MODELLING FINANCIAL INSTABILITY

Franklin Allen*

Financial instability can have large adverse effects on an economy. One major cause of instability is asset price bubbles. This paper starts by considering how such bubbles can arise due to the expansion of money and credit. The ways in which subsequent financial instability occurs are then discussed. Banking crises can arise due to panics or as a result of the business cycle. Contagion and financial fragility can cause small disturbances to have large effects. Finally, policy issues are touched upon.

1. Introduction

The UK has experienced a large increase in property prices in the past decade. At the same time consumer price inflation has been moderate. What has caused property prices to increase so much in an environment of low inflation? Is the rise in property prices due to the availability of easy credit or a change in fundamentals? If the increase is due to easy credit rather than an improvement in underlying fundamentals, does this mean that property prices will crash at some point and lead to financial instability? The purpose of this paper is to investigate these ideas in the context of recent models of financial instability.

There are a number of recent examples of monetary expansion and lending booms and apparent bubbles in real estate and stocks. Often these lending booms are triggered by financial deregulation. The subsequent bursting of these bubbles has led to financial crises (see also Borio, 2005, this volume). The idea that the amount of money and credit available is an important factor in the determination of asset prices is not new. In his description of historic bubbles Kindleberger (1978, p. 54) emphasises 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."

Perhaps the best known recent example of this type of phenomenon is the dramatic rise in real estate and stock prices that occurred in Japan in the late 1980s and their subsequent collapse in 1990. Financial liberalisation

throughout the 1980s and the desire to support the United States dollar in the latter part of the decade led to an expansion in credit. During most of the 1980s asset prices rose steadily, eventually reaching very high levels. For example, the Nikkei 225 index was around 10,000 in 1985. On December 19, 1989 it reached a peak of 38,916. A new Governor of the Bank of Japan, less concerned with supporting the US dollar and more concerned with fighting inflation, tightened monetary policy and this led to a sharp increase in interest rates in early 1990 (Frankel, 1993; Tschoegl, 1993). The bubble burst. The Nikkei 225 fell sharply during the first part of the year and by October 1, 1990 it had sunk to 20,222. Real estate prices followed a similar pattern. 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 1990s and 2000s have mostly been slightly positive or negative, in contrast to most of the postwar period when they were much higher.

Similar events occurred in Norway, Finland and Sweden in the 1980s (Heiskanen, 1993; Drees and Pazarbasioglu, 1995). In Norway the ratio of bank loans to nominal GDP went from 40 per cent in 1984 to 68 per cent in 1988. Asset prices soared while investment and consumption also increased significantly. The collapse in oil prices helped burst the bubble and caused the most severe banking crisis and recession since the war. In Finland an expansionary budget in 1987 resulted in massive credit expansion. The ratio of bank loans to nominal GDP increased from 55 per

* Wharton School, University of Pennsylvania. e-mail: allenf@wharton.upenn.edu

cent in 1984 to 90 per cent in 1990. Housing prices rose by a total of 68 per cent in 1987 and 1988. In 1989 the central bank increased interest rates and imposed reserve requirements to moderate credit expansion. In 1990 and 1991 the economic situation was exacerbated by a fall in trade with the Soviet Union. Asset prices collapsed, banks had to be supported by the Government and GDP shrank by 7 per cent. In Sweden a steady credit expansion through the late 1980s led to a property boom. In the fall of 1990 credit was tightened and interest rates rose. In 1991 a number of banks had severe difficulties because of lending based on inflated asset values. The Government had to intervene and a severe recession followed.

Mexico provides a dramatic illustration of an emerging economy affected by this type of problem. In the early 1990s the banks were privatised and a financial liberalisation occurred. Perhaps most significantly, reserve requirements were eliminated. Mishkin (1997) documents how bank credit to private nonfinancial enterprises went from a level of around 10 per cent of GDP in the late 1980s to 40 per cent of GDP in 1994. The stock market rose significantly during the early 1990s. In 1994 the Colosio assassination and the uprising in Chiapas triggered the collapse of the bubble. The prices of stocks and other assets fell and banking and foreign exchange crises occurred. These were followed by a severe recession.

These bubbles in asset prices typically have three distinct phases. The first phase starts with financial liberalisation or a conscious decision by the central bank to increase lending or some other similar event. The resulting expansion in credit is accompanied by an increase in the prices for assets such as real estate and stocks. This rise in prices continues for some time, possibly several years, as the bubble inflates. During the second phase the bubble bursts and asset prices collapse, often in a short period of time such as a few days or months, but sometimes over a longer period. The third phase is characterised by the default of many firms and other agents that have borrowed to buy assets at inflated prices. Banking and/or foreign exchange crises may follow this wave of defaults. The difficulties associated with the defaults and banking and foreign exchange crises often cause problems in the real sector of the economy which can last for a number of years. There is a significant interaction between the financial system and growth.

How can events such as these be understood? Allen and Gale (2000a; 2003; 2004) provide a theory of bubbles and ensuing crises based on the existence of an agency problem. Standard theories of asset pricing assume that investors purchase assets with their own wealth. In most financial systems, this is not the whole story. Intermediation is important. Many of the agents buying real estate, stocks, and other assets do so with other people's money. The purchase of real estate is usually debt financed. If the investment is successful, the borrower repays the loan and retains the difference between the value of the asset and the principal and interest. If the investment is unsuccessful, the borrower has limited liability and the lender bears the shortfall. Similarly, a large proportion of stocks is held by mutual funds, pension funds, and insurance companies. Money managers also have incentives to take risk. If their investment strategy is successful, they may be rewarded by a share of the returns, but most importantly they will attract new investors in the future. Because they receive management fees in proportion to the assets under their control, they will be significantly better off as a result of their good performance. If the investment strategy is unsuccessful, there is a limit to the downside risk that the manager bears. In the worst case, she will be fired but in any case her liability is limited. Thus, when intermediaries make investment decisions, the incentive scheme they face has convex payoffs.

