse to crises. This is why in most cases they go to such great lengths to avoid them. However, it is not clear that this is optimal. There are significant costs associated with regulations to avoid crises and in many cases the expected costs of crises are not very high. But what are these costs of regulation? Are crises always bad or can they sometimes be advantageous? Once again the key issue is what exactly is the market failure?

The Basel agreements illustrate the lack of agreement on the basic underlying market failure. An enormous amount of effort has been put into designing these rules. Billions of dollars have been expended by the banks in setting up systems to implement

them. They provide an example of regulation that is empirically rather than theoretically motivated. Practitioners have become experts at the details of a highly complex system for which there is no widely agreed rationale based in economic theory. What is the optimal capital structure? What market failure necessitates the imposition of capital adequacy requirements? Why can't the market be left to determine the appropriate level of capital? There are not good answers to these questions in the theoretical literature.

The key point is that just because there is asymmetric information of some kind does not necessarily mean there is a market failure and intervention is justified. It must be shown that the government can do better than the market. In the literature on capital adequacy, it is often argued that capital regulation is necessary to control the moral hazard problems generated by the existence of deposit insurance. Partial deposit insurance was introduced in the U.S. in the 1930s to prevent bank runs or, more generally, financial instability. Because banks issue insured debt-like obligations (e.g., bank deposits) they have an incentive to engage in risk-shifting behavior. In other words, the bank has an incentive to make excessively risky investments, because it knows that in the event of failure the loss is borne by the deposit insurance fund and in the event of success the bank's shareholders reap the rewards. The existence of bank capital reduces the incentive to take risks because, in the event of failure, the shareholders lose their capital. Thus, capital adequacy requirements are indirectly justified by the desire to prevent financial crises.

However, any analysis of optimal policy must weigh the costs and benefits of regulation. This can only be done in a model that explicitly models the possibility of crises. In the absence of explicit modeling of the costs of financial crises, it is difficult to

make a case for the optimality of intervention. As a corollary, it is difficult to make a case for capital adequacy requirements as a means of offsetting the risk taking generated by deposit insurance.

In this paper we consider the market failures that are associated with banking. The framework is based on that developed in Allen and Gale (2004a, 2004b, 2007) and Allen and Carletti (2006, 2007). We argue that the key issue that determines whether the financial system is a shock absorber or amplifier is whether there is a market failure. Without a market failure the financial system is a shock absorber. With a market failure, it is an amplifier.

2. Panics versus Fundamentals

Two approaches to crises can be developed. Both views of crises have a long history. One view, well expounded in Kindleberger (1978), is that they occur spontaneously as a panic. The modern version was developed by Bryant (1980) and Diamond and Dybvig (1983). The analysis is based on the existence of multiple equilibria. There is a panic in at least one equilibrium while in another there is not.

The second view asserts that crises arise from fundamental causes that are part of the business cycle (see, e.g., Mitchell (1941)). The basic idea is that when the economy goes into a recession or depression the returns on bank assets will be low. Given their fixed liabilities in the form of deposits or bonds banks may be unable to remain solvent. This may precipitate a run on banks. Gorton (1988) showed empirically that in the U.S. in the late nineteenth and early twentieth centuries, a leading economic indicator based on

the liabilities of failed businesses could accurately predict the occurrence of banking crises.

Panics

The panics view suggests that crises are random events, unrelated to changes in the real economy. The classical form of this view suggests that panics are the result of "mob psychology" or "mass hysteria" (see, e.g., Kindleberger (1978)). The modern version, developed by Bryant (1980) and Diamond and Dybvig (1983), is that bank runs are self-fulfilling prophecies. Given the assumption of first-come, first-served and costly liquidation of some assets there are multiple equilibria. If everybody believes no panic will occur only those with genuine liquidity needs will withdraw their funds and these demands can be met without costly liquidation of assets. However, if everybody believes a crisis will occur then it becomes a self-fulfilling prophecy as people rush to avoid being last in line. Which of these two equilibria occurs depends on extraneous variables or "sunspots". Although sunspots have no effect on the real data of the economy, they affect depositors' beliefs in a way that turns out to be self-fulfilling.

The key issue in theories of panics is which equilibrium is selected and in particular what is the equilibrium selection mechanism. Sunspots are convenient pedagogically but this explanation does not have much content. It does not explain why the sunspot should be used as a coordination device. There is no real account of what triggers a crisis. This is particularly a problem if there is a desire to use the theory for policy analysis.

