Three Essays on Financial Crises

by

Songklod Rastapana

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Economics

May 2018

THE UNIVERSITY OF WARWICK

Contents

List of Tables

List of Figures

Acknowledgments

I would like to express my deepest gratitude to my supervisors, Herakles Polemarchakis, Marcus Miller, and Ilan Kremer for their continuous support and patient advice for the entire duration of my PhD training. I am also extremely thankful to the Bank of Thailand for providing me with all support to pursue the degree. I would also like to thank the faculty members and colleagues at the Department of Economics, University of Warwick, for many helpful comments and suggestions. Finally, I am deeply indebted to my parents, Manoo Rastapana and Sudjit Rastapana, for their unconditional support and inspiration.

Declarations

This thesis is submitted to the University of Warwick in accordance with the requirements for the degree of Doctor of Philosophy. I declare that none of the chapters in this thesis has been submitted for any other degrees. Chapter 3 is jointly co-authored with my supervisor Marcus Miller and Lei Zhang. *Appendix D.1* from Chapter 4 is borrowed from an earlier version of Chapter 3. The rest of the thesis is solely my own work.

Abstract

This thesis analyses different aspects of financial crises with a focus on the role of pecuniary externalities.

The topics explored in these essays are as follows: Chapter 1 provides background on issues of illiquidity and insolvency, and discusses how the two can interact.

Chapter 2 studies pecuniary externalities in a 'bank run' model where banks supply credit in the form of marketable securities. An aggregate liquidity shock, which triggers 'fire sales' of such securities, can lead to insolvency when their value falls. So, in this type of model, a run on several banks can lead to insolvency driven pecuniary externalities.

Chapter 3 explores three explanations of the U.S. subprime crisis; insolvency due to externalities, insolvency due to cheating, and illiquidity driven by panic. We argue that these narratives should be treated as complements (rather than as substitutes), with each playing an important role at different stages of the crisis.

Chapter 4 studies the reversibility of shocks in a general equilibrium model of competitive markets with heterogeneous beliefs. I find that heterogeneous beliefs can amplify shocks; and, due to asymmetric adjustment of risky asset prices, they can also lead to systemic default when a group of optimistic agents exits the market.

Chapter 1

Illiquidity and Insolvency: Background

1.1 The social role of financial institutions

Before discussing the risks of illiquidity and insolvency facing financial institutions, we briefly summarise the important social role such institutions play in economies with incomplete contracts and borrowers with limited liability. According to Kashyap et al. [2014a], the financial system contributes to economic efficiency in three ways. First, it facilitates the extension of credit for productive investments, which appear as loans on the asset side of bank balance sheets. Second, on the liabilities side of bank balance sheets, liquid claims that facilitate financial transactions are created, with the law of large numbers offering efficiency gains in supplying liquidity insurance to depositors uncertain about the timing of their liquidity needs. Third, with the choice of deposits and bank equity on offer, the financial system facilitates risk sharing.

However, in serving the economy in these three ways, highly leveraged financial institutions are exposed to substantial risks of illiquidity and insolvency. These may, in fact, not be easily separable because, as Kashyap et al. [2014a] point out, the possibility of a bank run may be one of the mechanisms to ensure sound investments by the bank.[1] One problem to be discussed, however, is that this mechanism can be triggered accidentally, by unjustified panic.

[1] Another mechanism discussed in Holmstrom and Tirole [1997] is that the bank keeps sufficient 'skin in the game' in the form of own equity financing.

1.2 Liquidity crises

For the purpose of the analysis in this thesis, I classify a liquidity crisis as a panic-based phenomenon not justified by the fundamental investment activities of banking.[2] This, for example, includes the 'sunspot' bank run in the classic paper of Diamond and Dybvig [1983] driven by coordination failures, where – in the absence of deposit insurance – depositors 'run' on a bank simply because they expect others to do so. Such liquidity shocks may, however, trigger insolvency (as will be discussed below in connexion with Figure 1.1). Given the high leverage and maturity transformation involved in banking, the supply of liquid assets is limited.[3] This, absent 'lender of last resort' facilities, the recall of non-marketable loans before maturity can cause insolvency.

In contrast to Diamond and Dybvig [1983], Allen and Gale [2004] assume that bank lending takes the form of marketable securities. This would seem to solve the problem of illiquidity, as the individual bank can simply sell some assets to pay depositors. This is true for idiosyncratic liquidity shocks. But to the extent that liquidity shock is spread more widely, this selling may trigger changes in the price of these securities; such asset 'fire sales' may cause widespread insolvency, as will be discussed in Chapter 2.

Gorton [2010] argues that panic phenomena may arise simply as a result of the complexity of the financial securities involved. With regard to the U.S. subprime crisis, in particular, he argues that it was 'the loss of information' involved in securitisation and the 'opacity' of mortgage-backed securities in terms of the asset backing that le-d to what he describes as 'The Panic of 2007'. In discussing this view in Chapter 3, we cite the contrary view of Holmstorm, who contends the flawed business model being employed by the investment banks, not the 'loss' information, that caused the crisis.

1.3 Solvency crises

For the purpose of this thesis, a solvency crisis is treated as one that originates on the asset side of the bank balance sheet. There is a crisis when the value of the bank's asset falls so far below the value of its outstanding liabilities that it exhausts

[2] Allen and Gale [2004] distinguish such coordinated failure from 'rational' bank runs based on the poor performance of bank investments.

[3] When the banks share correlated fundamentals, the failure of one bank can be viewed as a 'sunspot', in which depositors consider other banks' failures as indicators of the fragility of their own banks, leading them to turn to an equilibrium in which all depositors run.

the equity buffer. (The issue of insolvency problems possibly leading to illiquidity will be discussed in Figure 1.2.)

The value of a bank asset may be compromised in number of different ways. An important case is a form of cheating in which the complexity of financial products lies at the heart of the problem that is, when a bank uses the financial resources at its disposal to invest in products that are much riskier than outsiders realise. As discussed in Chapter 3, Akerlof and Shiller [2015] argue that this was in fact the case for the U.S. subprime crisis. According to them, the credit rating agencies (CRAs) 'mined their reputation' to secure the fees from the investment banks by over-rating the quality of the securities involved, so that a large number of subprime mortgage securities were, in fact, rated as triple A. Once the risks concealed by the inflated ratings issued by CRAs were discovered, the collapse was inevitable.

The tendency of investors to ignore certain unlikely outcomes, a behavioural phenomenon that Gennaioli et al. [2012] have dubbed 'neglected risk', is worth considering as well.. Empirical evidence in Foote et al. [2012] indicates that the decline in housing prices during the crisis might have been an entirely neglected scenario, rather than being considered as having low probability ex-ante. Neglected risk is also consistent with the phenomenon discussed in Reinhart and Rogoff [2009] as the so-called 'this time is different' mantra. Almost every crisis comes as a surprise to the market because investors who ascribe to such a mantra fail to foresee a crisis when it inevitably comes in a different form. These findings remind us about the importance of unanticipated risk and so-called 'zero probability' events in economic modelling.[4]

Instead of appealing to asymmetric information, an important strand of recent literature emphasises the role of externalities as a potential trigger for insolvency. We focus on the model by Shin [2010], in which highly leveraged investment banks responding to news about the quality of assets generate price effects that greatly amplify the first-round effects. The details of the model and the calibration are supplied in Chapter 3, but it may be useful at this point to discuss the nature of such 'pecuniary externalities'.

A recent paper by Dávila and Korinek [2017] analyses pecuniary externalities in some detail. To illustrate their argument, they use the model of Kiyotaki and Moore [1997], in which there are no financial intermediaries; but the basis of the analysis also applies to models such as that of Shin [2010], in which highly leveraged financial institutions play a central role.

[4]See Lo [2017] for discussion of a new framework, the Adaptive Markets Hypothesis, in which rationality and irrationality coexist.

What they refer to as a 'collateral externalities' emerge when agents are subject to binding price-dependent financial constraints (e.g., borrowing constraints, regulatory capital requirements, Value at Risk (VaR)). Such constraints also include collateral constraints applied to farmers, the focus of study in Kiyotaki and Moore [1997]; and the constraint that takes the specific form of the Value at Risk (VaR) rules of Basel, applied to the investment banks in Shin [2010].[5] However, a common feature to both constraints involves fire sales that induce externalities. There is vast literature on fire sales including Allen and Gale [2004] and Diamond and Rajan [2011]. In fact, such pecuniary externalities generate inefficiencies, resulting in a situation in which 'constrained Pareto improvements' could be achieved by government intervention.[6] Dávila and Korinek [2017] specifically argue that actions to increase liquidity reserves and reduce fire sales could achieve welfare improvement.[7]

1.4 How illiquidity and insolvency interact

As noted above, problems on the liabilities side of a bank's balance sheet can lead to insolvency. Likewise, problems with bank assets can trigger liquidity runs. Figures 1.1 and 1.2 illustrate this situation, and also distinguish the information assumption used in the previously mentioned studies.

Figure 1.1 illustrates in schematic form how illiquidity can lead to insolvency whether or not information is symmetric. For Diamond and Dybvig [1983], information about asset quality is symmetric; but a problem of creditor coordination can lead to insolvency if the bank does not have enough liquid assets, or if it cannot find external financing to meet the withdrawals. Likewise, for Allen and Gale [2004] an aggregate liquidity shock that triggers fire sales of marketable securities can lead to insolvency when their value falls. But this is due to pecuniary externalities, and not to an informational asymmetry. For Gorton [2010], the loss of information is key. For him, information asymmetry generates panic in the asset market. The fall in asset value could be enough to cause insolvency.

[5] For completeness, we should mention that Dávila and Korinek [2017] discuss another case of pecuniary externalities leading to constrained inefficient allocations; this case, the so-called 'distributive externality', emerges when agents have different marginal rates of substitution (MRS) as result of incomplete markets in which the optimal allocation is not possible given the limited span of assets in the economy.

[6] 'Constrained' here refers to the government not having access to better information than market participants.

[7] See *Appendix A* for more discussion of pecuniary externalities and amplification mechanisms within financial systems.

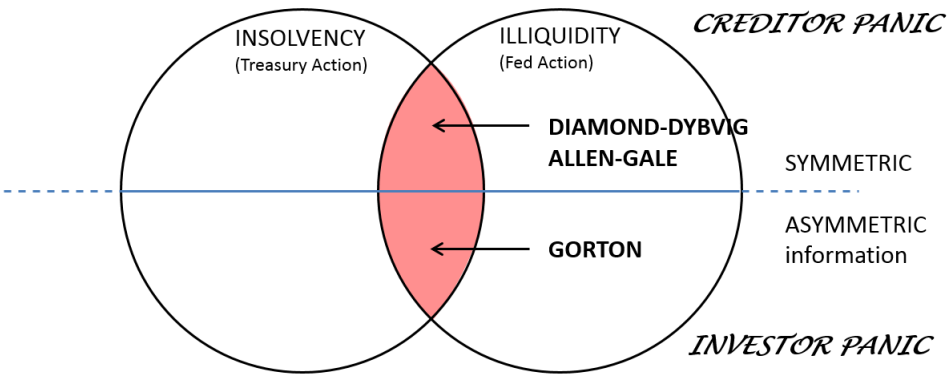


Figure 1.1: Preview of a familiar story (Diamond-Dybvig): illiquidity leading to insolvency (Notes: Gorton, 2009; Krugman 2018)

In symmetric fashion, insolvency shocks can induce illiquidity problems when deterioration in quality on the asset side of the balance sheet involves price falls lowering asset value below outstanding liabilities with mark-to-market accounting. This can occur whether or not information is symmetric. Figure 1.2 illustrates three factors that can lead to insolvency.

For Shin [2010], information is symmetric, but because investment banks hold no liquid assets, bad news about asset quality can trigger fire sales leading to insolvency. In practice (though not part of the model), this can easily lead to a 'silent bank run' in which creditors cease to roll over their lending. Likewise, for Fostel and Geanakoplos [2012] (as discussed in Chapter 4), responses to bad news about asset quality can also lead to substantial falls in asset prices as a result of pecuniary externalities. For Akerlof and Shiller [2015], however, the information asymmetry about asset quality leads to insolvency, as explained above and illustrated in Figure 1.2.

This interaction between shocks to the asset and liabilities side of the balance sheet can, of course, play an important role in accelerating financial crises.