The agency problem of excessive risk taking associated with limited liability is crucial for the analysis. If the penalties for default on debt or the reputational loss from being fired from an intermediary are sufficiently high then there will not be an incentive to take risks. Hence the theory can be thought of as applying to cases where these factors are insufficient to prevent risk taking.

If there is an agency problem of the type described the people making the investment decisions will have an incentive to take on risky projects. The fact that lenders are unable to observe the characteristics of a project means the borrowers can shift risk to the lenders and increase their own payoff. This causes investors to bid up the prices of risky assets above their fundamental values and there is a bubble. The more risky the asset the greater is the amount that can be shifted and the larger the bubble. This risk can come from two sources. The first is asset return risk. The second is financial risk. This is the risk associated with future financial conditions such as the amount of credit that will be available. A framework for analysing these effects is developed next.

2. A framework

A simple example is developed to illustrate the model in Allen and Gale (2000a).

For ease of exposition the example is slightly different from the model presented there. Standard models of asset pricing assume people invest with their own money. The price of an asset in this benchmark case is identified as the 'fundamental'. A bubble is said to occur when the price of an asset rises above this benchmark (see Allen, Morris and Postlewaite, 1993, for a full discussion of the definition of fundamental and bubble).

If the people making investment decisions borrow money, they are only interested in the upper part of the distribution of payoffs of the risky asset because of the possibility of default. As a result there is a risk shifting problem and the price of the risky asset is bid up above the benchmark so there is a bubble.

In the example, the people who make investment decisions do so with borrowed money. If they default there is limited liability. Lenders cannot observe the riskiness of the projects invested in so there is an agency problem. For the case of real estate this representation of the agency problem is directly applicable. For the case of stocks there are margin limits that prevent people directly borrowing and investing in the asset. However, a more appropriate interpretation in this case is that it is institutional investors making the investment decisions. This group constitutes a large part of the market in many countries. The agency problem that occurs is similar to that with a debt contract. First, the people that supply the funds have little control over how they are invested. Second, the reward structure is similar to what happens with a debt contract. If the assets the fund managers invest in do well, the managers attract more funds in the future and receive higher payments as a result. If the assets do badly, there is a limit to the penalty that is imposed on the managers. The worse that can happen is that they are fired. This is analogous to limited liability (Allen and Gorton, 1993).

Initially there are two dates t = 1, 2. There are two assets in the example. The first is a safe asset in variable supply. For each 1 unit invested in this asset at date 1 the output is 1.5 at date 2. The second is a risky asset in fixed supply that can be thought of as real estate or stocks. There is 1 unit of

Table 1. The basic example

Asset	Supply	Investment at date 1	Payoff at date 2
Safe	Variable	1	1.5
Risky	1	P	R = 6 with prob. 0.25
			= 1 with prob. 0.75
			ER = 2.25

Note: All agents in the model are assumed to be risk neutral.

this risky asset. For each unit purchased at price P at date 1 the output is 6 with prob. 0.25 and 1 with prob. 0.75 at date 2 so the expected payoff is 2.25. The details of the two assets are given in table 1.

The fundamental

Suppose each investor has wealth 1 initially and invests her own wealth directly. Since everybody is risk neutral the marginal returns on the two assets must be equated.

$$\frac{2.25}{P_F} = \frac{1.5}{1}$$

or

$$P_F = \frac{2.25}{1.5} = 1.5$$

The value of the asset is simply the discounted present value of the payoff where the discount rate is the opportunity cost of the investor. This is the classic definition of the fundamental. The benchmark value of the asset is thus 1.5 and any price above this is termed a bubble.

Intermediated case

Suppose next that investors have no wealth of their own. They can borrow to buy assets at a rate of 33.33 per cent. The most they can borrow is 1. If they borrow 1 they repay 1.33 if they are able to. If they are unable to pay this much the lender can claim whatever they have. As explained above, lenders cannot observe how loans are invested and this leads to an agency problem.

The first issue is can $P = 1.5$ be the equilibrium price? Consider what happens if an investor borrows 1 and invests in the safe asset.

Marginal return to safe asset
= 1.5 – 1.33
= 0.17

Suppose instead that she borrows 1 and invests in the risky asset. She purchases 1/1.5 units. When the payoff is 6 she repays the principal and interest of 1.33 and keeps what remains. When it is 1 she defaults and the entire payoff goes to the lender so she receives 0.

Marginal return to risky asset
$$= 0.25\left(\frac{1}{1.5} \times 6 - 1.33\right) + 0.75 \times 0$$
$$= 0.67$$

The risky asset is clearly preferred when $P = 1.5$ since $0.67 > 0.17$. The total expected payoff of 1.5 on the investment in 1 unit of the safe asset is the same as on the investment of $1/1.5$ units of the risky asset since $(1/1.5) \times 2.25 = 1.5$. The risky asset is more attractive to the borrower though. With the safe asset the borrower obtains 0.17 and the lender obtains 1.33. With the risky asset the borrower obtains 0.67 while the lender obtains $0.25 \times 1.33 + 0.75 \times 1 \times (1/1.5) = 1.5 - 0.67 = 0.83$. The risk of default allows 0.5 in expected value to be shifted from the lender to the borrower. This is the *risk shifting* problem. If the lender could prevent the borrower from investing in the risky asset he would do so but he cannot since this is unobservable.