Carlsson and van Damme (1993) showed how the introduction of a small amount of asymmetric information could eliminate the multiplicity of equilibria in coordination games. They called the games with asymmetric information about fundamentals global games. Their work showed that the existence of multiple equilibria depends on the players having common knowledge about the fundamentals of the game. Introducing noise ensures that the fundamentals are no longer common knowledge and thus prevents the coordination that is essential to multiplicity. Morris and Shin (1998) applied this approach to models of currency crises. Rochet and Vives (2004) and Goldstein and Pauzner (2005) have applied the same technique to banking crises.

Using a global games approach to ensure the uniqueness of equilibrium is theoretically appealing. It specifies precisely the parameter values for which a crisis occurs and allows a comparative static analysis of the factors that influence this set. This is the essential analytical tool for policy analysis. However, what is really needed in addition to logical consistency is empirical evidence that such an approach is valid. Currently there is a very limited empirical literature. This is in the context of currency crises and is broadly consistent with the global games approach (see Prati and Sbracia (2002), Tillman (2004), Bannier (2005), and Chen et al. (2007)).

In terms of answering the question what is the market failure, the coordination problem that leads to panics is one possible answer. The problem is that any serious policy analysis requires a theory of equilibrium selection. However, this is not something on which much progress has been made. Global games provide one possible approach but there is currently little evidence on how empirically relevant this approach is.

Fundamentals

An alternative to the sunspot view is that banking crises are a natural outgrowth of the business cycle. An economic downturn will reduce the value of bank assets, raising the possibility that banks are unable to meet their commitments. If depositors receive information about an impending downturn in the cycle, they will anticipate financial difficulties in the banking sector and try to withdraw their funds. This attempt will precipitate the crisis. According to this interpretation, crises are not random events but a response to unfolding economic circumstances.

A number of authors have developed models of banking crises caused by aggregate risk. For example, Chari and Jagannathan (1988) focus on a signal extraction problem where part of the population observes a signal about future returns. Others must then try to deduce from observed withdrawals whether an unfavorable signal was received by this group or whether liquidity needs happen to be high. Chari and Jagannathan are able to show crises occur not only when the outlook is poor but also when liquidity needs turn out to be high.

Building on the empirical work of Gorton (1988) that nineteenth century banking crises were predicted by leading economic indicators, Allen and Gale (1998) develop a model that is consistent with the business cycle view of the origins of banking crises. They assume that depositors can observe a leading economic indicator that provides public information about future bank asset returns. If there are high returns then depositors are quite willing to keep their funds in the bank. However, if the returns are sufficiently low they will withdraw their money in anticipation of low returns. There is thus a crisis.

Empirical evidence

What is the empirical evidence concerning whether runs are panic-based or fundamental-based? Friedman and Schwartz (1963) have written a comprehensive monetary history of the US from 1867-1960. Among other things, they argue that banking panics can have severe effects on the real economy. In the banking panics of the early 1930's, banking distress developed quickly and had a large effect on output. Friedman and Schwartz argued that the crises were panic-based and offered as evidence the absence of downturns in the relevant macroeconomic time series prior to the crises. Gorton (1988) showed that banking crises in the National Banking Era were predicted by a leading indicator based on liabilities of failed businesses. This evidence suggests banking crises are fundamental or business cycle related rather than panic-based. Calomiris and Gorton (1991) provide a wider range of evidence that crises are fundamental-based rather than panic-based. Wicker (1980, 1996) shows that, despite the absence of collapses in U.S. national macroeconomic time series, in the first two of the four crises identified by Friedman and Schwartz in the early 1930's there were large regional shocks and attributes the crises to these shocks. Calomiris and Mason (2003) undertake a detailed econometric study of the four crises using a broad range of data and conclude that the first three crises were fundamental-based while the fourth was panic-based.

Overall the evidence thus suggests that both types of banking crisis can occur in practice. However, the evidence for the U.S. in the nineteenth century and for the early 1930's suggests that fundamental-based crises are the most important type of crisis.

3. The Market Failure in Fundamental-based Models

Allen and Gale (2004a, 2007) develop a general equilibrium framework for understanding the normative aspects of crises. The model is a benchmark for investigating the welfare properties of financial systems. The interaction of banks and markets is considered. The markets are institutional markets in the sense that they are for banks and intermediaries to share risks and liquidity. Individuals cannot directly access these markets but instead invest their funds in banks that have access to the markets. Given the lack of a widely accepted theory of equilibrium selection they focus on fundamental shocks as the driver of financial crises – only *essential* crises are considered. In other words panics that are unnecessary in the sense that an equilibrium without a panic also exists are not considered.