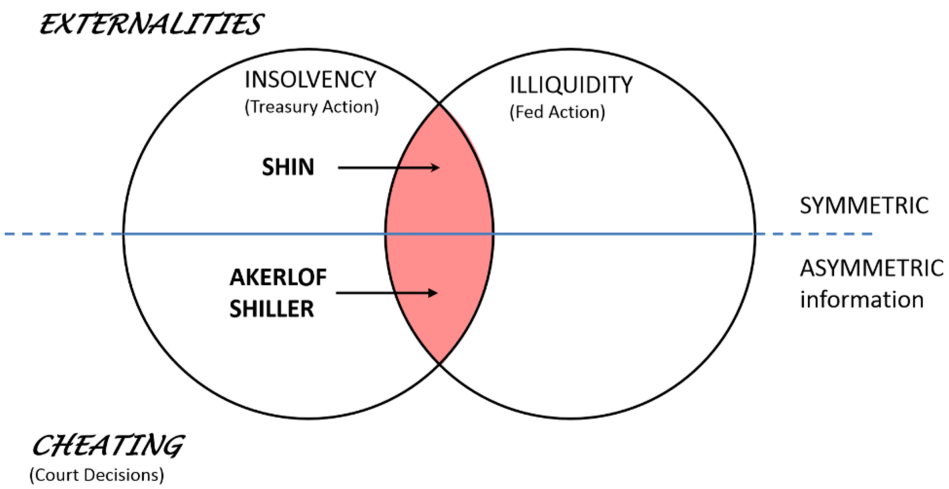


Figure 1.2: Preview of some new perspectives: insolvency leading to illiquidity (Notes: Shin, 2010; Akerlof and Shiller, 2015)

1.5 Defaults

Bankruptcy and default are usually at the heart of both illiquidity and solvency crises, and, in turn, financial crises. However, as mentioned in Goodhart and Tsomocos [2011], incorporating the possibility of default into the model is complex because it is hardly consistent with a number of elements of theoretical models that are currently in use, including complete markets, no aggregate uncertainty, and representative agents. This also highlights the limitations of modern economic theory. Moreover, this amplifies one reason why it is not surprising that the financial crisis was not commonly anticipated.

Early attempts to model default in a general equilibrium framework date back to Shubik [1973] and Shubik and Wilson [1977]. These studies are extended by Goodhart et al. [2006]; and Tsomocos [2003b,a] to incorporate incomplete markets, money, and default; and to highlight the trade-off between financial stability and economic efficiency. In fact, financial fragility in Goodhart et al. [2006] and Tsomocos [2003b,a] refers to a substantial number of defaults, though not full-fledged bankruptcy. In this sense, default is the strategic option. On the one hand, default provides an advantage by allowing borrowers to avoid repaying whole amount of their contractual obligations. On the other hand, it also induces welfare costs to the economy.

The default could be also modeled on collateralised loans where assets are pledged as collateral. This situation could occur when the equilibrium prices are

such that the collateral value is less than the obligated amount the borrowers need to repay. This type of default is further discussed in Chapter 4. Allen and Gale [2004] also show that insolvency and default can occur in equilibrium with the presence of an aggregate liquidity shock, as mentioned above. Indeed, high demand for liquidity could trigger a bank run, leading to an asset 'fire sale'. The fall in asset value could be such that the default is unavoidable. This type of default is further discussed in Chapter 2.

1.6 What's next

In summary, the following analysis of financial crises takes the form of three essays with an underlying focus on both the role of pecuniary externalities, and the possibility of default.

Chapter 2 studies pecuniary externalities in a 'bank run' model in which banks supply credit in the form of marketable securities. An aggregate liquidity shock, which triggers 'fire sales' of such securities, can lead to insolvency when their value falls. As a result, in this type of model, a run on several banks can lead to insolvency-driven pecuniary externalities.

Chapter 3 explores three explanations of the U.S, subprime crisis; insolvency due to externalities, insolvency due to cheating, and illiquidity driven by panic. We argue that these narratives should be treated not as substitutes, but as complements, with each playing an important role at different stages of the crisis.

Chapter 4 studies the reversibility of shocks in a general equilibrium model of competitive markets with heterogeneous beliefs. I find that, due to asymmetric adjustment of risky asset prices, heterogeneous beliefs can amplify shocks, and can lead to systemic default when a group of optimistic agents exits the market.

Chapter 2

Pecuniary Externalities in 'Bank Run' Models

2.1 Introduction

A common feature of many financial crises is the occurrence of bank runs. A key issue is what might trigger such events. As noted in the recent survey by Kashyap et al. [2014b];

> 'In the Diamond and Dybvig (1983) model, a run can occur because of a pure failure to coordinate by the patient depositors; in the jargon that is now popular, a random event like a "sunspot" could lead to a panic where all the patient depositors decide to run (Cass and Shell, 1983).'

This "sunspot" bank run is an idiosyncratic rather than a systemic phenomenon because a run on a particular bank depends on what its depositors expect each other to do. This is not to deny that many banks could experience runs at the same time, either by coincidence (perhaps by all observing the same sunspot), or by contagion spreading from one bank to the other via the interbank market, as discussed in Cifuentes et al. [2005] and [Allen and Gale, 2009, Chapter 10].

The Diamond-Dybvig approach emphasises the illiquidity of bank loans whose value is considerably reduced if recalled early to meet unexpected depositor demand so much so, indeed, that the bank may become insolvent when depositors panic, and the bank is forced to recall good but illiquid loans.

By contrast,Allen and Gale [2004] assume that bank investments are tradeable assets with endogenous liquidation values determined by the sales and purchases of all banks. So, they are explicitly considering systemic behaviour. This leads to

the possibility of insolvency due to 'fire sales' when a group of banks simultaneously sell tradeable securities in response to a depositor run. Indeed, their model can be interpreted to illustrate the role of securitisation when banks package loans into saleable securities, and seemingly eliminate the illiquidity of their balance sheets. While loans do indeed become saleable, this may not avert the risk of illiquidity if many banks are trying to sell at the same time – as in the U.S. subprime crisis. It is interesting to observe that this type of bank run can involve pecuniary externalities operating through price-dependent financial constraints as defined in Dávila and Korinek [2017].

A bank run could also occur in the presence of credit risk, as shown in Kashyap et al. [2017]. This study extends Diamond and Dybvig [1983] to incorporate endogenous credit and run risk, which interact with each other when banks determine their asset portfolio and capital structure. The possibility of a run arises as a result of maturity mismatch where the illiquid long-term loans are risky, being subject to the uncertain payoff of borrowers' projects. Credit risk arises because borrowers could potentially default, depending on the project payoff; and because borrowers' defaults might result in the bank's inability to repay depositors it can trigger runs on the banks. The study shows that more than one regulatory tool is needed to obtain the socially optimal allocations, and to correct distortions on the asset side and on the liability side of banks' balance sheets. This finding is also consistent with Kashyap et al. [2014a], in which more than one regulation is required to correct distortions associated with the possibility of a run.

The analysis in this chapter is related to the existing literature on bank runs, summarised above and in Chapter 1. The principal contribution will be to show how pecuniary externalities can play a role in cases of correlated or systemic crisis. Motivated by Allen and Gale [2004], I extend the model to introduce heterogeneity of banks in a framework in which markets and contracts are incomplete. Uncertainty in the demand of liquidity is also present in this framework as there are 'aggregate liquidity shocks'. Banks operate by offering deposit contracts to depositors in exchange for their endowments, which can be used for investment in short- or long-term assets. The deposit contract offers consumers an insurance against liquidity shocks. However, the incompleteness of the contracts could induce a bank run, leading to financial fragility.

I make two key modifications to the model of Allen and Gale [2004]. First, there are two types of banks: safe banks and risky banks, which are subject to different restrictions and commitments. Safe banks, which could be interpreted as 'narrow' banks, are so called because they choose portfolios such that they can

credibly promise depositors that there will be 'no runs', while the risky banks cannot make this promise. To this end, safe banks are restricted in their use of depositors' endowments, which they can only invest in short-term assets. In contrast, risky banks can only invest in long-term assets. Secondly, although safe and risky banks have different initial restrictions on the use of depositors' endowments, they can participate in an asset market to obtain their privately optimal portfolios. While Allen and Gale [2004] focus on studying asset price volatility in response to a small shock, the main focus of this chapter is in illustrating the source of inefficiency associated with bank runs.

The modifications can be shown to have three consequences. First, assuming that consumers cannot put their endowments into more than one bank, the amount of liquidity in the economy is determined by the fraction of consumers who choose to deposit their endowments into the safe banks. Second, in equilibrium, the safe banks will always be holding 'excess liquidity' as needed to meet demand by the risky banks. Because safe banks require a premium in supplying such liquidity, however, this implies a discounted value of long-term assets in the state when the aggregate demand of liquidity is high. Hence, as in Allen and Gale [2004], the fixed supply of liquidity leads to asset price volatility, even when the shock to the aggregate demand of liquidity is small. Third, given that available contracts are incomplete, the asset price volatility could induce illiquidity and bank runs if price falls trigger a sufficient number of depositors to withdraw at the same time. In this model, however, risky banks fail to internalise the effect of price change due to collective actions on the liquidation value of long-term assets given default. This implies that there will be pecuniary externalities operating through binding price-dependent financial constraints as discussed in Chapter 1. The contribution of this chapter is to illustrate the source of inefficiency associated with bank runs, and to demonstrate how inefficiency works in this type of bank run model.

The important role of pecuniary externalities has received attention in banking regulatory community regarding the need for international standards to address procyclicality that is, the feedback mechanisms between the financial and real sectors of the economy that could amplify the business cycle and potentially lead to credit crunch in recessions. In fact, in responses to the global financial crisis of 2008 the Basel Committee on Banking Supervision (BCBS) in 2010 introduced the countercyclical capital buffer under the Basel III proposal to ensure that banking-sector capital requirements take into account the macro-financial environment in which banks operate. However, sources of procyclicality could be attributed not only to the inherit nature of the financial system, but also to the regulatory stan-

dards themselves. Since before the crisis, this critique has been discussed in the literature. For example, Catarineu-Rabell et al. [2005] use a general equilibrium to study the implication of new risk-based capital requirement proposed by the BCBS. The proposal is based on a bank's decision on a loan-rating scheme in determining the default probability of a borrower, and whether the bank will make the capital requirement procyclical, countercyclical, or neutral. The study indicates that while the bank will not choose a more stable approach, a procyclical approach is a possible choice. Thus, it is important to provide incentives for banks to employ more stable rating schemes. Repullo and Suarez [2013] use a dynamic equilibrium model to study the trade-offs of capital requirements. In fact, on the one hand, their primary microprudential role is to provide a buffer to contain banks' risk of failure. On the other hand, capital requirements potentially induce a procyclical supply of credit, especially in bad times when a credit crunch is possible due to seemingly high failure probabilities. The study indicates interesting findings: while Basel II's reliance on a risk-based capital requirement is more procyclical than Basel I, in which the requirement is flat, Basel II makes banks safer. In addition, Basel III seems to address the trade-offs by aiming at higher capital requirements with less cyclical variation via the implementation of capital conservation and countercyclical buffers. Borio et al. [2001]; Pederzoli et al. [2010]; and Behn et al. [2016] provide further discussion on procyclicality.

The rest of the chapter is organised as follows. Section 2.2 describes the model setting and provides the basic ideas of economies in which there is no aggregate uncertainty for demand of liquidity. Section 2.3 describes the efficient allocation in the presence of aggregate uncertainty. Section 2.4 illustrates market equilibrium in the economy without banks, and discusses determination of asset prices. Section 2.5 demonstrates banking equilibrium in the economy in which two types of banks are distinguished by restrictions with respect to their investments, and by types of deposit contracts. Section 2.6 concludes the chapter by providing numerical examples and discussion.

2.2 The model

2.2.1 Dates and assets

There are three dates, t = 0, 1, and 2. At each date, there is a single commodity that can be used for consumption or for investment, and it serves as a numeraire. There are two types of assets: 1) short-term or liquid assets, with storage technology in which one unit of the commodity invested yields one unit in the next period; 2)

long-term or illiquid assets with productive investment technology, in which one unit of the commodity invested at date 0, produces $R > 1$ units at date 2. Therefore, the investment in long-term assets can be done only at date 0, while the investment in short-term assets can be done at either date 0 or date 1. The returns of both assets are assumed to be certain. The usual trade-off applies between liquidity and returns, with longer-term investments providing higher return, but requiring longer to mature.

2.2.2 Consumers

There are two types of continuum agents: consumers and banks. Consumers are located along a continuum where the measure of consumers is normalized to unity. Each consumer is endowed with 1 unit of consumption of a good at date 0, and nothing else later. Consumers only have access to short-term technology. They are uncertain about their time preference of consumption—whether they want to consume only at date 1 or only at date 2 with probability λ_S and $(1-\lambda_S)$ respectively. λ_S depends on the state of nature and can be λ_H and λ_L having a probability π and $(1-\pi)$ respectively and $0 < \lambda_L < \lambda_H < 1$. All uncertainty is resolved at date 1. The true state is publicly observed, and each consumer learns his type — that is, whether he is an early or a late consumer. An individual's type is private information. The consumers' expected utility is given by:

$$\pi(\lambda_H U(C_{1H}) + (1-\lambda_H)U(C_{2H})) + (1-\pi)(\lambda_L U(C_{1L}) + (1-\lambda_L)U(C_{2L}))$$

where $U(.)$ is a neoclassical utility function (increasing, strictly concave, twice continuously differentiable).