What is the equilibrium price of the risky asset given this agency problem? In an equilibrium where the safe asset is used, the price of the risky asset, P, will be bid up since it is in fixed supply, until the expected profit of borrowers is the same for both the risky and the safe asset:

$$0.25\left(\frac{1}{P} \times 6 - 1.33\right) + 0.75 \times 0 = 1.5 - 1.33$$

so

$$P = 3.$$

There is a bubble with the price of the risky asset above the benchmark of 1.5.

The idea that there is a risk shifting problem when the lender is unable to observe how the borrower invests the funds is not new (see, for example, Jensen and Meckling, 1976; Stiglitz and Weiss, 1981). However, it has not been widely applied in the asset pricing literature. Instead of the standard result in corporate finance textbooks that debt-financed firms are willing to accept negative net present value investments, the manifestation of the agency problem here is that the debt-financed investors are willing to invest in assets priced above their fundamental.

The amount of risk that is shifted depends on how risky the asset is. The greater the risk the greater the potential to shift risk and hence the higher the price will be. To illustrate this consider the previous example but suppose the payoff on

Table 2. A mean-preserving spread of the basic example

Asset	Supply	Investment at date 1	Payoff at date 2
Safe	Variable	1	1.5
Risky	1	P	R = 9 with prob. 0.25
			= 0 with prob. 0.75
			ER = 2.25

the risky asset is a mean-preserving spread of the original payoffs (table 2).

Now the price of the risky asset is given by

$$0.25\left(\frac{1}{P} \times 9 - 1.33\right) + 0.75 \times 0 = 1.5 - 1.33$$

so

$$P = 4.5.$$

More risk is shifted and as a result the price of the risky asset is bid up to an even higher level.

It is interesting to note that in both the stock market boom of the 1920s and the one in the 1990s the stocks that did best were 'high-tech' stocks. In the 1920s it was radio stocks and utilities that were the star performers (see White, 1990). In the 1990s it was telecommunications, media and entertainment and technology stocks that did the best. It is precisely these stocks which have the most uncertain payoffs because of the nature of the business they are in.

One of the crucial issues is why the banks are willing to lend to the investors given the chance of default. To see this consider again the case where the payoffs on the risky asset are those in table 1 and $P = 3$. In this case the quantity of the risky asset purchased when somebody borrows 1 is $1/P = 1/3$. In the equilibria considered above the investors are indifferent between investing in the safe and risky asset. Suppose for the sake of illustration the fixed supply of the risky asset is $4/3$. The amount of funds depositors have is 10 and the number of borrowers is 10. In the equilibrium where $P = 3$, 4 of the borrowers invest in the risky asset and 6 in the safe in order for the fixed supply of $4/3$ units of the risky asset to be taken up. In this case 40 per cent of borrowers invest in risky assets and 60 per cent invest in safe assets. A bank's expected payoff from lending one unit is then given by the following expression.

$$\text{Bank's expected payoff}$$
$$= 0.4[0.25 \times 1.33 + 0.75 \times (1/3) \times 1] + 0.6[1.33]$$
$$= 1.03$$

The first term is the payoff to the bank from the 40 per cent of investors in the risky asset. If the payoff is 6, which occurs with probability 0.25, the loan and interest is repaid in full. If the payoff is 1, which occurs with probability 0.75, the borrower defaults and the bank receives the entire proceeds from the $1/3$ unit owned by the borrower. The bank's payoff is thus $(1/3) \times 1$. The 60 per cent of investors in the safe asset are able to pay off their loan and interest of 1.33 in full.

If the banking sector is competitive, the receipts from lending, 1.03, will be paid out to depositors. In this case it is the depositors that bear the cost of the agency problem. In order for this allocation to be feasible, markets must be segmented. The depositors and the banks must not have access to the assets that the investors who borrow invest in. Clearly, if they did, they would be better off just to invest in the safe asset rather than put their money in the bank or lend to the investors. The segmented markets assumption is designed to capture the fact that an investor like the investment bank Goldman Sachs can year after year earn a much higher return on its equity than the return an average individual can make. Even if the individual is willing to take the same risk, she cannot earn the same expected return as Goldman Sachs because they have much better access to some markets.

Credit and interest rate determination

The quantity of credit and the interest rate have so far been taken as exogenous. These factors are incorporated in the example next to illustrate the relationship between the amount of credit and the level of interest rates. The simplest case to consider is where the central bank determines the aggregate amount of credit B available to banks. It does this by setting reserve requirements and determining the amount of assets available for use as reserves. For ease of exposition it is simply assumed the central bank sets B. The banking sector is competitive. The number of banks is normalised to 1 and the number of investors is also normalised to 1. Each investor will therefore be able to borrow B from each bank.

The safe asset pays a fixed return r to the investor: if x is invested in the safe asset at date 1 the return is rx at date 2. The safe asset can be interpreted in a number of ways. One possibility is that it is debt issued by the corporate sector. The investors treat the rate of return as fixed because they are small relative to the size of the corporate sector. In equilibrium, competition will ensure that the rate of return on the bonds is equal to the marginal product of capital.