Both financial intermediaries and markets play an important role in the model. Financial intermediaries provide liquidity insurance to consumers against idiosyncratic liquidity shocks. Markets allow financial intermediaries and their depositors to share aggregate liquidity and return shocks.

In understanding the market failures that can justify regulation a key role is played by complete versus incomplete markets and contracts. If financial markets are complete it is possible for intermediaries to hedge all aggregate risks in the financial markets. Complete markets involve state-contingent Arrow securities or their equivalent in terms of derivative securities or dynamic trading opportunities. In contrast incomplete markets mean that the amount of consumption in each possible aggregate state cannot be independently varied. If the contracts between intermediaries and consumers are complete then they can also be conditioned on aggregate risks. An incomplete contract

would be something like debt where the payoff on the contract does not depend on the aggregate state. Given these definitions Allen and Gale (2004a) show the following result.

Result 1: When markets are complete and contracts are complete the allocation of resources is *incentive* efficient.

The result provides the important benchmark of circumstances where Adam Smith's invisible hand works despite the presence of asymmetric information. As usual it involves comparing the allocation a decentralized market system with an allocation implemented by a central planner. The reason that the allocation is incentive efficient is that the idiosyncratic liquidity shocks to depositors cannot be directly observed by the intermediaries in the case of the market or the planner in the case of direct allocation. The depositors must have the correct incentives to reveal the information if this is necessary in the efficient allocation. Hence the notion of incentive efficiency rather than full efficiency is used.

In this ideal world of complete markets and complete contracts there is no market failure. Moreover financial crises do not occur because banks and other intermediaries can balance assets and liabilities state by state. In this case there is no need for regulation or government intervention of any kind. It is the analog to the first fundamental theorem of welfare economics in the context of financial intermediation.

So far we have assumed complete contracts between banks and other intermediaries and their customers. Many contracts observed in practice between

intermediaries and consumers such as debt and deposit contracts are incomplete. However, even if this is the case it is possible to show a result concerning efficiency.

Result 2: When contracts are incomplete and markets are complete the allocation is *constrained* efficient.

Again the invisible hand of the market works in the sense that a planner constrained to use incomplete contracts with consumers could not do any better then the market provided financial markets are complete. What is more it can be shown that in the equilibrium with incomplete contracts there can be financial crises. For example, if a bank uses a deposit contract then there can be a banking crisis. This demonstrates that crises are not everywhere and always bad. In some cases they can increase effective state contingencies and improve the possibilities for risk sharing and hence the allocation of resources. Of course, it is not the case that crises are always good, only that in some cases they can be, in particular when financial markets are complete and contracts between intermediaries and consumers are incomplete.

Once again there is no market failure and no justification for regulation or any other kind of intervention. This is another important benchmark. It shows that some crises can be good. Moreover the possibility of crisis does not always justify intervention. Having said that though there is of course another case to be considered and that is when financial markets are incomplete. This is what we turn to next. As we shall see there is indeed a market failure here. Now crises can be bad and regulations and other forms of intervention have the possibility of improving the allocation of resources.

The difference between complete and incomplete markets essentially determines whether the financial system is a shock absorber or an amplifier. With complete markets it is a shock absorber. The completeness allows risks to be born efficiently by everybody. With incomplete markets, however, shocks even very small ones, can be amplified and significant inefficiencies can result.

4. Incomplete Markets

The two results in the previous section show that if there are complete markets then there is no market failure. This is true whether contracts between banks and other intermediaries are complete or incomplete. Of course, welfare is usually higher with complete contracts than incomplete contracts but there is no market failure. With incomplete markets, it turns out there is indeed a market failure. This can take a number of different forms as we shall see. There can be financial fragility, contagion, and asset price bubbles.

The essential problem with incomplete markets is that liquidity provision is inefficient. The nature of risk management to ensure that the bank or intermediary has the correct amount of liquidity changes significantly from the case of complete markets. When markets are complete it is possible to use Arrow securities or equivalently a full set of derivatives or dynamic trading strategies to ensure liquidity is received when it is needed. The price system ensures adequate liquidity is provided in every state and is priced properly state by state. To understand how this works it is helpful to conceptualize complete markets in terms of Arrow securities that are traded at the initial date and pay off in a particular state. In this case banks and other intermediaries buy liquidity in states

where it is scarce by selling liquidity in states where it is plentiful for them. The complete markets allow risk sharing and insurance. The financial system acts as a shock absorber. If risk is increased it is spread around efficiently by the complete markets.