2.2.3 Banks

Banks exist to provide investment and liquidity services to consumers. Banks have access to both short and long technology as a result of their ability to pool the consumers' resources and to invest these resources into short-term and long-term assets. Banks can offer a better combination of risks and asset returns to consumers. In addition, banks also have access to the asset market, while consumers do not. So, banks can buy or sell long-term assets at date 0 and date 1. Banks compete by offering deposit contracts (e.g. demand deposits) to consumers in exchange for their endowments, and consumers respond by choosing the most attractive contracts offered to them. The deposit contract promises a fixed amount of goods, (d_1, d_2). If a consumer withdraws at date 1, he will receive d_1 units of consumption goods,

and he will receive d_2 units of consumption units if he withdraws at date 2. Free entry ensures that the banking market is perfectly competitive and earns zero profit in equilibrium. Therefore, deposit contracts offered by banks in equilibrium must maximise the utility of consumers. As banks can do anything that consumers can do, there is no loss of generality in assuming that consumers deposit their entire endowment in a bank at date 0.

There are two types of banks: safe banks and risky banks. At date 0, safe banks can only use goods from depositors to invest in short-term assets, while risky banks can only use goods from depositors to invest in long-term assets. However, both types of banks have access to the market at date 0. So, they can have access to both technologies. In offering deposit contracts to depositors, a safe bank is considered safe because it can commit no run, while risky banks cannot commit no run at date 0. Consumers cannot diversify by putting their money into more than one bank.

2.2.4 Uncertainty

There are two sources of uncertainty in the model. First, each consumer faces the idiosyncratic uncertainty about their time preference of consumption, by which they could be early or late consumers. Second, banks face an aggregate uncertainty about the fraction of early and late consumers; this fraction depends on the state of nature S. By pooling the consumers' resources, the 'law of large numbers' suggests that the fraction of early consumers in state S among banks is identically equal to the probability λ_S.

2.2.5 Markets

There exist asset markets for banks at date 0 and date 1. Therefore, banks can use the asset market at date 0 to achieve better risk sharing. However, they cannot hedge against uncertainty in an aggregate demand for liquidity because the market is incomplete while the state-contingent securities are not available. Banks will use the market at date 1 when all uncertainty is revealed to obtain liquidity or to discard excess liquidity as required. As such, the market is incomplete at date 0, and it is complete at date 1, when all uncertainty is realised.

2.2.6 Economies without aggregate uncertainty

This section aims to point out the basic role of banks in offering a better combination of risks and returns to consumers where there is no aggregate uncertainty. Let us

assume for now that a consumer has access to both short-term and long-term assets, rather than having access only to short-term assets As noted here and in other literature, in the absence of aggregate uncertainty, the banking solution is efficient because consumers can use the market at date 1 to diversify the idiosyncratic risk away. However, the market solution in the absence of banks in the economy is efficient only in the specific case $U(C) = ln(C)$. This is because the market provision of liquidity is inefficient. For equilibrium to exist, consumers must be willing to hold both assets at date 0. However, they will do so only when $P = 1$, because otherwise, one asset is dominated by the other. As a result, the market at date 0 fails to reveal how much investors would be willing to pay contingent on knowing their types. The details about solutions in different cases could be described as follows:

The **efficient solution** is described as $\frac{U'(C_1)}{U'(C_2)} = R$, while $C_1 = \frac{y}{\lambda}$ and $C_2 = \frac{(1-y)R}{1-\lambda}$. If we assume that $U(C) = ln(C)$, the planner's solution is $(y, C_1, C_2) = (\lambda, 1, R)$. The allocation is efficient when the marginal benefit of liquidity in smoothing the consumption, and the marginal cost of lower expected consumption are balanced. The optimal liquidity at date 0 in this special case is equal to the probability of being an early consumer λ. The optimal consumption of early and late consumers is $1 and R$, respectively. With the assumption of $U(C) = ln(C)$ or $\sigma = 1$ under the Constant Relative Risk Aversion (CRRA) utility function, an interesting result is that the optimal level of investment is equal to the probability of being an early consumer λ, while the return on long-term assets plays no role in determining the optimal y. This is because the optimal consumption plan matches the return of short-term and long-term assets; this return is $(C_1, C_2) = (1, R)$. In particular, the optimal allocation doesn't require the return on long-term assets to be shared with the early consumer. Also, the return from investment in short-term assets will be used only for the early consumer's consumption, while the return from investment in long-term assets will be used only for the early consumer's consumption. As a result, the planner will choose $y = \lambda$ to satisfy the fraction, λ, of consumers who are going to be early consumers and receive a consumption level equal to 1, which matches the return on short-term assets.

In the **Autarky economy**, in which consumers are unable to trade assets and have to consume the returns generated by their portfolio, the solution is$(y, C_1, C_2) = (\frac{\lambda R}{R-1}, \frac{\lambda R}{R-1}, (1-\lambda)R)$. The solution is inefficient because a consumer has no mechanism to transfer wealth when the uncertainty is resolved at date 1. As shown in *Appendix B.1*, in the autarky economy, $C_2 = (1-\lambda)R < R$, while $C_1 = \frac{\lambda R}{R-1}$, which can be greater or less than 1.

In the **market economy**, we further assume that a consumer also has access

to the asset market in which consumers can trade with each other once uncertainty is revealed at date 1. In equilibrium, the price of the asset at date 1, P, is equal to 1. The market equilibrium allocation would give $(C_1, C_2) = (1, R)$. Given that the consumer's utility function belongs to the CRRA, the market's provision of liquidity is efficient only in a special case, when $\sigma = 1$. The market is inefficient if $\sigma > 1$, when a risk-averse consumer values consumption smoothing more than when $\sigma = 1$, given the same expected consumption. Therefore, the efficient solution in this case would be $C_1 > 1$ and $C_2 < R$; that is, the efficient solution requires the higher return of a long-term asset to be shared between the early consumer and the late consumer. Similarly, the market is inefficient if $\sigma < 1$, when the consumer becomes less risk averse, and prefers to increase the expected consumption. Thus, the efficient solution requires $C_1 < 1$ and $C_2 > R$.

This reminds us about the trade-off between liquidity and return. In particular, it shows the benefit of smoothing consumption, and the cost of foregone returns, which result in lower expected consumption. In other words, intertemporal insurance against liquidity shock, in which consumer might need liquidity at date 1 with probability λ, is costly. In the efficient solution, higher risk aversion implies a higher marginal utility of liquidity, $\lambda U'(C_1)$, which should be equal to the marginal cost of liquidity in forgoing returns, $\lambda R U'(C_2)$.

It must be noted that the perfectly insured allocation becomes efficient when the preference is Leontief, and $\sigma \to \infty$. I will show later that when there exists an aggregate uncertainty, perfectly insured allocation is also possible even if the preference is not Leontief.

In the absence of aggregate uncertainty, the banking solution is obviously efficient, since by pooling the consumers' endowments, banks allow consumers to share higher returns of long-term assets, thereby providing a better combination of risks and returns. In particular, the banking problem becomes similar to the planner problem.

To sum up, a consumer is better off when he has access to the asset market rather than when he is in autarky. Since a consumer in autarky must consume the returns generated by his portfolio, his constraints are $C_1 = y$ and $C_2 = y + (1 - y)R$ for $y \in [0, 1]$. Therefore, the maximum consumption for an early consumer is obtained when $y = 1$ and $C_1 = 1$, and the maximum consumption for a late consumer is obtained when $y = 0$ and $C_2 = R$. The autarky allocation is dominated by the market allocation since the market allocation will provide greater consumption on one date, and at-least-as-good consumption on the other date. As a result, access to the asset market improves the expected utility of the consumer. However, the

market solution is only efficient in a specific case – when $\sigma = 1$.
Proof. See *Appendix B.1.* for completeness.

2.3 The first best equilibrium

The planner's problem is to maximise the utility of the representative consumer by making investment decisions on short-term and long-term assets at date 0, and determining the consumption at date 1 and date 2. At date 0, a planner decides the allocation of a representative consumer with endowment of 1 unit of goods into y units of short-term assets and $1y$ units of long-term assets, where $y \in (0, 1)$. The planner also assigns the consumption bundle, $C = \{C_L, C_{2L}, C_{1H}, C_{2H}\}$, to a representative consumer. At date 1, the available consumption amount for the representative consumer is given by the amount y invested in the short-term asset. Given that λ represents a fraction of early consumers, the feasible allocation of the planner is $\lambda_S C_{1S} \leq y$, where $S \in \{H, L\}$. At date 2, the available consumption amount for the representative consumer is given by the return from investment for the long-term asset, $(1-y)R$, and the remaining goods from date 1 , $y-\lambda_S C_{1S}$. The remaining goods is restored using short-term technology to transfer consumption to a late consumer at date 2. Therefore, the available consumption amount at date 2 is given by $(1-\lambda_S)C_{2S} \leq (y-\lambda_S C_{1S}) + (1-y)R$. The equation holds with equality, as a strictly concave assumption implies that all available goods from investment will be used up in equilibrium. The social planner's problem is described by:

$$\max_y \pi(\lambda_H U(C_{1H}) + (1-\lambda_H)U(C_{2H})) + (1-\pi)(\lambda_L U(C_{1L}) + (1-\lambda_L)U(C_{2L})) \quad (2.1)$$

subject to:

$$\begin{aligned} \lambda_H C_{1H} &= y & (1-\lambda_H)C_{2H} &= (1-y)R \\ \lambda_L C_{1L} &\leq y & (1-\lambda_L)C_{2L} &= (1-y)R + (y - \lambda_L C_{1L}) \end{aligned}$$

First, let us assume that the constraints in both the states are binding. Given the optimal holding of liquidity, y, the efficient allocation ***across time*** provides $C_{1H} < C_{1L}$ and $C_{2H} > C_{2L}$ as $\lambda_H > \lambda_L$. Then, if $C_{1L} < C_{2L}$, the consumption bundle becomes $C_{1H} < C_{1L} < C_{2L} < C_{2H}$. If $C_{1L} \geq C_{2L}$, then the consumption bundle becomes $C_{1H} < C_{1L} = C_{2L} < C_{2H}$ as the consumer will be better off by postponing consumption from date 1 to date 2 until $C_{1L} = C_{2L}$. In the latter case, the feasible constraint is not binding in state L, and consumers receive fully insured consumption bundles regardless of their types. Fully insured consumption is equal to

the depositor's expected consumption $y+(1-y)R$. Therefore, the efficient allocation across time ensures binding constraints in state H, while the constraint in state L is either binding or not binding, depending on the optimal C_{1L} and C_{2L}. Intuitively, it is never optimal for the planner to carry forward liquidity from date 1 to date 2 in both states, as the planner can do better by allocating more goods into long-term assets, and can enjoy higher return and higher expected consumption. Also, the efficient solution, where $C_{2S} \geq C_{1S}$, provides an *incentive-compatible outcome* in the sense that there are no incentives for the late consumer for pretending to be an early consumer. Therefore, we can relax the assumption that a planner needs to know the investors' type. The solution can be summarised by the following consumption functions:

$$\begin{aligned} C_{1H} &= \frac{y}{\lambda_H} & C_{2H} &= \frac{(1-y)R}{(1-\lambda_H)} \\ C_{1L} &= \min\{\frac{y}{\lambda_L},\ y+(1-y)R\} & C_{2L} &= \max\{\frac{(1-y)R}{1-\lambda_L},\ y+(1-y)R\} \end{aligned} \tag{2.2}$$

From the solution above, we see that the aggregate liquidity will be used up at date 1 to supply only for early consumers as $\lambda_S C_{1S} = y$, except when $C_{1L} > C_{2L}$, in which case the planner can do better by shifting some consumption to the late consumers until $C_{1L} = C_{2L}$. Then, it is easy to show that a consumer obtains higher expected utility in state L than that in state H due to the lower consumption volatility in state L given that the same expected consumption of $y + (1-y)R$.[1] In sum, all these features are provided by the liquidity shock in terms of aggregate uncertainty on demand for liquidity.

Second, the efficient allocation ***across states*** also suggests the optimal holding of liquidity, y, where the marginal benefit of y in state H is equal to the marginal cost in state L. In particular, the increase in y results in a higher expected utility in state H (positive marginal utility), but it comes with the cost of a lower expected utility in state L (negative marginal utility). As mentioned in the case of without aggregate uncertainty, the optimal allocation, $\sigma = 1$, is described as $(y, C_1, C_2) = (\lambda, 1, R)$. Given that this is a benchmark for the economy with aggregate uncertainty, the aggregate liquidity in state H emphasises the benefit of expected consumption as $C_{1H} < 1$ and $C_{2H} > R$. In state L, the aggregate liquidity in the market emphasises the benefit of intertemporal smoothing as $C_{1L} > 1$ and $C_{2L} < R$, where the return of long-term asset is required to be shared with the

[1] Depositor' expected consumption $E[C_S] = \lambda_S C_{1S} + (1-\lambda_S)C_{2S} = y + (1-y)R$

early consumers. The efficiency requires the balance between the benefit in state H (i.e. cost in state L) and in state L (i.e. cost in state H). This is the standard phenomenon in this type of model in the economy in which the supply of liquidity is fixed by the initial portfolio choices of liquidity y, but there are subsequent shocks to the demand for liquidity. I will show that in market equilibrium (in Section 2.4) and in banking equilibrium (in Section 2.5) that this fixed supply can cause substantial asset-price volatility. In sum, increasing liquidity in the market results in greater consumption smoothing, but with the cost of lower expected consumption.