The return on the safe asset is determined by the marginal product of capital in the economy. This in turn depends on the amount of the consumption good x that is invested at date 1 in the economy's productive technology to produce $f(x)$ units at date 2. The total amount that can be invested is B and the amount that is invested at date 1 in the risky asset since there is 1 unit is P. Hence the date 1 budget constraint implies that

$$x = B - P.$$

It is assumed

$$f(x) = 3x^{0.5} = 3(B-P)^{0.5}. \tag{1}$$

Provided the market for loans is competitive, the interest rate r will be bid up by investors until it is equal to the marginal product of investment,

$$r = f'(B-P) = 1.5(B-P)^{-0.5}. \tag{2}$$

At this level the safe asset will not yield any profits for investors. If it were lower than this there would be an infinite demand for the safe asset and if it were higher than this there would be zero demand.

The amount the investors will be prepared to pay for the risky asset assuming its payoffs are as in table 1 is then given by

$$0.25\left(\frac{1}{P}\times 6 - r\right) + 0.75\times 0 = 0.$$

Using (2) in this,

$$P = 4(B-P)^{0.5}.$$

Solving for P gives

$$P = 8(-1 + \sqrt{1 + 0.25B}). \tag{3}$$

When $B = 5$ then $P = 4$ and $r = 1.5$. The relationship between P and B is shown by the solid line in chart 1. By controlling the amount of credit, the central bank controls the level of interest rates and the level of asset prices. Note that this relationship is different from that in the standard

Chart 1. Credit and asset prices
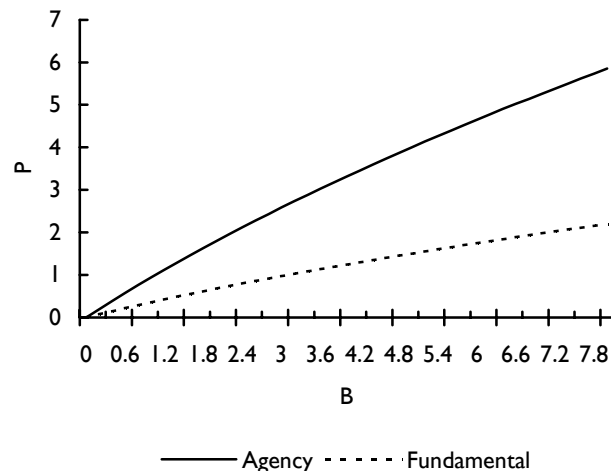
asset pricing model when the price of the risky asset is the discounted expected payoff,

$$P_F = \frac{2.25}{r}.$$

This case is illustrated by the dotted line in chart 1. A comparison of the two cases shows that the fundamental is relatively insensitive to the amount of credit compared to the case where there is an agency problem. Changes in aggregate credit can cause relatively large changes in asset prices when there is an agency problem.

Financial risk

The previous section assumed that the central bank could determine the amount of credit B. In practice, the central bank has limited ability to control the amount of credit and this means B is random. In addition there may be changes of policy preferences, changes of administration, and changes in the external environment which create further uncertainty about the level of B. This uncertainty is particularly great in countries undergoing financial liberalisation. In order to investigate the effect of this uncertainty an extra period is added to the model. Between dates 1 and 2 everything is the same as before. Between dates 0 and 1 the only uncertainty that is resolved is about the level of B at date 1. Thus between dates 0 and 1 there is financial uncertainty. The uncertainty about aggregate credit B at date 1 causes uncertainty about prices at date 1. Given that investors are borrowing from banks at date 0 in the same way as before this price uncertainty again leads to an agency problem and risk shifting. The price of the risky asset at date 0 will reflect this price uncertainty and can lead the asset price to be even higher than at date 1.

Suppose that there is a 0.5 probability that $B = 5$ and a 0.5 probability that $B = 7$ at date 1. Then using (2) and (3) the prices and interest rates are as shown in table 3.

The pricing equation at date 0 is

$$0.5\left(\frac{1}{P_0} \times 5.27 - r_0\right) + 0.5 \times 0 = 0,$$

Table 3. The basic example extended to include financial risk

Probability	B	P	r
0.5	5	4	1.5
0.5	7	5.27	1.14

Table 4. A mean-preserving spread of the example extended with financial risk

Probability	B	P	r
0.5	4	3.14	1.81
0.5	8	5.86	1.03

where r_0, the date 0 interest rate, is given by (2) with B and P replaced by B_0 and P_0. Substituting for r_0 and simplifying

$$P_0 = \frac{5.27}{1.5(B_0 - P_0)^{-0.5}}.$$

Taking $B_0 = 6$ and solving for r_0 and P_0 gives

$$r_0 = 1.19$$

$$P_0 = 4.42.$$

As when the uncertainty is due to variations in asset payoffs, the greater the financial uncertainty the greater is P_0. Consider a mean preserving spread on the financial uncertainty so that table 3 is replaced by table 4.

In this case it can be shown

$$r_0 = 1.27$$

$$P_0 = 4.61.$$

The risk shifting effect operates for financial risk in the same way as it does for asset payoff risk. Although the expected payoff at date 2 is only 2.25, the price of the risky asset at date 1 in the last case is 4.61. The possibility of credit expansion over a period of years may create a great deal of uncertainty about how high the bubble may go and when it may collapse. This is particularly true when economies are undergoing financial liberalisation. As more periods are added it is possible for the bubble to become very large. The market price can be much greater than the fundamental.