In contrast when markets are incomplete, liquidity provision is achieved by selling assets in the market when the liquidity is required. Asset prices are determined by the available liquidity or in other words by the "cash in the market". It is necessary that people hold liquidity and stand ready to buy assets when they are sold. These suppliers of liquidity are no longer compensated for the cost of providing liquidity state by state. Instead the cost must be made up on average across all states and this is where the problem lies.

The providers of liquidity have the alternative of investing in a productive long asset. There is an opportunity cost to holding liquidity since this has a lower return than the productive long asset. In order for people to be willing to supply liquidity they must be able to make a profit in some states. If nobody held liquidity then when banks and intermediaries sold assets to acquire liquidity their price would collapse to zero. This would provide an incentive for people to hold liquidity since they can acquire assets cheaply. In equilibrium prices will be bid up to the level where the profit in these states where banks and intermediaries sell is sufficient to compensate the providers of liquidity for all the other states where they do not use the liquidity and simply bear the opportunity cost of holding it. In other words, prices are low in the states where banks and intermediaries need liquidity. But this is exactly the wrong time from an efficiency point of view for there to be a transfer from the banks and intermediaries who need liquidity to the providers of liquidity. There is in effect negative insurance and suboptimal risk

sharing. Allen and Carletti (2006, 2007) explain in detail how this pricing mechanism works.

With incomplete markets the financial system thus acts as an amplifier. Large shocks can lead to more price volatility and this can cause significant problems in terms of bankruptcy and so forth.

To summarize when markets are incomplete asset prices must be volatile to provide incentives for liquidity provision. This asset price volatility can lead to costly and inefficient crises. There is a market failure that potentially provides the justification for regulation and other kinds of intervention to improve the allocation of resources.

5. The Symptoms of Market Failure

The problems in liquidity provision that result from incomplete markets can result in a number of phenomena that are associated with financial crises. These are financial fragility, contagion and asset price bubbles. Financial fragility is where a small shock can have a large effect and lead to a crisis. With contagion a shock in one region can spread to other regions and have a damaging effect. With asset price bubbles the inefficient provision of liquidity by the market can be exacerbated by the inefficient provision of liquidity by the central bank and this can result in deviations of asset prices from fundamentals. We consider each of these symptoms of market failure in turn.

Financial fragility

There are many historical illustrations of situations where small shocks have a significant impact on the financial system. For example, Kindleberger (1978, pp. 107-

108) argues that the immediate cause of a financial crisis:

> "...may be trivial, a bankruptcy, a suicide, a flight, a revelation, a refusal of credit to some borrower, some change of view which leads a significant actor to unload. Prices fall. Expectations are reversed. The movement picks up speed. To the extent that speculators are leveraged with borrowed money, the decline in prices leads to further calls on them for margin or cash, and to further liquidation. As prices fall further, bank loans turn sour, and one or more mercantile houses, banks, discount houses, or brokerages fail. The credit system itself appears shaky and the race for liquidity is on."

Recent examples provide a stark illustration of how small events can cause large problems. In August 1998 the Russian government announced a moratorium on about 281 billion roubles ($13.5 billion) of government debt. Despite the small scale of the default, it triggered a global crisis and caused extreme volatility in many financial markets. The hedge fund Long Term Capital Management (LTCM) came under extreme pressure. Despite LTCM's small size in relation to the global financial system, the Federal Reserve Bank of New York was sufficiently worried about the potential for a crisis if LTCM went bankrupt that it helped arrange for a group of private banks to purchase the hedge fund and liquidate its positions in an orderly way. The Fed's concern was that, if LTCM went bankrupt, it would be forced to liquidate all its assets quickly. LTCM held many large positions in fairly illiquid markets. In such circumstances, prices

might fall a long way if large amounts were sold quickly. This could put strain on other institutions, which would be forced to sell in turn, and this would further exacerbate the problem, as Kindleberger describes in the passage above.

Allen and Gale (2004b) show how the interaction of financial intermediaries and markets can lead to financial fragility. Small events, such as minor liquidity shocks, can have a large impact on the financial system because of the interaction of banks and markets. The role of liquidity is crucial. In order for financial intermediaries to have an incentive to provide liquidity to a market, asset prices must be volatile. Intermediaries that are initially similar may pursue radically different strategies, both with respect to the types of asset they invest in and their risk of default. The interaction of banks and markets provides an explanation for systemic or economy-wide crises, as distinct from models, such as Bryant (1980) and Diamond and Dybvig (1983) that explain individual bank runs.