The solution to the planner's problem consists of a portfolio choice of, y^*, and a consumption bundle, $C^* = \{C^*_L, C^*_{2L}, C^*_{1H}, C^*_{2H}\}$, such that y^* solves the portfolio choice problem Eq(2.1), which is given by Eq(2.3), and the consumption bundle satisfies Eq(2.2).

$$\frac{\pi U'(C_{1H}) + (1-\pi)U'(C_{1L})}{\pi U'(C_{2H}) + (1-\pi)U'(C_{2L})} = R \tag{2.3}$$

Supposing that $U(C) = ln(C)$, the solution to the planner's problem is described by:

$$y* = \begin{cases} \frac{((R-1)(1+\pi\lambda_H)+\pi)-\sqrt{(R-1)(1+\pi\lambda_H)+\pi)^2-4\pi\lambda_H R(R-1)}}{2(R-1)} & \text{if} \quad R \leq \hat{R} \\ \pi\lambda_H + (1-\pi)\lambda_L & \text{if} \quad R > \hat{R} \end{cases}$$

and

$$C^*_{1H} = \frac{y^*}{\lambda_H} \qquad C^*_{2H} = \frac{(1-y^*)R}{(1-\lambda_H)}$$

$$C^*_{1L} = \begin{cases} y^* + (1-y^*)R & if \quad R \leq \hat{R} \\ \frac{y^*}{\lambda_L} & if \quad R > \hat{R} \end{cases} \qquad C^*_{2L} = \begin{cases} y^* + (1-y^*)R & if \quad R \leq \hat{R} \\ \frac{(1-y^*)R}{(1-\lambda_L)} & if \quad R > \hat{R} \end{cases}$$

where

$$\hat{R} \equiv \frac{(1-\lambda_L)(\pi\lambda_H + (1-\pi)\lambda_L)}{\lambda_L(1-(\pi\lambda_H + (1-\pi)\lambda_L))}$$

and λ_L represents the probability of being an early consumer in state L, and $\pi\lambda_H + (1-\pi)\lambda_L$ represents the probability of being an early consumer. In addition, there exists a critical long-term asset return, $\hat{R}$, such that if R is lower than the critical value, $C_{1L} = C_{2L}$.

Proof. See *Appendix B.2.*

2.3.1 Comparative statics

When $U(C) = ln(C)$, the comparative statics in Figure 2.1 suggests that in the case of a binding feasible constraint, when no storage technology is being used at date 1 and $C_{1L} < C_{2L}$, the optimal holding of liquidity in the economy increases with $\pi, \lambda_H,$ and λ_L, while R plays no role in determining the optimal liquidity. This is because the binding constraints in both the states imply that the optimal allocation doesn't require returns of each asset to be shared between the two types of consumers. In other words, the returns from short-term assets will be delivered to the early consumers, while the returns from long-term assets will be delivered to the late consumers; this is similar to the efficient allocation when there is no aggregate uncertainty. In case of the non-binding solution, when $C_{1L} = C_{2L}$, the optimal holding of liquidity in the economy increases with π and λ_H, and decreases with R, while λ_L plays no role in determining the optimal liquidity because, in state L, consumers receive fully insured consumption bundle regardless of their types equal to the expected consumption, $y + (1 - y)R$.

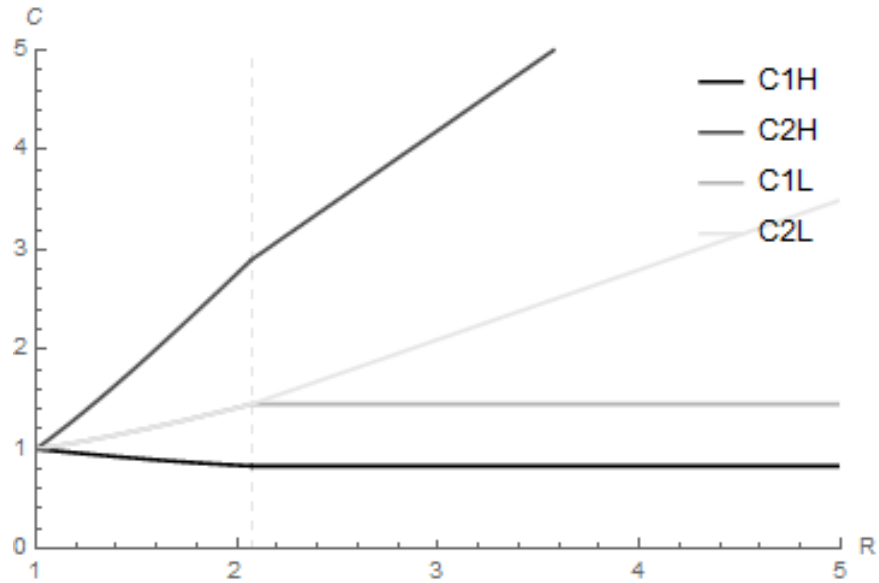


Figure 2.1: Consumption functions at date 1 and 2 as a function of R
Given: $\pi = 0.6$; $\lambda_H = 0.7$; $\lambda_L = 0.4$; $U(C) = ln(C)$

2.4 Market allocation

This section examines the market solution in the absence of banks in the economy. I make two initial assumptions:first, a consumer has access to both short- and

long-term technology at date 0, and, as usual, a consumer has access to the short technology at date 0. Thus, he decides upon initial portfolio choices of short-term asset, y, and long-term asset, $1-y$, at date 0. Second, there also exists a market for consumers to trade long-term assets with each other at date1 once the uncertainty is revealed. In particular, early consumers can get rid of their holdings of long-term assets, while late consumers can get rid of their holdings of short-term assets using the market. Let P_S denote the price of long-term assets at date 1 in state S, where $S \in \{H, L\}$. A consumer's problem is:

$$\max_{y} \quad \pi(\lambda_H U(C_{1H}) + (1-\lambda_H)U(C_{2H})) + (1-\pi)(\lambda_L U(C_{1L}) + (1-\lambda_L)U(C_{2L})) \tag{2.4}$$

subject to:

$$\begin{aligned} C_{1H} &= y + P_H(1-y) \quad C_{2H} = \left((1-y) + \frac{y}{P_H}\right) R \\ C_{1L} &= y + P_L(1-y) \quad C_{2L} = \left((1-y) + \frac{y}{P_L}\right) R \end{aligned} \tag{2.5}$$

Constraints are binding with equality because an early consumer wants to consume as much as possible at date 1, and a late consumer wants to consume as much as possible at date 2. Therefore, an early consumer will sell all his long-term assets at date 1, and his consumption will be equal to the present value of their portfolio at date 1, which is $y + P_S(1-y)$. And, the late consumer will use all of the consumption goods to buy long-term assets at date 1, given that its return is not less than the return of short-term assets. The late consumer's consumption at date 2 is, thus equal to $\left((1-y) + \frac{y}{P_S}\right) R$.

2.4.1 Asset prices and market clearing

The price of long-term assets is determined by their demand and supply in the market. Demand for long-term assets comes from late consumers who want to consume at date 2, while early consumers will never want to hold on to long-term assets since they only value consumption at date 1. At date 1, both asset returns are certainly known since all uncertainty is revealed. The return of short-term assets for one unit of goods invested is equal to 1, while the return from long-term assets is equal to $\frac{R}{P_S}$. With access to storage technology at date 1, late consumers are willing to exchange their goods, received from investment in short-term assets at date 0 and date 1, for long-term assets if the returns on the long-term assets are greater than or equal to the return from short-term assets. In particular, if $\frac{R}{P_S} < 1$, the late consumer will prefer short-term assets and there will be no demand for long-term

assets. If $\frac{R}{P_S} = 1$, late consumers will be indifferent about holding short-term assets and long-term assets, and the demand for long-term assets becomes perfectly elastic. Since the total available goods in the economy coming from late consumers at date 1 can be used to buy the long-term assets is $(1-\lambda)y$, the upper bound of aggregate demand for long-term assets, where $R = P_S$, becomes $\frac{(1-\lambda)y}{R}$. If $\frac{R}{P_S} > 1$, the late consumers will only want to hold on to the long-term assets and will inelastically supply all their goods, exhibiting 'cash-in-the-market pricing'. Aggregate demand for long-term assets becomes $\frac{(1-\lambda)y}{P_S}$. The aggregate demand for long-term asset is, thus, described by:

$$D(P_S) = \begin{cases} 0 & if \quad \frac{R}{P_S} < 1 \\ [0, \frac{(1-\lambda_S)y}{R}] & if \quad \frac{R}{P_S} = 1 \\ \frac{(1-\lambda_S)y}{P_S} & if \quad \frac{R}{P_S} > 1 \end{cases}$$

Early consumers will inelastically supply long-term assets at date 1 as there is no incentive in carrying forward any long-term assets to date 2 because they only value consumption at date 1. As fraction λ of consumers are late consumers, and each of them holds $(1-y)$ units of long-term assets, the aggregate supply of long-term assets is described by:

$$S = \lambda(1-y)$$

Combining the aggregate demand and the aggregate supply of long-term assets, we can see that there are two possible equilibrium prices, which are (1) $P = R$, when $S \leq D(R)$, in which excess demand for long-term assets is possible, and late consumers are indifferent between short-term and long-term assets, and (2) $P = \frac{(1-\lambda_S)y}{\lambda_S(1-y)}$, when $S > D(R)$. The latter case demonstrates *the cash-in-the-market pricing* as mentioned earlier. The market clearing gives the following price in equilibrium and can be shown in Figure 2.2.

$$P_S = \min\left\{R\,,\ \frac{(1-\lambda_S)y}{\lambda_S(1-y)}\right\} \tag{2.6}$$

It is worthwhile to note that in this case where assets in the market represent two technologies with trade-off between returns and investment maturity, excess demand or excess supply are possible when consumers are indifferent between both types of assets at date 1. In this case, when $\frac{R}{P_S} = 1$, the excess demand for long-term assets in the market is equal to $\frac{(1-\lambda_S)y}{P_S} - \lambda_S(1-y)$, which can also be interpreted as excess supply of short-term assets, and is equal to $(1-\lambda_S)y - \lambda_S(1-y)R$.

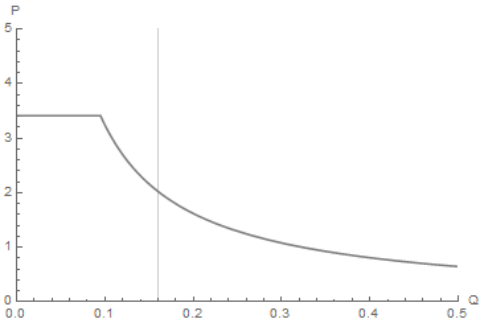


Figure 2.2: Demand and supply of long asset at date1

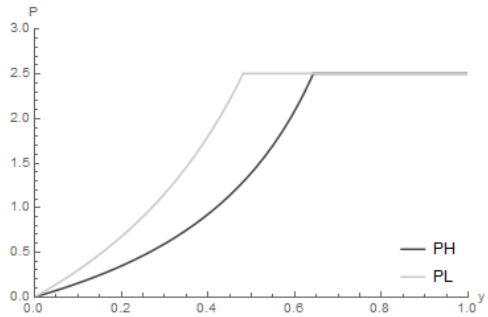


Figure 2.3: Price of long asset in state H and L as a function of y

The market-clearing condition requires $P_H < 1$ and $1 < P_L \leq R$. For a market to be clear at date1, both short-term and long-term assets must be held by consumers at date 0. Given that the market-clearing condition gives $P_S = \min\{R, \frac{(1-\lambda_S)y}{\lambda_S(1-y)}\}$, where $\lambda_H > \lambda_L$, we know that $P_L > P_H$. Therefore, $P_H < 1$ because if $P_H \geq 1$, then $P_L > P_H \geq 1$. Then, the long-term assets dominate the short-term assets at date 0, and no one wants to hold short-term assets. This cannot be an equilibrium. In addition, $1 < P_L \leq R$ because if $P_L \leq 1$, then $P_H < P_L \leq 1$. Then, the short-term assets dominate the long-term assets at date 0, and no one wants to hold long-term assets. This again cannot be an equilibrium. See Figure 2.3 for the prices of long-term asset as a function of aggregate liquidity, y.

2.4.2 Equilibrium

First, given the optimal holding of liquidity, y, the market allocation provides efficient outcome *across time* because agents can price in their willingness to buy/sell long-term assets given their types. As previously shown in the efficient allocation, the market allocation across time ensures binding constraints in state H when there exists cash-in-the-market pricing, that is $P_H = \frac{(1-\lambda_H)y}{\lambda_H(1-y)}$, while the constraint in state L is binding (or not) depending upon the optimal C_{1L} and C_{2L}. If $C_{1L} < C_{2L}$, then the constraint is binding, and there is cash-in-the-market-pricing in state L, that is $P_L = \frac{(1-\lambda_L)y}{\lambda_L(1-y)}$. However, if $C_{1L} \geq C_{2L}$, the constraint is not binding, and the price reaches the upper bound, that is $P_L = R$. In the former case, the consumption bundle becomes $C_{1H} < C_{1L} < C_{2L} < C_{2H}$, while in the latter case, the consumption bundle becomes $C_{1H} < C_{1L} = C_{2L} < C_{2H}$, in which consumers, regardless of their types, receive fully insured consumption bundles in state L equal to the expected consumption $y + (1-y)R$.