These examples illustrate that what is important in determining the risky asset's price at date 0 is expectations about aggregate credit at date 1. If aggregate credit goes up, then asset prices will be high and default will be avoided. However, if aggregate credit goes down, then asset prices will be low and default will occur. The issue here is what is the dynamic path of aggregate credit? The point is that the expectation of credit expansion is already taken into account in the investors' decisions about how

much to borrow and how much to pay for the risky asset. If credit expansion is less than expected, or perhaps simply falls short of the highest anticipated levels, the investors may not be able to repay their loans and default occurs. In Allen and Gale (2000a) it is shown that even if credit is always expanded then there may still be default. In fact it is shown that there are situations where the amount of credit will be arbitrarily close to the upper bound of what is anticipated and widespread default is almost inevitable.

Discussion

A simple example has been developed where a borrower chooses the type of investments (safe or risky) and the lender is unable to observe how the funds are invested. As in Jensen and Meckling (1976) and Stiglitz and Weiss (1981), these assumptions imply there is a risk shifting problem. By buying risky assets, the borrower can shift downside risk onto the lender, but retains the right to any upside returns. The more risky the asset, the more attractive it becomes. When a significant proportion of investors in the market have these incentives, the equilibrium asset price will be high relative to the 'fundamental' value of the asset, which is defined as the price that would obtain in the standard asset pricing model where everybody is investing his own wealth. The difference between the equilibrium price and the fundamental value is the 'bubble'. Two factors are particularly important in determining the size of the bubble. One is the amount of credit that is available to finance speculative investment. The other is the degree of uncertainty in the market. The greater is either of these factors, the greater is the bubble.

The relationship between credit and asset prices is relatively straightforward in real estate markets. An expansion of credit reduces the interest rate at which investors can borrow and this in turn increases the prices they are willing to pay. In stock markets, the relationship is more subtle. Margin restrictions imply that only a proportion of the total investment can be financed with borrowed funds. However, if credit expands, investors may be willing to borrow a greater amount against the houses, cars, and other assets they buy and put more money into intermediaries such as mutual funds. As explained above, the incentives that money managers face are similar to those that would be created if the money were directly borrowed and, again, asset prices will be bid up as a result.

The relationship between credit and asset prices becomes even more complex in a dynamic context. In deciding how much he should pay for an asset today, an investor will consider the future price of the asset and the possibility of capital gains. The future price will depend in part on the level of credit that is anticipated in future periods. If an expansion of credit is anticipated, asset prices are likely to rise and this expectation will feed back into current prices. Thus, it is not only current credit expansion but anticipated future expansion that feeds the bubble in asset prices.

There is another aspect of future credit expansion that has a direct impact on current asset prices. It is unlikely that the future level of credit can be perfectly anticipated. There may in fact be a great deal of uncertainty about future credit expansion. This arises from factors such as the central bank's limited ability to control the amount of credit, changes of policy preferences, changes of administration, and changes in the external environment, all of which may alter the amount of credit that will be created. The more uncertainty that is associated with future credit, the more uncertain future asset prices will be. Because of the risk shifting problem, uncertainty makes assets more attractive to the debt financed investor, and this results in a higher asset price and a larger bubble.

The theory thus predicts that bubbles will tend to occur when the current credit levels are high, when future credit levels are expected to be higher, and when future credit levels are expected to be uncertain. This is consistent with the fact that many asset bubbles associated with recent crises were preceded by financial liberalisation. In the Scandinavian countries, there was a move away from restricted financial systems towards market oriented ones. This led to an immediate expansion in credit and also considerable uncertainty about the future level of credit. In Japan, the Government continually eased regulation on banks and the financial markets throughout the 1980s. Similar deregulation occurred in many emerging economies, such as Mexico and the South East Asian economies.

The second phase of the financial crisis involves the bursting of the bubble and a collapse in asset prices. In some of the episodes recounted in the Introduction, it appears that the collapse was precipitated by a real shock. An example is the collapse in oil prices that triggered the bursting of the bubble in Norway. In other cases, the crisis appears to have been triggered by an event in the financial sector. A good example is Japan's tightening of credit in 1990, which precipitated the collapse in asset prices.

The effect of a real shock is easy to understand. Anything that affects the health of the businesses that make up the

economy will clearly have a direct impact on asset prices. Furthermore, uncertainty about these factors will lead to uncertainty about stock prices. The effect of a financial shock is more complex.

The model in Allen and Gale (2000a) suggests that a critical determinant of asset prices is the expected value and the volatility of credit expansion. In many cases financial liberalisation leads to an expansion of credit which feeds a bubble in asset prices. These higher prices are in turn supported by the anticipation of further increases in credit and asset prices. Any faltering of this cumulative process may cause the bubble to burst and lead to a financial crisis. What is critical is the relationship between actual and expected credit expansion. Since anticipated expansion has been built into current asset prices, continued expansion is required to allow speculators to repay their debts. In fact, a positive level of credit expansion may be required to prevent the bubble from bursting. Allen and Gale (2000a) call a credit regime robust if there is no financial crisis as long as the level of credit does not contract. A fragile regime is one in which credit is actually required to expand at a positive rate in order to prevent a financial crisis. It is fairly easy to construct examples of fragile regimes. In fact, examples can be constructed where an arbitrarily high rate of credit expansion is necessary to prevent a crisis. In this case, the probability of a crisis is close to one.

The third phase of the crisis occurs after asset prices have collapsed. At this stage there will be widespread default and the banking system will come under severe strain. If the fall in asset prices is not too large, the banking system may be able to survive intact. However, in more extreme cases either many banks will fail and be liquidated or the Government will be forced to step in and rescue the banks. For small countries there may also be a currency crisis as the Government is forced to choose between lowering interest rates to save the banking system or raising them to protect the exchange rate. Even if rates are raised there may still be an exodus of capital. A moderate increase in interest rates may not be sufficient to prevent capital flight because of the weakened state of the banking system and the uncertainty that often accompanies financial turbulence.