As described in the previous section, the central idea is that when markets are incomplete financial institutions are forced to sell assets in order to obtain liquidity. Because the supply of and demand for liquidity are likely to be inelastic in the short run, a small degree of aggregate uncertainty can cause large fluctuations in asset prices. Holding liquidity involves an opportunity cost and the suppliers of liquidity can only recoup this cost by buying assets at fire sale prices in some states of the world; so, the private provision of liquidity by arbitrageurs will always be inadequate to ensure complete asset-price stability. As a result, small shocks can cause significant asset-price volatility. If the asset-price volatility is severe enough, banks may find it impossible to meet their fixed commitments and a full-blown crisis will occur.

Contagion

Financial contagion refers to the process by which a crisis that begins in one region or country or industry spreads to an economically linked region or country or another industry. There are a number of different reasons contagion can occur. For example, one basis for contagion is information (see, e.g., Kodres and Pritsker (2002), Calvo and Mendoza (2000a, 2000b) and Calvo (2002)). Here we will focus on a second type of contagion that is due to incompleteness that is laid out in Allen and Gale (2000a). Again the problem is concerned with liquidity provision but in a somewhat different way than that discussed in the context of financial fragility. The possibility of this kind of contagion arises from the overlapping claims that different regions or sectors of the banking system have on one another. When one region suffers a bank crisis, the other regions suffer a loss because their claims on the troubled region fall in value. If this spillover effect is strong enough, it can cause a crisis in the adjacent regions. In extreme cases, the crisis passes from region to region, eventually having an impact on a much larger area than the region in which the initial crisis occurred.

Suppose the economy consists of a number of regions. The number of early and late consumers in each region fluctuates randomly, but the aggregate demand for liquidity is constant. This allows for inter-regional insurance as regions with liquidity surpluses provide liquidity for regions with liquidity shortages. One way to organize the provision of insurance is through the exchange of interbank deposits. Suppose that region A has a large number of early consumers when region B has a low number of early consumers, and vice versa. Since regions A and B are otherwise identical, their deposits are perfect substitutes. The banks exchange deposits at the first date, before they observe the

liquidity shocks. If region A has a higher than average number of early consumers at date 1, then banks in region A can meet their obligations by liquidating some of their deposits in the banks of region B. Region B is happy to oblige, because it has an excess supply of liquidity, in the form of the short asset. At the final date, the process is reversed, as banks in region B liquidate the deposits they hold in region A to meet the above-average demand from late consumers in region B.

Inter-regional cross holdings of deposits work well as long as there is enough liquidity in the banking system as a whole. If there is an excess demand for liquidity, however, the financial linkages caused by these cross holdings can turn out to be a disaster. While cross holdings of deposits are useful for reallocating liquidity within the banking system, they cannot increase the total amount of liquidity. If the economy-wide demand from consumers is greater than the stock of the short asset, the only way to provide more consumption is to liquidate the long asset. In this case liquidation refers to technological or physical liquidation rather than selling the asset in a market. There is a limit to how much can be liquidated without provoking a run on the bank, however, so if the initial shock requires more than this buffer, there will be a run on the bank and the bank is forced into bankruptcy. Banks holding deposits in the defaulting bank will suffer a capital loss, which may make it impossible for them to meet their commitments to provide liquidity in their region. Thus, what began as a financial crisis in one region will spread by contagion to other regions because of the cross-holdings of deposits.

Whether the financial crisis does spread depends crucially on the pattern of inter-connectedness generated by the cross holdings of deposits. The interbank network is said to be complete if each region is connected to all the other regions and incomplete if each

region is connected with a small number of other regions. In a complete network, the amount of interbank deposits that any bank holds is spread evenly over a large number of banks. As a result, the initial impact of a financial crisis in one region may be attenuated. In an incomplete network, on the other hand, the initial impact of the financial crisis is concentrated in the small number of neighboring regions, with the result that they easily succumb to the crisis too. As each region is affected by the crisis, it prompts premature liquidation of long assets, with a consequent loss of value, so that previously unaffected regions find that they too are affected.

It is important to note the role of a free rider problem in explaining the process of contagion. Cross holdings of deposits are useful for redistributing liquidity, but they do not create liquidity. So when there is excess demand for liquidity in the economy as a whole each bank tries to meet external demands for liquidity by drawing down its deposits in another bank. In other words, each bank is trying to "pass the buck" to another bank. The result is that all the inter-bank deposits disappear and no one gets any additional liquidity.