Second, the market allocation in states H and L determines an aggregate

amount of liquidity in the market, y. The market provision of liquidity is efficient when $\sigma = 1$, when the market outcome equalises marginal benefits of liquidity in state H and the marginal costs in state L. However, when $\sigma \neq 1$, market provision of liquidity is inefficient because the incompleteness of the market prevents efficient risk sharing across states. Accordingly, market fails to price in agents' willingness to buy/sell long-term assets contingent on their types. However, if the market is complete, and the consumer can trade contingent securities at date 0, the prices of contingent securities should reveal the willingness of market participants to buy and sell the long-term assets, and should allow efficient outcome *across states*.

Overall, *the market provides efficient allocation across time given the holding of* y, since early consumers can liquidate long-term assets, while late consumers can acquire more long-term assets at date1. However, *the market provision of liquidity is inefficient, except when* $\sigma = 1$, because there are no state-contingent securities to provide efficient risk sharing across states.

The market allocation consists of a portfolio choice, y, and a consumption bundle, $C = \{C_{1L}, C_{2L}, C_{1H}, C_{2H}\}$ such that y solves the portfolio-choice problem Eq(2.4), which is given by Eq(2.7), and the consumption bundle satisfies Eq(2.5).

$$\frac{(\pi U'(C_{1H}) + (1-\pi)U'(C_{1L})) - (\pi U'(C_{1H})(\frac{1-\lambda_H}{1-y}) + (1-\pi)U'(C_{1L})(\frac{1-\lambda_L}{1-y}))}{(\pi U'(C_{2H}) + (1-\pi)U'(C_{2L})) - (\pi U'(C_{2H})(\frac{\lambda_H}{y}) + (1-\pi)U'(C_{2L})(\frac{\lambda_L}{y}))} = R \tag{2.7}$$

where the price function is given by Eq(2.6).

2.5 Banking equilibrium

Banks exist to improve consumers' welfare by offering them the deposit contracts that provide a better combination between risks and returns. Given that the competitive banking market ensures zero profit of the bank in equilibrium, without loss of generality, I can assume that late consumers will be resident claimants who receive the residue of the bank's assets at date 2. As a result, the deposit contract offered at date 0 is characterised by the promised payment at date 1, d, and the remaining assets, including the returns of long-term assets, will be the consumption of the late withdrawer.

2.5.1 Safe bank

At date 0, the safe bank that has a 'no run' commitment offers a deposit contract that specifies the promised payment at date 1, in exchange for consumption goods. The

consumption goods received from consumers in exchange for the deposit contract will be first invested in short-term assets by the safe bank; the safe bank can have access to the long-term assets by participating in the assets market to share risk with the risky bank at the end of date 0. Let y^S denote the amount of short-term assets held by the safe bank at the end of date 0. For every unit of consumption goods deposited in a safe bank, the amount of long-term assets held by a safe bank is described by $\frac{1-y^S}{P_0}$, where P_0 denotes the price at date 0 of the long-term assets. At date 1, as the safe bank has a 'no run' commitment, it will deliver promised payments of d^S to consumers regardless of the state of nature. At date 2, the deposit contract provides the consumer who withdraws at date 2 the residue of the bank's assets. The problem of the safe bank is:

$$\max_{y^S, d^S} \quad \pi(\lambda_H U(C^S_{1H}) + (1-\lambda_H)U(C^S_{2H})) + (1-\pi)(\lambda_L U(C^S_{1L}) + (1-\lambda_L)U(C^S_{2L})) \quad (2.8)$$

subject to:

$$C^S_{1L} = d^S \qquad C^S_{2L} = \frac{y^S - \lambda_L d^S + P_L(\frac{1-y^S}{P_0})}{(1-\lambda_L)(\frac{P_L}{R})}$$

$$C^S_{1H} = d^S \qquad C^S_{2H} = \frac{y^S - \lambda_H d^S + P_H(\frac{1-y^S}{P_0})}{(1-\lambda_H)(\frac{P_H}{R})}$$

2.5.2 Risky bank

In contrast to a safe bank, a risky bank does not have a 'no run' commitment to a depositor. The deposit contract offered by the risky bank at date 0 is characterised by the promised payment at date 1, d^R, and the remaining assets, including the returns of long-term assets will be the late withdrawer's consumptions. However, goods received from consumers in exchange for the deposit contract will be first invested in long-term assets, and the risky bank can have access to the short-term assets by participating in the asset market at the end of date 0. Let y^R denote the amount of short-term assets held by the risky bank at the end of date 0. The amount of long-term assets held by the risky bank is described by $1 - \frac{y^R}{P_0}$.

Since at date 1, the risky bank cannot make a 'no run' commitment, we will explore under what conditions the bank will be forced to default on its promise to depositors. In the event of bankruptcy, the risky bank is required to liquidate all its assets in an attempt to provide the promised payment, d^R, to the consumers who withdraw at date 1. Unlike Diamond-Dybvig, in which the liquidation value of long-

term assets at date 1 is exogenous, the liquidation value is endogenously determined in equilibrium due to the presence of a competitive market in this model.[2] The liquidation value of the risky bank's assets is, thus, described by $y^R + P_H(1 - \frac{y^R}{P_0})$.

Once all uncertainty is revealed at date 1, consumers learn their types, and state S is realised. While an early consumer will withdraw at date 1 for sure (since he only values consumption at date 1), a late consumer has an option of withdrawing either at date 1 and using the short-term assets to transfer his consumption to date 2, or withdrawing at date 2. The former case generates the possibility of a bank run if the late consumer prefers to withdraw at date 1, rather than waiting until date 2. The late consumer is willing to do so if waiting until date 2 give him less consumption than withdrawing at date1. In particular, there will be a bank run if and only if $\frac{y^R + P_S(1-\frac{y^R}{P_0}) - \lambda_S d^R}{(1-\lambda_S)} \frac{R}{P_S} < d^R$, where $S \in \{H, L\}$. The left-hand side of the inequality represents the consumption if late consumers withdraw at date 2. In what follows, the no-default constraint is described by:

$$\lambda_S d^R + (1 - \lambda_S)\frac{P_S}{R} d^R \quad \leq \quad y^R + P_S(1 - \frac{y^R}{P_0}) \tag{2.9}$$

In case of 'no run', the deposit contract provides the consumer who withdraws at date 2 the residue of the bank's asset. The problem of the risky bank can be written as

$$\max_{y^R, d^R} \quad \pi(\lambda_H U(C^R_{1H}) + (1 - \lambda_H)U(C^R_{2H})) + (1 - \pi)(\lambda_L U(C^R_{1L}) + (1 - \lambda_L)U(C^R_{2L})) \tag{2.10}$$

subject to:

$$C^R_{1L} = d^R \qquad\qquad C^R_{2L} = \frac{y^R - \lambda_L d^R + P_L(1 - \frac{y^R}{P_0})}{(1 - \lambda_L)(\frac{P_L}{R})}$$

$$C^R_{1H} = \begin{cases} d^R & \text{if (2.9) is satisfied} \\ y^R + P_H(1 - \frac{y^R}{P_0}) & \text{otherwise} \end{cases} \qquad C^R_{2H} = \begin{cases} \frac{y^R - \lambda_H d^R + P_H(1-\frac{y^R}{P_0})}{(1-\lambda_H)(\frac{P_H}{R})} & \text{if (2.9) is satisfied} \\ y^R + P_H(1 - \frac{y^R}{P_0}) & \text{otherwise} \end{cases}$$

The model setting is summarised in Figure 2.4 below.

[2]The presence of aggregate uncertainty in this model represents another difference from Diamond-Dybvig.

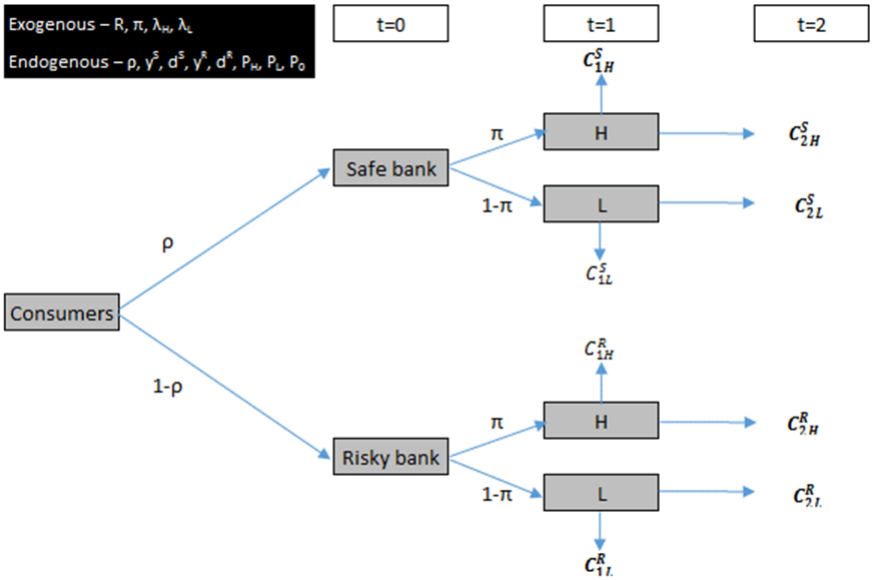


Figure 2.4: Model settings in banking economy

2.5.3 Market clearing

Safe and risky banks use the market at date 1 to obtain liquidity, or to discard excess liquidity. The price of the long-term assets in state S is again denoted by P_S, where $S \in \{H, L\}$. Let us assume that fraction ρ and $1-\rho$ of consumers put their endowments into safe and risky banks, respectively.

Since a bank run is possible in equilibrium, I split the market clearing into two cases: first, in equilibrium without a bank run, the market clearing requires $P_L = R$ and $P_H < 1$ in order to have both assets held by banks in the economy. In this equilibrium, the amount of liquidity in the market is just enough to provide banks' promises to early consumers in state H. Therefore, there is excess supply of liquidity in state L, and P_L is bid up to R when the returns of long-term and short-term assets are the same, and when a bank is indifferent between short-term and long-term asset. For a market to be cleared at date 0, P_H needs to be strictly lower than 1; otherwise the long-term assets will dominate the short-term assets at date 0. Since only safe banks can have access to short technology at date 0, the aggregate supply of liquidity in the market is equal to ρ, while the aggregate demand for liquidity in state H is equal to $\lambda_H(\rho d^S + (1-\rho)d^R)$. I will show in the next section that in equilibrium without run, the safe bank and the risky bank make identical choices (y and d) at date 0, and they hold just enough liquidity to keep their promise in state H. Although the asset market is not essential in this case, we

will still need it in order for equilibrium at date0 to exist. The aggregate demand for liquidity in state L is equal to $\lambda_L(\rho d^S + (1-\rho)d^R)$, strictly less than aggregate liquidity in the economy. Therefore, the inelasticity of demand and of the supply of liquidity mean that small aggregate liquidity shocks can cause large fluctuations in asset prices. I will show later that bank runs provide less volatile asset prices. The market clearing conditions are described as follows:

$$\begin{aligned}
&\text{Market clearing in state L} && P_L = R \\
&\text{Market clearing in state H} && \rho(y^S - \lambda_H d^S) = (1-\rho)(\lambda_H d^R - y^R) \\
&\text{Market clearing at date0} && \rho(1-y^S) = (1-\rho)y^R
\end{aligned} \tag{2.11}$$

Second, in equilibrium with bank run, the market clearing requires that $1 < P_L \leq R$ and $P_H < 1$. In this case, the amount of liquidity in the market is just enough to provide the banks' promises to early consumers in state L, while there is a bank run in state H. For the market to be cleared at date 0, any asset cannot dominate the other. As $\lambda_H > \lambda_L$, we know that $P_L > P_H$. Therefore, $P_H < 1$ because if $P_H \geq 1$, then $P_L > P_H \geq 1$. Then long-term assets dominate short assets at date 0, and no one wants to hold short-term assets. Also, $1 < P_L \leq R$ because if $P_L \leq 1$, then $P_H < P_L \leq 1$. Then, short-term assets dominate long-term assets at date 0, and no one wants to hold long-term assets.