Perhaps the most important aspect of the third phase is the spillover of the financial crisis into the real economy. In practice, financial crises are often associated with a significant fall in output or at least a reduction in the rate of growth. Output fell dramatically in the South East Asian economies that were subject to crises. This was also the case in the Scandinavian countries. However, the

Scandinavian countries quickly rebounded. In Japan, although the initial effect of the 1990 crash was relatively mild, growth has been depressed for a long period of time and the situation has continued to deteriorate.

3. Theories of financial instability

Banking crises

There are two traditional views of banking crises. One is that they are random events, unrelated to changes in the real economy. The classical form of this view suggests that crises are panics and are the result of 'mob psychology' or 'mass hysteria' (see, for example, Kindleberger, 1978). The modern version, developed by Diamond and Dybvig (1983) and others, 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 everyone believes that a banking panic is about to occur, it is optimal for each individual to try to withdraw her funds. Since each bank has insufficient liquid assets to meet all of its commitments, it will have to liquidate some of its assets at a loss. Given first-come, first-served, those depositors who withdraw initially will receive more than those who wait. On the one hand, anticipating this, all depositors have an incentive to withdraw immediately. On the other hand, if no one believes a banking panic is about to occur, only those with immediate needs for liquidity will withdraw their funds. Assuming that banks have sufficient liquid assets to meet these legitimate demands, there will be no panic. 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.

When there are multiple equilibria sunspots provide one way to select the equilibria that will occur. However, this approach is not very satisfactory since the equilibrium that is selected is essentially arbitrary. In an important paper, Morris and Shin (1998) show how the global games approach can be used to select an equilibrium in the context of currency crises. The approach relies on there being a lack of common knowledge about fundamentals. Goldstein and Pausner (2005) and Rochet and Vives (2004) have shown how a similar approach can be used to select a unique equilibrium in the context of banking crises.

An alternative to the view that banking crises are panics is that they 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, panics are not random events but a response to unfolding economic circumstances.

A number of authors have developed models of banking panics caused by aggregate risk. Wallace (1988; 1990), Chari (1989) and Champ, Smith, and Williamson (1996) extend Diamond and Dybvig (1983) by assuming the fraction of the population requiring liquidity is random. Chari and Jagannathan (1988), Jacklin and Bhattacharya (1988), Hellwig (1994), and Alonso (1996) introduce aggregate uncertainty which can be interpreted as business cycle risk. 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 unfavourable signal was received by this group or whether liquidity needs happen to be high. Chari and Jagannathan are able to show panics occur not only when the outlook is poor but also when liquidity needs turn out to be high. Jacklin and Bhattacharya (1988) also consider a model where some depositors receive an interim signal about risk. They show that the optimality of bank deposits compared to equities depends on the characteristics of the risky investment.

Building on the empirical work of Gorton (1988) and Calomiris and Gorton (1991) 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 panics. In their model, crises can improve risk sharing but they also involve deadweight costs if they cause projects to be prematurely liquidated. A central bank can avoid these deadweight costs and implement an optimal allocation of resources through an appropriate monetary policy. By creating fiat money and lending it to banks, the central bank can prevent the inefficient liquidation of investments while at the same time allowing optimal sharing of risks.

Contagion and financial fragility

The prevalence of financial crises has led many to conclude that the financial sector is unusually susceptible to shocks. One theory is that small shocks can have a large impact. A shock that initially affects only a particular region or sector, or perhaps even a few institutions, can spread by contagion to the rest of the financial sector and then infect the larger economy. There are a number of different types of contagion that have been suggested in the literature. The first is contagion through interlinkages between banks and financial institutions. The second is contagion through asset prices in financial markets. De Bandt and Hartmann (2002) contains a survey of this literature.

Banks are linked in several ways including payments systems and interbank markets. These linkages can lead to a problem of contagion. We start by considering models of payment system contagion. Building on a locational model of payment systems developed by McAndrews and Roberds (1995), Freixas and Parigi (1998) have considered contagion in net and gross payment systems. In a net payment system banks extend credit to each other within the day and at the end of the day settle their net position. This exposes banks to the possibility of contagion if the failure of one institution triggers a chain reaction. In a gross system transactions are settled on a one-to-one basis with central bank money. There is no risk of contagion but banks have to hold large reserve balances. A net payment system is preferred when the probability of banks having low returns is small, the opportunity cost of holding central bank money reserves is high, and the proportion of consumers that have to consume at another location is high. Freixas, Parigi and Rochet (2000) use this model to examine the conditions under which gridlock occurs. They show that there can be gridlock when the depositors in one bank withdraw their funds, anticipating that other banks cannot meet their netting obligations if all their depositors have also withdrawn their funds. Rochet and Tirole (1996a) consider the role of the too-big-to-fail policy in preventing contagion.

Allen and Gale (2000b) focus on a channel of contagion that arises from the overlapping claims that different regions or sectors of the banking system have on one another through interbank markets. When one region suffers a banking 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 and becomes a contagion. Lagunoff and Schreft (2001) study the spread of crises in a probabilistic model. Financial linkages are modeled by assuming that each project requires two participants and each participant requires two projects. When the probability that one's partner will withdraw becomes too large, all participants simultaneously withdraw and this is interpreted as a financial crisis. Van Rijckeghem and Weder (2000) document linkages through banking centers empirically. Rochet and Tirole (1996b) use monitoring as a

means of triggering correlated crises; if one bank fails, it is assumed that other banks have not been properly monitored and a general collapse occurs.