The only solution to a global shortage of liquidity (withdrawals exceed short assets), is to physically liquidate long assets. Each bank has a limited buffer that it can access by physically liquidating the long asset. If this buffer is exceeded, the bank must fail. This is the key to understanding the difference between contagion in complete and incomplete networks. When the network is complete, banks in the troubled region have direct claims on banks in every other region. Every region takes a small hit (physically liquidates a small amount of the long asset) and there is no need for a global crisis. When the network is incomplete, banks in the troubled region have a direct claim only on the

banks in adjacent regions. The banks in other regions are not required to liquidate the long asset until they find themselves on the front line of the contagion. At that point, it is too late to save themselves.

There are a number other ways that contagion can occur. For example, Allen and Carletti (2006) analyze how financial innovation can create contagion across sectors and lower welfare relative to the autarky solution. They focus on the structure of liquidity shocks hitting the banking sector as the main mechanism generating contagion. Differently, Allen and Carletti (2007) focus on the impact of different accounting methods and show that mark-to-market accounting can lead to contagion in situations where historic cost based accounting values do not.

Bubbles

The idea that the amount of liquidity available is an important factor in the determination of asset prices has a long history. In addition to the liquidity provided by the market, the liquidity in the form of money and credit provided by the central bank also plays an important role. It is this aspect of liquidity provision that is the focus here. In his description of historic bubbles Kindleberger (1978; p. 54) emphasizes the role of this factor: “Speculative manias gather speed through expansion of money and credit or perhaps, in some cases, get started because of an initial expansion of money and credit.”

In many recent cases where asset prices have risen and then collapsed dramatically an expansion in credit following financial liberalization appears to have been an important factor. Perhaps the best known example of this type of phenomenon is the dramatic rise in real estate and stock prices that occurred in Japan in the late 1980's

and their subsequent collapse in 1990. The next few years were marked by defaults and retrenchment in the financial system. The real economy was adversely affected by the aftermath of the bubble and growth rates during the 1990's were typically slightly positive or negative, in contrast to most of the post war period when they were much higher.

This and other examples suggest a relationship between the occurrence of significant rises in asset prices or positive bubbles and the provision of liquidity. They also illustrate that the collapse in the bubble can lead to severe problems because the fall in asset prices leads to strains on the banking sector. Banks holding real estate and stocks with falling prices (or with loans to the owners of these assets) often come under severe pressure from withdrawals because their liabilities are fixed. This forces them to call in loans and liquidate their assets, which in turn appears to exacerbate the problem of falling asset prices. In other words there may be negative asset price bubbles as well as positive ones. These negative bubbles where asset prices fall too far can be very damaging to the banking system. This can make the problems in the real economy more severe than they need have been.

Despite the apparent empirical importance of the relationship between liquidity and asset price bubbles there is no widely agreed theory of what underlies these relationships. Allen and Gale (2000b) provide a theory of this based on the existence of an agency problem. Many investors in real estate and stock markets obtain their investment funds from external sources. If the ultimate providers of funds are unable to observe the characteristics of the investment, there is a classic risk shifting problem. Risk shifting increases the return to investment in risky assets and causes investors to bid up

prices above their fundamental values. A crucial determinant of asset prices is thus the amount of credit that is provided. Financial liberalization, by expanding the volume of credit and creating uncertainty about the future path of credit expansion, can interact with the agency problem and lead to a bubble in asset prices.

When the bubble bursts either because returns are low or because the central bank tightens credit, banks are put under severe strain. Many of their liabilities are fixed while their assets fall in value. Depositors and other claimants may decide to withdraw their funds in anticipation of problems to come. This will force banks to liquidate some of their assets and this may result in a further fall in asset bubbles because of a lack of liquidity in the market. It can be shown that when there is a market for risky assets then their price is determined by "cash-in-the-market pricing" in some states and can fall below their fundamental value. This leads to an inefficient allocation of resources. The central bank can eliminate this inefficiency by an appropriate injection of liquidity into the market.

6. Discussion

We have identified two market failures. The first concerns a coordination problem associated with panics. The problem in analyzing this from a policy perspective is that there is no widely accepted method for selecting equilibria. Global games are one promising approach but as yet there is little empirical evidence to support this methodology. The second market failure concerns the incompleteness of financial markets. The essential problem here is that the incentives to provide liquidity lead to an inefficient allocation of resources. We have discussed three manifestations of market

failure associated with liquidity provision. These are financial fragility, contagion, and asset price bubbles.