In state L, the supply of liquidity comes from the safe bank which supplies the excess liquidity $(y^S - \lambda_L d^S)$ to the risky banks, which lacks in liquidity and demand $(\lambda_L d^R - y^R)$to satisfy their early customers. In this case, the supply and demand of liquidity are, thus, inelastic. In state H, the supply of liquidity also comes from the safe banks who use the excess liquidity, $(y^S - \lambda_H d^S)$, to buy the long-term assets liquidated by risky banks at 'fire sale' prices. The total demand of liquidity, given that risky banks need to liquidate all long-term assets, is, thus, elastic and equal to $P_H(1 - \frac{y^R}{P_0})$. At date 0, the supply of liquidity coming from a safe bank, which only holds liquidity in the economy, is equal to $(1 - y^S)$, and the demand for liquidity coming from the risky bank, which only holds the long-term assets in the economy, is equal to y^R. The market-clearing conditions are described as follows:

$$\begin{aligned}
&\text{Market clearing in state L} && \rho(y^S - \lambda_L d^S) = (1-\rho)(\lambda_L d^R - y^R) \\
&\text{Market clearing in state H} && \rho(y^S - \lambda_H d^S) = (1-\rho)P_H(1 - \frac{y^R}{P_0}) \\
&\text{Market clearing at date0} && \rho(1-y^S) = (1-\rho)y^R
\end{aligned} \tag{2.12}$$

Consumers must be indifferent between safe banks and risky banks in order

to deposit their endowment into both types of banks. The consumers' indifference condition is:

$$\begin{aligned}&\pi(\lambda_H U(C_{1H}^S) + (1-\lambda_H)U(C_{2H}^S)) + (1-\pi)(\lambda_L U(C_{1L}^S) + (1-\lambda_L)U(C_{2L}^S))\\ &= \pi(\lambda_H U(C_{1H}^R) + (1-\lambda_H)U(C_{2H}^R)) + (1-\pi)(\lambda_L U(C_{1L}^R) + (1-\lambda_L)U(C_{2L}^R))\end{aligned} \tag{2.13}$$

2.5.4 Equilibrium without run

The private decision of a safe bank is described by the optimal amount of short-term assets holding, y^S, as in Eq(2.14) and deposit contracts, characterised by d^S, as in Eq(2.15).

$$\begin{aligned}y^S = &\frac{(P_H(P_L-(1-\pi)P_0)-\pi P_L P_0)((P_L-P_0)\lambda_H+(P_0-P_H)\lambda_L)}{(P_L-P_0)(P_0-P_H)(P_0(\lambda_H-\lambda_L)-P_L\lambda_H+p_H\lambda_L)}\\ &+\frac{(P_H-P_L)P_0\lambda_H\lambda_L(P_H\pi-P_0+P_L(1-\pi))}{(P_L-P_0)(P_0-P_H)(P_0(\lambda_H-\lambda_L)-P_L\lambda_H+p_H\lambda_L)}\end{aligned} \tag{2.14}$$

$$d^S = \frac{(P_H-P_L)(\pi(\lambda_H-\lambda_L)+\lambda_L)}{P_0(\lambda_H-\lambda_L)-P_L\lambda_H+P_H\lambda_L} \tag{2.15}$$

The private decision of a risky bank is described by the optimal amount of short-term assets holding, y^R, as in Eq(2.16) and deposit contracts, characterised by d^R, as in Eq(2.17).

$$\begin{aligned}y^R = &\frac{(\pi P_H P_L+(P_H(1-\pi)-\pi)P_L P_0-P_H P_0^2(1-\pi))((P_L-P_0)\lambda_H-(P_0-P_H)\lambda_L)}{(P_L-P_0)(P_0(\lambda_H-\lambda_L)-P_L\lambda_H+p_H\lambda_L)}\\ &+\frac{(P_H-P_L)P_0\lambda_H\lambda_L(P_H\pi-P_0+P_L(1-\pi))}{(P_L-P_0)(P_0(\lambda_H-\lambda_L)-P_L\lambda_H+p_H\lambda_L)}\end{aligned} \tag{2.16}$$

$$d^R = \frac{(P_H P_L-(1-P_H)P_L P_0+P_H P_0^2(1-\pi))(\pi(\lambda_H-\lambda_L)+\lambda_L)}{P_0(P_0(\lambda_H-\lambda_L)-P_L\lambda_H+P_H\lambda_L)} \tag{2.17}$$

To describe the system of equations that characterise equilibrium, the equilibrium without bank run is determined by the market-clearing condition, as in eq(2.11), and the consumer-indifference condition, as in eq(2.13). With four unknowns and four equations, we can solve for (P_H, P_L, P_0, ρ). We can then determine (y^S, d^S, Y^R, d^R) accordingly.

In this equilibrium, safe banks and risky banks make identical choices (y and d). Safe and risky banks hold just enough liquidity to satisfy their depositors in state

H; that is, $y = \lambda_H d$. Therefore, the asset market at date 1 is not essential. However, the role of price here is to ensure the optimal holding of y and d. The opportunity costs of the commitment to "no run occur in state L, when both banks need to hold excess liquidity and carry forward using storage technology from date 1 to date 2. So, they need to forgo higher returns from holding more long-term assets in state L, resulting in low expected consumption. This is the source of inefficiency in ensuring 'no run' in equilibrium. The market at date 0 is for risk sharing given restrictions to access the technologies.

The comparative statics, shown in *Appendix B.3*, suggest that the aggregate liquidity in the economy, as represented by ρ, increases with π, λ_H, and λ_L, and decreases with R. This is similar to the motives that underpin a bank's optimal holding of liquidity (y) in equilibrium since both types of banks make identical choices. These results are obvious, as the increase in the fraction of early consumers, represented by λ, encourages a bank to hold more liquidity to satisfy early consumers. Also, an increase in the probability of state H means a higher chance of liquidity shock, encouraging a bank to hold more liquidity. However, an increase in the return of long-term assets discourages a bank from maintaining liquidity, since the long-term assets become more attractive.

The optimal deposit contract (d) increases with π and λ_L, and decreases with R and λ_H. As a higher probability of state H causes a bank to hold more liquidity, this also results in offering an early consumer a higher deposit contract because the liquidity will be used up in state H. In addition, the larger the liquidity shock (e.g. increase in λ_H and decrease in λ_L), the lower the deposit contract offered to an early consumer. An increase in the long-term assets' return results in a deposit contract that holds less appeal to an early consumer because long-term assets become more attractive.

For the consumption bundle offered to a consumer, an increase in R makes the long-term assets more attractive, and a bank will hold less liquidity and offer less promise of a payoff to an early consumer. This will result in lower consumption among early consumers, and greater consumption for late consumers. An increase in the probability of state H, which means a higher chance of liquidity shock, encourages a bank to hold more liquidity, and to offer deposit contracts with more promise of a payoff to early consumers, resulting in greater consumption among early consumers and less consumption among late consumers. The larger the liquidity shock, the lower the consumption of both an early consumer, and a late consumer in state L, and the higher the consumption of a late consumer in state H.

2.5.5 Equilibrium with run

The private decision of a safe bank is similar to the case when there is no run, and is described by the optimal amount of short-term assets holding, y^S, as in Eq(2.14), and deposit contract, characterised by d^S, as in Eq(2.15).

The private decision of a risky bank when the incentive-compatibility constraint Eq(2.9) is violated, could be described by the optimal amount of short-term assets holding, y^R, as in Eq(2.18), and a deposit contract, characterised by d^R, as in eq(2.19).

$$y^R = \frac{\pi P_L}{P_L - P_0} - \frac{(1-\pi)P_H P_0}{P_0 - P_H} \tag{2.18}$$

$$d^R = \frac{(1-\pi)((1+P_H)P_L P_0 - P_H P_L - P_H P_0^2)}{P_0(P_0 - P_H)} \tag{2.19}$$

To describe the system of equations that characterise equilibrium, the equilibrium with bank run is determined by the market-clearing condition, as in Eq(2.12), and the consumer-indifference condition, as in Eq(2.13). With four unknowns and four equations, we can solve for (P_H, P_L, P_0, ρ). We can then determine (y^S, d^S, y^R, d^R) accordingly.

In this equilibrium, a risky bank offers deposit contracts with greater amount of promise a payoff to early consumers as compared to when there is no bank run, $d^R > d^S$. Aggregate liquidity in the economy is just enough to satisfy early customers in state L, while it is less than the aggregate demand in state H, resulting in the run on risky banks. In this equilibrium, safe banks need to hold excess liquidity even in state H in order to buy liquidated long-term assets at 'fire sale' prices, when the long-term assets' return in state is greater than R since $P_H < 1$. This is to compensate for the opportunity in state L, where a safe bank needs to hold excess liquidity to rescue the risky banks when $1 < P_L < R$. Thus, there is no holding of short assets from date 1 to date 2. The source of inefficiencies, in this case, comes from the default cost associated with a risky bank which needs to liquidate all its long-term assets early. However, there is an efficiency gain by increasing the contingency of the contract.

The comparative statics shown in *Appendix B.4* suggest that the aggregate liquidity in the economy, as represented by ρ, increases with π, λ_H, λ_L, and R. In contrast to the equilibrium without run, aggregate liquidity in this economy increases with R. Although the increase in long-term assets' return provides more incentive to hold less liquidity, an equilibrium requires greater holding of liquidity for better outcomes, as shown in *Appendix B.4*. The reason behind this is that the

increase in long-term assets' return results in a higher opportunity cost of a safe bank holding excess liquidity in state L. As a result, the safe bank needs to hold even more liquidity to compensate for such opportunity cost by buying long-term assets at 'fire sale' prices in state H. Therefore, a safe bank holds more liquidity, y^S, at the end of date 0, while the risky bank holds less liquidity, y^R. In terms of the promise of optimal deposit contract, R doesn't affect the decision of a safe bank, but encourages the risky bank to offer higher consumption for early consumers, who want to withdraw earlier.

The equilibrium prices, P_H and P_0, increase with R, π, and λ_L, while they decrease with λ_H. P_L however increases with R, π, and λ_H, while it decreases with λ_L.

2.6 Numerical results and discussions

While in Diamond and Dybvig [1983], a bank run emerges as the result of a coordination failure among depositors resulting in inefficient equilibrium, this study, however, shows that a bank run occurs as a result of endogenous liquidation price, as shown earlier in Allen and Gale [2004]. In particular, late depositors will rationally run on risky banks, when the long-term asset price is so low that the incentive compatibility constraints are violated, and risky banks cannot meet their obligation to pay the promise according to a deposit contract. In this sense, price-taking bank defaults because of exogenous aggregate shocks on demand for liquidity, not because of coordination failure which is unanticipated at date 0. When risky banks default, they need to liquidate their holdings of long-term assets, which results in low asset prices. This, in turn, causes a group of banks to default.

2.6.1 When is equilibrium with run better?

The numerical examples in Table 2.1 suggest that bank run equilibrium could provide better outcome than equilibrium without run. Given the contract is incomplete, default can improve welfare by increasing contract contingency as shown in Zame [1993]. In particular, the costs of insuring no run in holding excess aggregate liquidity dominate the cost of consumption smoothing in holding more risky assets. Depending on the parameter, the equilibrium without run can also provide a better outcome as indicated in Table 2.2.

Intuitively, the comparative statics shown in Figure 2.5 indicate that equilibrium with bank run is likely to provide greater outcome when the probability of large shock, π, are low, and when the payoff of long-term assets, R, is high. In

	Planner	Banking Equilibrium	
		Without Run	With Run
$E[U]^S$	0.089	0.075	0.081
$E[U]^R$		0.075	0.081
P_L		2.000	1.503
P_H		0.465	0.498
P_0		1.000	1.019
ρ		0.885	0.872
Aggregate Holding of y	0.868	0.885	0.872
y^S		0.885	0.933
y^R		0.885	0.455
d^S		0.983	0.992
d^R		0.983	1.259

Table 2.1: Better outcomes in equilibrium with bank run
Given: $\pi = 0.35;\ \lambda_H = 0.9;\ \lambda_L = 0.85;\ R = 2;\ U(C) = ln(C)$

	Planner	Banking Equilibrium	
		Without Run	With Run
$E[U]^S$	0.160	0.145	0.139
$E[U]^R$		0.145	0.139
P_L		2.000	1.551
P_H		0.490	0.545
P_0		1.000	1.077
ρ		0.784	0.786
Aggregate Holding of y	0.768	0.784	0.786
y^S		0.784	0.883
y^R		0.784	0.428
d^S		0.980	0.992
d^R		0.980	1.251

Table 2.2: Better outcomes in equilibrium without run
Given: $\pi = 0.35;\ \lambda_H = 0.8;\ \lambda_L = 0.75;\ R = 2;\ U(C) = ln(C)$

addition, better outcomes are more likely when the probability of being an early consumer in state L, λ_L, are low, while the probability of being early consumer in state H, λ_H, are high. This indicates that with larger shocks, bank run equilibrium becomes more preferable.