There are a number of papers that consider contagion through financial markets. King and Wadhwani (1990) considered a situation where information is correlated between markets. Price changes in one market are perceived to have implications for asset values in other markets. Kodres and Pritsker (2000) use a multi-asset rational expectations model to show how macroeconomic risk factors and country-specific asymmetric information can combine to produce contagion. Kyle and Xiong (2001) present a model of contagion in financial markets due to the existence of a wealth effect.

The notion of financial fragility is closely related to that of contagion. When a financial system is fragile a small shock can have a big effect. The shock may be spread by contagion. A financial crisis may rage out of control and bring down the entire economic edifice. The memory of the Great Depression and earlier crises is still with us and it powerfully reinforces belief in financial fragility. Financial multipliers are modelled by Kiyotaki and Moore (1997). In their model, the impact of illiquidity at one link in the credit chain travels down the chain and has a big impact. Chari and Kehoe (2000) show that herding behavior can cause a small information shock to have a large effect on capital flows.

4. Policy issues

The theories of crises outlined above raise at least two important public policy issues. The first is how bubbles in asset prices can be prevented. The second is how to deal with the banking system and minimise the loss of output after an asset bubble has occurred and precipitated a banking crisis. Each of these is discussed in turn.

Although it has long been recognised that there is a link between monetary policy, inflation and asset prices, there has only recently been an active debate concerning the extent to which central banks should target asset prices. The standard analysis of the link between stock prices and inflation suggests that when the money supply is increased, prices and wages will in the long run increase in line with the standard quantity theory of money. Depending on the relative speeds of adjustment of prices in the output and input markets, profits and hence stock prices can be increased or decreased by inflation. The empirical evidence suggests that a rise in inflation (realised, expected or unexpected) reduces stock prices. This type of theory does not provide much guidance to central banks in how to

target asset prices beyond suggesting that if inflation is controlled asset prices will be determined by fundamentals.

The theory outlined in Section 2 provides a rather different perspective on the relationship between monetary policy and asset prices. The theory emphasises the importance of the level and volatility of credit for asset price determination and thus provides an important role for monetary policy and the reserve requirements of banks in preventing the development of bubbles in asset prices. Governments and central banks should try to avoid unnecessary expansion of credit as well as unnecessary uncertainty about the path of credit expansion. This suggests that financial liberalisation is a particularly risky exercise, as experience confirms (see also empirical work surveyed in Demirgüç-Kunt and Detragiache, 2005, this volume). In a liberalisation regime, credit tends to increase dramatically and, because there is no experience with the new regime, uncertainty also increases significantly. If financial liberalisation is to be undertaken, it should be done slowly and carefully. To the extent possible, the central bank should make clear how the volume of credit will evolve over time.

The second policy issue concerns how the Government should intervene to deal with problems caused by a banking crisis and minimise the spillovers into the real economy. The collapse of a bubble can cause a significant debt overhang. The value of the option to continue together with the difficulty of liquidating loans for their fair value means that banks will try to remain in business as long as possible. In order to maintain levels of capital consistent with regulation, banks will reduce the volume of new loans and this will lead to a credit crunch. Goodhart (2005, this volume) discusses how the new Basel 2 rules may aggravate this problem. The reduction in output and the further negative impact this will have on the creditworthiness of other borrowers can lead to a significant reduction in output. To offset these negative effects, the Government can try to recapitalise the banking system. This can involve direct infusions of funds or outright nationalisation of the banking system. A comparison of Scandinavia and Japan provides an interesting contrast between swift intervention in Scandinavia and the non-interventionist attitude of the Japanese Government. In the Scandinavian countries, swift intervention and recapitalisation of the banking systems lead to quick recoveries. In contrast in Japan, the hesitation to intervene has resulted in an economy that has stagnated for well over a decade.

Designing policies to prevent the adverse consequences of financial instability is one of the most important tasks

facing policymakers. Understanding the causes of financial instability is an important prerequisite to the design of such policies. This remains a crucial and ongoing area of research.

REFERENCES

Allen, F. and Gale, D. (1998), 'Optimal financial crises'. *Journal of Finance*, 53, pp. 1245–84.

—(2000a), 'Bubbles and crises', *Economic Journal*, 110, pp. 236–55.

—(2000b), 'Financial contagion', *Journal of Political Economy*, 108, pp. 1–33.

—(2003), 'Asset price bubbles and stock market interlinkages', in Hunter, W., Kaufman, G. and Pomerleano, M. (eds), *Asset Price Bubbles*, MIT Press, pp. 323–36.

—(2004), 'Asset price bubbles and monetary policy', in Desai, M. and Said, Y. (eds), *Global Governance and Financial Crises*, London, Routledge, pp. 19–42.

Allen, F. and Gorton, G. (1993), 'Churning bubbles', *Review of Economic Studies*, 60, pp. 813–36.

Allen, F., Morris, S. and Postlewaite, A., (1993), 'Finite bubbles with short sale constraints and asymmetric information', *Journal of Economic Theory*, 61, pp. 206–29.

Alonso, I. (1996), 'On avoiding bank runs', *Journal of Monetary Economics*, 37, pp. 73–87.

Calomiris, C. and Gorton, G. (1991), 'The origins of banking panics, models, facts, and bank regulation', in Hubbard, G. (ed.), *Financial Markets and Financial Crises*, Chicago, IL, University of Chicago Press, pp. 109–173.

Champ, B., Smith, B. and Williamson, S. (1996), 'Currency elasticity and banking panics: theory and evidence', *Canadian Journal of Economics*, 29, pp. 828–64.