The framework we have developed allows some insight into the question of when the financial system acts a shock absorber and when it acts as an amplifier. When markets are complete and there is no market failure, it acts as a shock absorber. Risks are spread efficiently across economic agents. In this sense risks are absorbed. When there is a market failure the financial system can act as an amplifier. In the case of panics, there is an extreme amplification effect. Sunspots are shocks that by themselves have no effect. However, if they are used as coordination devices they can have an extreme effect on the equilibrium allocation and in that sense the financial system acts as an amplifier.

The second market failure of incomplete markets in fundamental-based models also acts as an amplifier. Financial fragility is another extreme example. Here small shocks can again lead to large changes in asset prices. This volatility in turn can lead to significant disruption and crises. With contagion, there is again amplification. A shock in one region can spill over to other regions and have a much larger effect than the original effect. Finally, asset price bubbles can also lead to large economic problems and in that sense are amplifiers.

Having identified when there is a market failure, the natural question that follows is whether there exist policies that can correct the undesirable effects of such failures. With the first market failure of panics, one of the main points that Diamond and Dybvig (1983) made was that deposit insurance was a way of eliminating the multiplicity of equilibria. In practice deposit insurance is not complete since typically only small depositors are covered. As a result actual deposit insurance schemes do not prevent the

possibility of panics. The analysis of deposit insurance as a way of eliminating crises is something that deserves more attention. It potentially provides an underpinning for why deposit insurance is needed and this in turn justifies the need for capital regulation. In standard analyses of capital regulation, the need for this is usually justified by the existence of deposit insurance but this is simply assumed. A full analysis requires the need for deposit insurance to be properly modeled.

In the context of the market failure due to incomplete markets in fundamental-based models, Allen and Gale (2004a, 2007) and Gale and Özgür (2005) consider two types of regulation. The first is the regulation of bank liquidity and the second is the regulation of bank capital. Allen and Gale (2004a) investigate bank liquidity regulation and show requiring banks to hold more liquidity than they would choose to is welfare improving if relative risk aversion is above 1. Gale and Özgür (2005) investigate simple examples with consumers who have constant relative risk aversion, when financial markets are incomplete. It is shown that the effect of bank capital regulation depends critically on the degree of relative risk aversion. When relative risk aversion is sufficiently low (below 2) increasing levels of bank capital above what banks would voluntarily hold can make everybody better off. The informational requirements for these kinds of intervention are high. Thus it may be difficult to improve welfare through these kinds of regulation as a practical matter.

Financial fragility, contagion, and asset price bubbles are also manifestations of market failures. The policies required for dealing with these are rather different. These issues have not been analyzed very much. However, it seems likely that they require provision of liquidity by the central bank to overcome them. The relationship between

monetary policy and the control of crises is not well understood. For the case of financial fragility the problem is the price volatility that arises from the private incentives for liquidity provision. By injecting monetary liquidity into the market the central bank may be able to change the price volatility and hence the financial fragility. With contagion, the problem is again a lack of liquidity. By injecting liquidity into the interbank market, the central bank may be able to prevent the spread of crises. Also asset price bubbles represent an important area where the central bank may be able to use monetary policy to solve the market failure.

The development of microeconomic banking models with monetary channels is at an early stage. Allen and Gale (1998, 2007) and Diamond and Rajan (2006), among others, have made steps in this direction. However, the role of monetary policy in solving these market failures and turning the financial system into a shock absorber rather than an amplifier represents an important topic for future research.

References

Allen, F. and D. Gale (1998). "Optimal Financial Crises." *Journal of Finance* 53, 1245-1284.

Allen, F. and D. Gale (2000a). "Financial Contagion," *Journal of Political Economy* 108, 1-33.

Allen, F. and D. Gale (2000b). "Bubbles and Crises," *Economic Journal* 110, 236-255.

Allen, F. and D. Gale (2004a). "Financial Intermediaries and Markets," *Econometrica* 72, 1023-1061.

Allen, F. and D. Gale (2004b). "Financial Fragility, Liquidity, and Asset Prices," *Journal of the European Economic Association* 2, 1015-1048.

Allen, F. and D. Gale (2007). *Understanding Financial Crises,* Clarendon Lecture Series in Finance, Oxford: Oxford University Press.

Allen, F and E. Carletti (2006). "Credit Risk Transfer and Contagion," *Journal of Monetary Economics* 53, 89-111.

Allen, F and E. Carletti (2007). "Mark-to-Market Accounting and Liquidity Pricing," *Journal of Accounting and Economics*, forthcoming.

Bannier, C. (2005). "The Role of Information Disclosure and Uncertainty in the 1994/95 Mexican Peso Crisis: Empirical Evidence," *Review of International Economics*, forthcoming.