2.6.2 Identifying the source of inefficiencies

Pecuniary externalities play a role in the inefficiencies in this model. However, as mentioned in Chapter 1, according to Dávila and Korinek [2017], there are two types of externalities: 'distributive externalities' as a result of an inequalised marginal rate

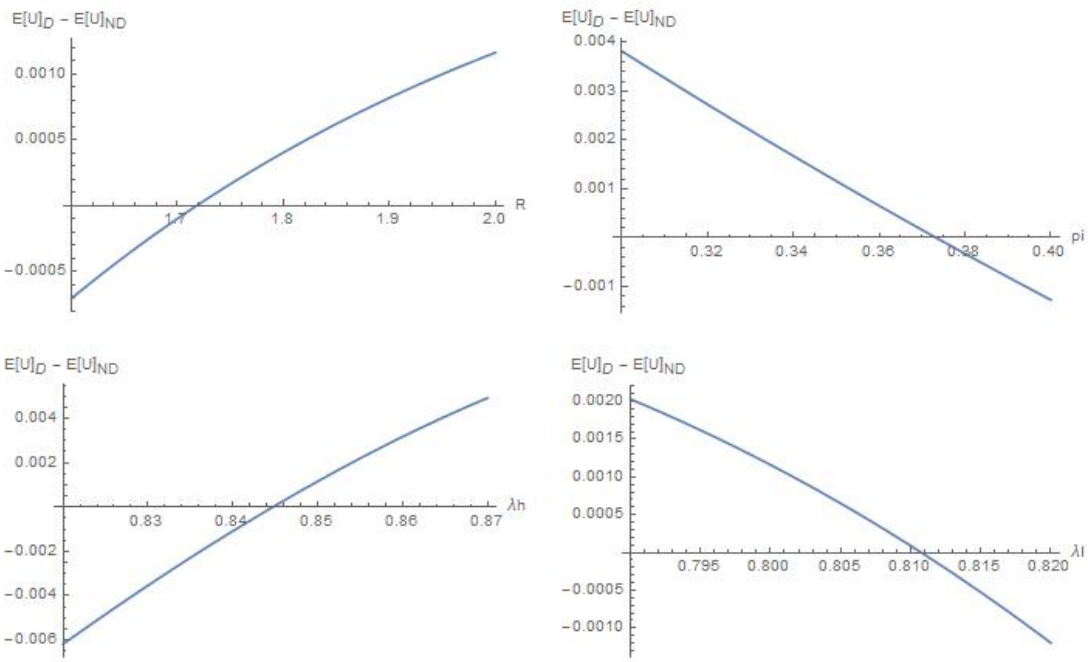


Figure 2.5: Welfare comparisons, equilibrium with bank run over with no run
Given: $R = 2; \pi = 0.35;\ \lambda_H = 0.85;\ \lambda_L = 0.8;\ U(C) = ln(C)$

of substitution (MRS), and 'price-dependent constraint externalities' as a result of the binding price-dependent constraint. In this model, the first inefficiency works through 'distributive externalities' as a result of incompleteness of the market. When there is no market for banks to achieve efficient risk sharing in the absence of state-contingent securities, the market at date 0 fails to reveal how much investors would be willing to buy/sell assets contingent on knowing their types. As a result, the market provision of liquidity in the initial portfolio choices of banks is inefficient. In fact, the numerical results, as shown in Table 2.1 and Table 2.2, indicate *towards 'excessive provision of liquidity'* in market equilibrium.

More importantly, this paper shows that in equilibrium with bank run, when the asset price is so low that the incentive-compatible constraints are violated, there arises the second inefficiency: 'price-dependent constraint externalities', working through binding incentive-compatible constraints, emerge when risky banks fail to internalise the effect of price change due to collective actions on the liquidation of long-term assets given the default. As a result, the price of long-term assets at date 0 is greater than 1. Although safe and risky banks are restricted to different technologies, in the absence of price-dependent constraint externalities, having access to the technologies at date 0 should result in the price of long-term assets at date 0

being equal to 1, as we can see in equilibrium without run. As a robustness check, we can vary the values of the parameters and illustrate that the price of long-term assets at date 0 is greater than 1; this is shown in *Appendix B.4* Figure B.5.[3]

In addition, excessive holdings of liquidity due to distributive externalities usually result in more fluctuation of asset prices because the greater cost of having excess liquidity in state L necessitates higher compensation for safe banks, which buy long-term assets at 'fire sale' prices in state H. In other words, the first type of externalities leads to lower prices in state H, which could introduce the second type of externalities, when the incentive-compatibility constraints are violated. However, the price of assets in state H in equilibrium with run is greater than in equilibrium without run because the increasing contingency of the contract by run helps to improve welfare by lessening the extent of over provision of liquidity.

[3]Other comparative statics of banking equilibrium are provided in *Appendix B.3 and B.4*.

Chapter 3

Subprime Assets and Financial Crisis: Theory, Policy and the Law[1]

Joint With: Marcus Miller and Lei Zhang[2]

3.1 Introduction

That the period of macroeconomic stability known as the Great Moderation should have ended in a financial cataclysm was a nasty shock – especially for those who believed in the inherent efficiency of financial markets! But how to account for the fact that the spark for the crisis came not from emerging markets but from within the United States itself, where monetary affairs had, for many years, been in the hands of Mr Greenspan, doyen of central bankers?

Earlier financial shocks and external factors doubtless played a role, with US interest rates being cut after the high-tech bubble collapsed in 2000; and then kept low as funds flowed in from the 'savings glut' in East Asia. But here we focus on factors specific to US housing finance to see how subprime mortgage lending, sponsored by misguided policy and aided by febrile financial innovation, could undermine the integrity of the US financial system.

The objective in revisiting these issues is not to allocate blame; rather to see

[1]Acknowledgements: for valuable comments we are grateful to George Akerlof, Robert Akerlof, Sacha Becker, Huberto Ennis, Peter Hammond, Stephania Paredes-Fuentes, Herakles Polemarchakis, Alistair Milne, Peter Spencer and Yifan Zhang. An earlier version of this paper is available as CEPR DP No 11533.

[2]Marcus Miller: Department of Economics, University of Warwick; Lei Zhang: The School of Economics, Sichuan University

how a repeat may be avoided. If the problem was essentially financial panic, for example, as Gorton [2010] and others maintain, the remedy would be ample provision of liquidity. But if, as Mian and Sufi [2015] contend, policy-makers were using the provision of cheap credit as an elixir to cure growing income inequality and shadow banks 'joined the party' with sophisticated products that would only work when house prices were rising, then the analysis and policy response needs to go much deeper. To set the scene, we begin with some institutional and policy background.

Housing finance: getting onto the housing ladder

The development of US banking in the late twentieth century, according to Calomiris and Haber [2014], involved a 'bargain' between banks and society: banks were permitted to merge and grow so long as they promoted home ownership for low-income mortgage applicants.

> "Once branching limits were removed, bankers had ambitious plans for mergers. Their plans were, however, subject to a political constraint: they needed to be judged good citizens of the communities they served in order to gain approval from the Federal Reserve Board. Good citizenship came to be defined as being in compliance with the 1997 Community Reinvestment Act ...For activist groups seeking to direct credit to their memberships and constituencies, the good-citizenship merger criterion was a powerful lever in negotiations with merging banks. The bankers and the activists forged a coalition that consolidated the American banking industry into a set of megabanks that were too big to fail." [Calomiris and Haber, 2014, p. 208].

Such a deal had the apparent advantage of helping to offset the stagnation of median incomes and growing inequality as earnings at the top of the income distribution raced ahead. Instead of taxes and subsidies to redistribute income, the idea was that those on lower incomes would borrow to get on the housing ladder so – with time and house price appreciation – they could extract equity to increase consumption.

As the authors go on to point out, however:

> "Other partners had to be drawn into the coalition in order to make it stable. Banks would not make unlimited commitments to their activist partners: Community Reinvestment Act loans implied higher levels of risk for the bank

than traditional mortgage loans. Thus, under pressure from activist groups, Congress began to place regulatory mandates on government-sponsored enterprises (GSEs) that purchased and securitised mortgages... Fanny Mae and Freddie Mac, in particular, were required to repurchase mortgage loans made to targeted groups (i.e. individuals who had low incomes or lived in urban locations that were defined as underserved). In order to meet these targets, Fannie and Freddie had to weaken their underwriting standards." [Calomiris and Haber, 2014, p. 209].

Under the Clinton and Bush administrations, the mandate on GSEs for low income housing steadily increased, from 42% of assets in 1995 to 56% in 2004. Indeed, it has been estimated that:

"by 2008, the mortgage giants, the FHA and various other government programs were exposed to about $2.7trillion in subprime and Alt-A loans, approximately 59% of total loans in these categories. ... As money from the government-sponsored agencies flooded into financing or supporting low income housing, the private sector joined the party. ... Unfortunately, the private sector, aided and abetted by agency money, converted the good intentions behind the affordable housing mandate and the push towards an ownership society into a financial disaster."
[Rajan, 2011, p. 38-39]

Despite substantial political pressure to extend home ownership by poorer households, the subprime share of mortgage market remained around 10% until 2003. With the development of private label securitisation (PLS), however, the subprime experiment in 'dynamic credit enhancement' for low-income borrowers accelerated sharply. As can be seen from Figure 3.1, the share of subprime mortgages rapidly doubled to over 20% of all mortgage originations in 2006. But when the house price bubble burst, the share of subprime mortgages fell precipitously, with virtually none being securitised in 2008.

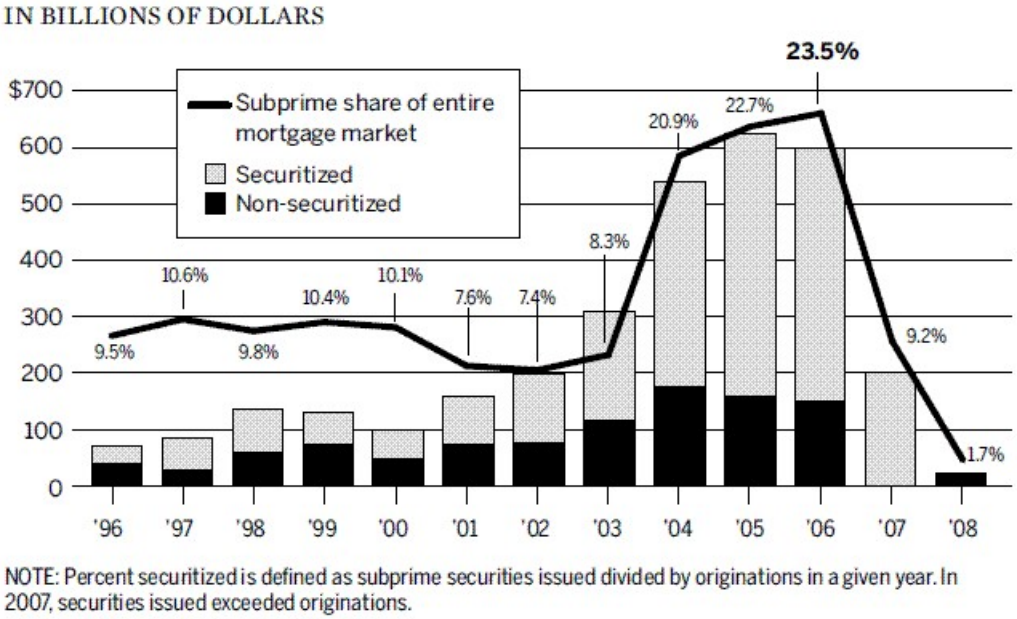


Figure 3.1: U.S. Subprime mortgage originations.
Source: Financial Crisis Inquiry Commission Report, p. 70, Figure 5.2.

Could policy makers and regulators not have stopped this vast expansion of subprime lending - by changing the mandates, for example; or by imposing higher prudential capital requirements on such loans? In the view of [Calomiris and Haber, 2014, p. 281]:

> "they could ... but they chose not to do so. Instead, regulators stood by and watched: in essence they subcontracted regulation of banking to private firms that sold ratings and whose incentives were therefore aligned with those issuers and purchasers, who wanted to have inflated ratings."

If this is an accurate assessment of the policy and regulatory framework, then the onset of financial crisis seems as inevitable as the fate of Santiago Nasar in Marquez's Chronicle of a death foretold.

Brief overview of some relevant literature

In a prescient paper delivered at the Jackson Hole Conference on The Greenspan Era in 2005, Raghuram Rajan raised the issue of whether financial innovation was making the world a riskier place. The focus of his concern was on leverage and

asymmetric information in financial intermediation, and how distorted incentives could lead to excessive risk-taking. Though Rajan's concern was met with general scepticism from other delegates, it was supported by Hyun Shin, on the ground that, even with common knowledge, high leverage could lead to instability on account of 'pecuniary externalities'. He used the internal dynamics of the Millenium Bridge to illustrate how shocks can be greatly amplified in financial markets - 'the supreme example of an environment where individuals react to what's happening around them, and where individuals' actions affect the outcomes themselves'. In another influential paper delivered at Jackson Hole soon after the crisis broke,Gorton [2010] argued that the lack of transparency in financial innovation could trigger financial panic in the form of a bank run.

Fostel and Geanakoplos [2012] also stressed the role of financial development; but, in marked contrast to the bank-focussed perspectives just discussed, theirs is a general equilibrium approach. They stress the role of heterogeneous beliefs as the driver for leverage as optimists borrow from pessimists; and how the sequential introduction of financial innovations is, in and of itself, enough to cause boom and bust.[3] Another fast-growing branch of the literature, lying between detailed partial equilibrium models of banking and 'institution free' general equilibrium, focuses on adding 'financial frictions' to DSGE models cast in the Gali/Woodford tradition of modern macroeconomics. We make no attempt to analyse these contributions here. For a good illustration of the DSGE approach, with a helpful summary of other papers in this burgeoning field, the reader may be referred to Coimbra and Rey [2017]; and a concise version of the general equilibrium approach is provided in Miller et al. [2016].