Chari, V. (1989), 'Banking without deposit insurance or bank panics: lessons from a model of the U.S. national banking system', *Federal Reserve Bank of Minneapolis Quarterly Review*, 13(Summer), pp. 3–19.

Chari, V. and Jagannathan, R. (1988), 'Banking panics, information, and rational expectations equilibrium', *Journal of Finance*, 43, pp. 749–60.

Chari, V. and Kehoe, P. (2000), 'Financial crises as herds', working paper, Federal Reserve Bank of Minneapolis.

De Bandt, O. and Hartmann, P. (2000), 'Systemic risk: a survey', in Goodhart, C. and Illing, G. (eds), *Financial Crises, Contagion, and the Lender of Last Resort*, Oxford, Oxford University Press, pp. 249–97.

Diamond, D. and Dybvig, P. (1983), 'Bank runs, deposit insurance, and liquidity', *Journal of Political Economy*, 91, pp. 401–19.

Drees, B. and Pazarbasioglu, C. (1995), 'The nordic banking crises: pitfalls in financial liberalization?', Working Paper 95/61, International Monetary Fund, Washington, DC.

Frankel, J. (1993), 'The Japanese financial system and the cost of capital', in Takagi, S. (ed.), *Japanese Capital Markets: New Developments in Regulations and Institutions*, Oxford, Blackwell, pp. 21–77.

Freixas, X. and Parigi, B. (1998), 'Contagion and efficiency in gross and net interbank payment systems', *Journal of Financial Intermediation*, 7, pp. 3–31.

Freixas, X., Parigi, B. and Rochet, J. (2000), 'Systemic risk, interbank relations and liquidity provision by the central bank', *Journal of Money, Credit and Banking*, 32, pp. 611–38.

Goldstein, I., and Pauzner, A. (2005), 'Demand deposit contracts and the probability of bank runs', *Journal of Finance* (*forthcoming*).

Gorton, G. (1988), 'Banking panics and business cycles', *Oxford Economic Papers*, 40, pp. 751–81.

Heiskanen, R. (1993), 'The banking crisis in the Nordic countries', *Kansallis Economic Review*, 2, pp. 13–19.

Hellwig, M. (1994), 'Liquidity provision, banking, and the allocation of interest rate risk', *European Economic Review*, 38, pp. 1363–89.

Jacklin, C. and Bhattacharya, S. (1988), 'Distinguishing panics and information-based bank runs: welfare and policy implications', *Journal of Political Economy*, 96, pp. 568–92.

Jensen, M. and Meckling, W. (1976), 'Theory of the firm: managerial behavior, agency cost and ownership structure', *Journal of Financial Economics*, 3, pp. 305–60.

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

King, M. and Wadhwani, S. (1990), 'Transmission of volatility between stock markets', *Review of Financial Studies*, 3, pp. 5–33.

Kiyotaki, N. and Moore, J. (1997), 'Credit chains', *Journal of Political Economy*, 105, pp. 211–48.

Kodres L. and Pritsker, M. (2002), 'A rational expectations model of financial contagion', *Journal of Finance*, 57, pp. 768–99.

Kyle, A. and Xiong, W. (2001), 'Contagion as a wealth effect', *Journal of Finance*, 56, pp. 1401–40.

Lagunoff, R. and Schreft, S. (2001), 'A model of financial fragility', *Journal of Economic Theory*, 99, pp. 220–64.

McAndrews, J. and Roberds, W. (1995), 'Banks, payments and coordination', *Journal of Financial Intermediation*, 4, pp. 305–27.

Mishkin, F. (1997), 'Understanding financial crises: a developing country perspective', *Annual World Bank Conference on Development Economics 1996*, Washington, DC, The International Bank for Reconstruction and Development, pp. 29–61.

Morris, S. and Shin, H. (1998), 'Unique equilibrium in a model of self-fulfilling currency attacks', *American Economic Review*, 88, pp. 587–97.

Rochet, J. and Tirole, J. (1996a), 'Interbank lending and systemic risk', *Journal of Money, Credit and Banking*, 28, pp. 733–62.

—(1996b), 'Controlling risk in payment systems', *Journal of Money, Credit and Banking*, 28, pp. 832–62.

Rochet, J. and Vives, X. (2004), 'Coordination failures and the lender of last resort: was Bagehot right after all? ', *Journal of the European Economic Association*, 2, pp. 1116–47.

Stiglitz, J. and Weiss, A. (1981), 'Credit rationing in markets with imperfect information', *American Economic Review*, 71, pp. 393–410.

Tschoegl, A. (1993), 'Modeling the behavior of Japanese stock indices', in Takagi, S. (ed.), *Japanese Capital Markets: New Developments in Regulations and Institutions*, Oxford: Blackwell, pp. 371–400.

Van Rijckeghem, C. and Weder, B. (2000), 'Spillovers through banking centers: a panel data analysis', IMF Working Paper WP/00/88, Washington, D.C., International Monetary Fund.

Wallace, N. (1988), 'Another attempt to explain an illiquid banking system: the Diamond and Dybvig model with sequential service taken seriously', *Federal Reserve Bank of Minneapolis Quarterly Review*, 12(Fall), pp. 3–16.

—(1990), 'A banking model in which partial suspension is best', *Federal Reserve Bank of Minneapolis Quarterly Review*, 14(Fall), pp. 11–23.

White, E. (1990), *Crashes and Panics: The Lessons from History*, Homewood, IL, Dow Jones Irwin.