Bordo, M., B. Eichengreen, D. Klingebiel and M. Martinez-Peria (2001). "Is the Crisis Problem Growing More Severe?" *Economic Policy*, April 2001, 53-82 + Web Appendix.

Boyd, J., S. Kwak, and B. Smith (2005). "The Real Output Losses Associated with Modern Banking Crises," *Journal of Money, Credit, and Banking* 37, 977-999.

Bryant, J. (1980). "A Model of Reserves, Bank Runs, and Deposit Insurance," *Journal of Banking and Finance* 4, 335-344.

Calomiris, C. and G. Gorton (1991). "The Origins of Banking Panics, Models, Facts, and Bank Regulation," in R. G. Hubbard, ed., *Financial Markets and Financial Crises*, Chicago, IL: University of Chicago Press.

Calomiris, C. and J. Mason, "Fundamentals, Panics, and Bank Distress during the Depression," *American Economic Review* 93, 1615-1647.

Calvo, G. (2002). "Contagion in Emerging Markets: When Wall Street is a Carrier," proceedings from the *International Economic Association Congress*, vol. 3, Buenos Aires, Argentina 2002. Also in G. Calvo (2005), *Emerging Capital Markets in Turmoil: Bad Luck or Bad Policy?* Cambridge, MA: MIT Press.

Calvo, G. and E. Mendoza (2000a). "A Rational Contagion and the Globalization of Securities Markets," *Journal of International Economics* 51, 79-113.

Calvo, G. and E. Mendoza (2000b). "Capital-Markets Crises and Economic Collapse in Emerging Markets: An Informational Frictions Approach," *American Economic Review* 90, 59-64.

Carlsson, H. and E. van Damme (1993). "Global Games and Equilibrium Selection," *Econometrica* 61, 989-1018.

Chari, V. and R. Jagannathan (1988). "Banking Panics, Information, and Rational Expectations Equilibrium," *Journal of Finance* 43, 749-60.

Chen, Q., I. Goldstein, and W. Jiang (2007). "Payoff Complementarities and Financial Fragility: Evidence from Mutual Fund Flows," working paper, University of Pennsylvania.

Diamond, D. and P. Dybvig (1983). "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy* 91, 401-419.

Diamond, D. and R. Rajan (2006). "Money in a Theory of Banking," *American Economic Review* 96, 30-53.

Friedman, M. and A. Schwartz (1963). *A Monetary History of the United States, 1867-1960*, Princeton, NJ: Princeton University Press.

Gale, D. and O. Özgür (2005). "Are Bank Capital Ratios Too High or Too Low: Risk Aversion, Incomplete Markets, and Optimal Capital Structures," *Journal of the European Economic Association* 3, 690-700.

Goldstein, I. and A. Pauzner, (2005). "Demand-Deposit Contracts and the Probability of Bank Runs," *Journal of Finance* 60, 1293-1327.

Gorton, G. (1988). "Banking Panics and Business Cycles," *Oxford Economic Papers* 40, 751-781.

Hoggarth, G., R. Reis, and V. Saporta (2002). "Costs of Banking System Instability: Some Empirical Evidence," *Journal of Banking and Finance* 26, 825-855.

Honohan, P., and L. Laeven (2005). *Systemic Financial Crises: Containment and Resolution*, Cambridge, UK: Cambridge University Press.

Kindleberger, C. (1978). *Manias, Panics, and Crashes: A History of Financial Crises*, New York: Basic Books.

Kodres L. and M. Pritsker (2002). “A Rational Expectations Model of Financial Contagion,” *Journal of Finance* 57, 768-99.

Mitchell, W. (1941). *Business Cycles and Their Causes*, Berkeley: University of California Press.

Morris, S. and H. Shin (1998). “Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks,” *American Economic Review* 88, 587-597.

Prati, A. and M. Sbracia (2002). “Currency Crises and Uncertainty about Fundamentals,” IMF Working Paper 02/3.

Rochet, J. and X. Vives, (2004). “Coordination Failures and the Lender of Last Resort: Was Bagehot Right after All?” *Journal of the European Economic Association* 2, 1116-1147.

Tillman, P. (2004). “Disparate Information and the Probability of Currency Crises: Empirical Evidence,” *Economics Letters* 84, 61-68.

Wicker, E. (1980). “A Reconsideration of the Causes of the Banking Panic of 1930,” *Journal of Economic History* 40, 571-583.

Wicker, E. (1996). *The Banking Panics of the Great Depression*, Cambridge: Cambridge University Press.