Goodhart et al. [2010] model a housing and mortgage crisis by incorporating heterogenous financial intermediations and households. By allowing endogenous credit and default, which are key features of many crises, the simulation results in this study show a number of realistic outcomes, including: that the crisis becomes more severe if banks are risk-loving, that government support helps stabilise the economy, and that prices and trade quantities are reduced when money is tight. The paper also highlights the role of the interbank market in spreading individual default.

However, the majority of papers on the financial crisis take a partial equilibrium perspective – with a focus on the institutional aspects of 'shadow banking'

[3]The innovation that can trigger collapse is the availability of 'naked' Credit Default Swaps (CDS) contracts which allow for insurance against failure, so non-asset holders can 'short' investment in risk assets. So-called 'naked' CDS contracts do not require ownership of the assets being 'insured'.

in particular. As is typical for banking models, there is a split between those, like Gary Gorton, who emphasises the role of shocks to liquidity in a setting where fundamentals are essentially well-founded, and those who focus on structural flaws in incentives and/or regulatory structure capable of precipitating widespread insolvency due to excess risk-taking.

Prudential regulation to check excess risk-taking by highly-leveraged institutions (HLIs) had been widely discussed well before the subprime crisis, as Goodhart [2011] testifies. A key issue in debate was whether the value at risk (VaR) rules adopted in Basel II to check risk-taking by individual banks would be sufficient to guarantee systemic stability; or whether it could be flawed for ignoring externalities. Danielsson et al. [2001], argued that balance sheet rules, designed to ensure prudential behaviour at the level of the individual bank, could lead to systemic instability when common, 'macroeconomic' shocks are amplified by 'pecuniary externalities' in the form of asset price changes which affect bank equity in pro-cyclical fashion.

A masterly survey of the literature on the problems posed by such externalities is provided by Brunnermeier et al. [2012]. They leave on one side, however, the issue of distorted incentives due to asymmetric information analysed earlier by Holmstrom and Tirole [1997] and Hellmann et al. [2000]. How financial innovation could exacerbate these issues, as high-lighted by Rajan [2006], was emphasised by Foster and Young [2010] – who showed how financial derivatives could be used by fund-managers of average ability to mimic the performance of star traders, taking on tail risks to do so.[4] In the context described above, where monitoring of asset quality had been delegated to unregulated, private-enterprise Credit Rating agencies (CRAs), Akerlof and Shiller [2015] argued that investment banks had an alternative strategy for making their investments appear superior: getting them rated as AAA by compliant agencies. As with mimicry, however, getting high returns involved taking on significant risk.

Structure of the paper

These topics – externalities, distorted incentives and creditor panic – are analysed in some detail in Sections 3.2 and 3.3, considering in particular whether each one could itself be sufficient to cause banking crisis.

The first threat of insolvency examined in Section 3.2 involves the role of externalities. We focus in particular on the Investment Banking model of Shin [2010] which emphasises how 'pecuniary externalities' can amplify unexpected shocks to the

[4]This strategy offers the prospect of high returns for some time followed by substantial losses as tail risks finally materialise.

quality of investments they hold.[5] To check on the robustness of US-style shadow banking in the face of shocks, we ask: could these externalities prove sufficiently strong that the simple reversal of 'good news' might lead to widespread insolvency and banking collapse?

The second threat of insolvency examined involves the distorted incentives for risk-taking in HLIs, particularly after the switch from partnerships to limited liability in US Investment Banking, as discussed in Akerlof and Shiller [2015]. The focus here being on the role of asymmetric information in the marketing of and investment in highly risky assets, we apply the adverse selection approach of Akerlof [1970] to the marketing of subprime assets. Relaxing the 'rational expectations' constraint imposed in that paper allows for risks to be concealed by inflated ratings issued by CRAs who are 'mining their reputation' to secure the fees on offer for rating subprime loans; and to lead to financial collapse when it is discovered that these loans were worth a lot less than previously thought.

We note that these two threats are in fact complementary: just as successful exploitation of asymmetric information to take on excess risk underpinned the credit boom, so revelation of the risks sub-prime lending really involved was the downside shock that brought on the crisis.

In Section 3.3, the 'confidence crisis' view is discussed and we ask: was the rise in the cost of insuring subprime assets a matter of mindless panic as suggested by Gorton? Or was it not due to a realisation of faulty fundamentals?

But what if these various perspectives are high-lighting different aspects of a complex reality? This may recall the ancient Hindu parable of the blind men and the elephant, where the former – each guided only by touching a different part of the animal, be it a tusk, the tail, an ear, or a leg – give a series of correct but partial characterisations of the noble beast. The conclusion in the Rigveda, cited above, namely that Reality is one, though wise men speak of it variously, tempts one to ask: should these seemingly conflicting accounts not be combined? For an answer we turn not to theory, nor to econometric tests of theory, but to the evidence of law courts and the actions of policy-makers in the Fed and Treasury. What did the extraordinary policy actions taken by these agencies reveal about the nature of the crisis? Did those charged to dispense justice find evidence of misbehaviour sufficient to prosecute the players involved?

To balance these three perspectives – and to see whether in practice they proved complementary – Section 3.4 summarises key official policy actions taken

[5]Such externalities are sometimes discussed under the heading of contagion. See for example Allen and Carletti [2012].

in response to the crisis; and subsequent findings in the law courts against both CRAs and Investment Banks.[6] In some versions of the parable a sighted observer appears to reconcile the various conflicting perspectives. In this sprit, the view-with-hindsight of the current chair of the Federal Reserve, as expressed at Jackson Hole 2017, is also cited.

After a brief account of possible steps to increase risk-sharing in housing finance, section 3.5 concludes.

3.2 The risk of insolvency - two views

3.2.1 Insolvency with Value-at-Risk: common shocks and 'pecuniary externalities'

In this section, the Investment Banking model of Shin [2010] and Adrian and Shin [2010] is used to examine the contention[7] that VaR based regulation is no guarantee of systemic stability. We find by simulation that the representative Investment Bank could become insolvent when a significant upgrade in risky asset quality is followed by a subsequent reversal.

In what, for convenience, will be referred to simply as the Shin model, there are two groups of investors; (1) risk averse agents with mean-variance preferences, who do not use leverage to finance investments such as pension funds and mutual funds; and (2) risk neutral investors, who can finance investments with leverage subject to a Value-at-Risk (VaR) constraint. For present purposes, we will treat the latter as homogenous and highly-leveraged investment banks. But, in reality, such active leveraged investors include hedge funds and foreign banks, as well as U.S. investment banks.

There are two assets: (1) a riskless bond with its rate of return normalised to 0; and (2) a risky asset with random payoff Q, uniformly distributed over $[q-z, q+z]$ where $q > 0$, with moments denoted by:

$$E[Q] = q$$
$$Var[Q] = \frac{z^2}{3}$$

Both types of investors are endowed with initial equity equal to e. Investors' portfolio payoff (end of period wealth) is $W = Qy + (e - py)$, where y represents

[6] Why the legal settlements have taken the form of 'deferred prosecution agreements' with the companies involved, rather than the criminal prosecution of high-level individuals, is also discussed.

[7] Made by Shin and others in "An academic response to Basel II", Danielsson et al. [2001].

quantity of the risky asset holdings and p is the price of the risky asset.

Passive investors

As they do not borrow to finance their investments, risk averse investors are categorised as passive. Their 'mean-variance' preferences are described by:

$$U(W) = E(W) - \frac{1}{2\tau}\sigma_W^2$$

where τ represents risk tolerance and, since their portfolios comprise of only riskless bonds and risky asset, portfolio variance is $\sigma_W^2 = \frac{y^2 z^2}{3}$. Risk averse investor's optimisation thus becomes:

$$\max_y \quad qy + (e - py) - \frac{y^2 z^2}{6\tau}$$

The demand function of passive investors becomes:

$$y_p = \begin{cases} \frac{3\tau}{Z^2}(q-p) & if \quad q > p \\ 0 & if \quad otherwise \end{cases} \tag{3.1}$$

Note that because of the assumption on mean-variance preferences, the demand for risky asset by the passive investors is independent of their wealth.[8]

Active investors: Investment Banks

Risk neutral investors are active as they use leverage to finance their investments, subject to a VaR constraint. Specifically, investment banks' optimisation is described as:

$$\max_y \quad E(W)$$

$$\text{s.t.} \quad VaR = (p - (q - z)) \leq e$$

where $E(W) = (q-p)y + e$ and the VaR constraint implies the borrowing is no greater than the worst realised payoff on the risky asset, $py - e \leq (q-z)y$.

[In a more general case where the distribution of Q has unbounded support, the VaR constraint becomes probabilistic: i.e., $Prob(VaR = (p - Q)y \geq e) \leq \alpha$, where α is the probability of losing the entire equity. Under this modified VaR con-

[8] In Shin's model, the specific formulation of VaR implies no default ex post as the distribution of Q has bounded support. However, if the support of Q is not bounded, ex post default is possible; so the wealth of the passive investors may be affected. But with these mean-variance preferences, the determination of equilibrium asset prices is not affected.

straint and the limited liability, the expected payoffs of the active investors become: $E(W) = (1-\alpha)[(q-p)y+e]$. However, the active investors can purchase Credit Default Swaps(CDS) to insure against the tail risks associated with losses beyond the VaR and so avoid bankruptcy. Let the cost of CDS be β and the CDS is used to insure against Q falls below $q-z$, then the VaR constraint in the text is restored, and the expected payoffs of active investors now become: $E(W) = (q-p)y+e-\beta$. So in the presence of CDS, the formulation used in the text can also apply to the case even if the support of Q is unbounded.]

Since $E(W)$ is linear in y, then for $q > p$, so long as the VaR constraint is binding, the demand for risky asset by investment banks becomes:

$$y_A = \begin{cases} \frac{e}{p-(q-z)} & if \quad > p \\ 0 & if \quad otherwise \end{cases} \tag{3.2}$$

Market clearing

For $q > p$ and assuming that aggregate supply of risky assets is fixed and equal to 1, the market clearing condition $y_P + y_A = 1$ gives the equilibrium price:

$$p = q - \frac{z}{2}\left[\frac{z}{3\tau} + 1 - \sqrt{(\frac{z}{3\tau}-1)^2 + \frac{4e}{3\tau}}\right] \tag{3.3}$$

For a given supply of risk assets (normalised to one) on the horizontal axis, various market equilibria are illustrated in Figure 3.2, calibrated broadly using figures gleaned from Shin [2010], as shown in Table 3.1, using the formulae provided in *Appendix C.1.* The construction is that the demand by passive investors, measured from the right-hand axis, lies below the mean, with a slope that reflects their degree of risk aversion; while the demand curve for active investors is measured from the left hand axis. The kink reflects their initial equity e and the downward slope indicates, not risk aversion, but the effect of the VaR rule: a fall in price allows more assets to be held as there is less risk per asset, measured as $p-(q-z)$, to be covered by their equity. Equilibrium is where total demand matches supply. The outcome shown in the middle of the diagram is labelled L to indicate the Low quality risk assets available; that on the right, labelled H to indicate a much higher quality, shows the considerable expansion of holdings by investment banks triggered thereby; the outcome on the left, where Investment Banks go out of business, is labelled I for insolvency.

The low quality of assets available at L refers to a downside risk of 0.13 relative to an expected payoff of 1.06 which gives the minimum payoff of $q-z =$

0.93 indicated by the dashed red line near the foot of the figure. The demand schedule from Investment banks, subject to VaR with equity of 0.024, has a kink at $\frac{e}{z} = 0.024/0.13 = 0.18$ and descends as a rectangular hyperbola towards 0.93 as its lower asymptote. It intersects the demand from passive investors at a price just above unity, giving investment banks a market share of about 30%.

What if, for reasons to be discussed below, there is an unanticipated increase in the quality of risk assets, known to all participants, which narrows the downside risk substantially to only 0.06, lifting the minimum payoff to $q-z = 1.0$ (as indicated by the dashed line near the middle of the Figure 3.2). The reduction of perceived risk will increase the demand by mean/variance investors, as shown by the clockwise rotation of their demand schedule. The demand from investment banks will increase for two reasons. First because, with lower downside risk per unit, the initial equity can cover the risk on a larger quantity of assets; and second because, with mark to market accounting, capital gains from the price rise for risky assets on their balance sheets will raise their equity value. The combined effect is a marked shift to the right in demand curve for Investment Banks operating under the VaR rules, as shown by the upper rectangular hyperbola in the figure.

Given the parameter values indicated, market-clearing equilibrium is at H, with the price lying very close to the top of the narrow ' band' of 6% between q and $q - z$, and with the lower downside risk fully covered by the higher equity. In this case, meant to represent pre-crisis boom, demand by the risk-neutral Investment banks, holding about two thirds of the risky assets with a leverage ratio of almost 20, has virtually eliminated the risk-premium $q - p$ on these assets.[9]

[9] In line with Crockett's dictum, that 'risk exposure is built up in the boom but is only manifest in the bust'.

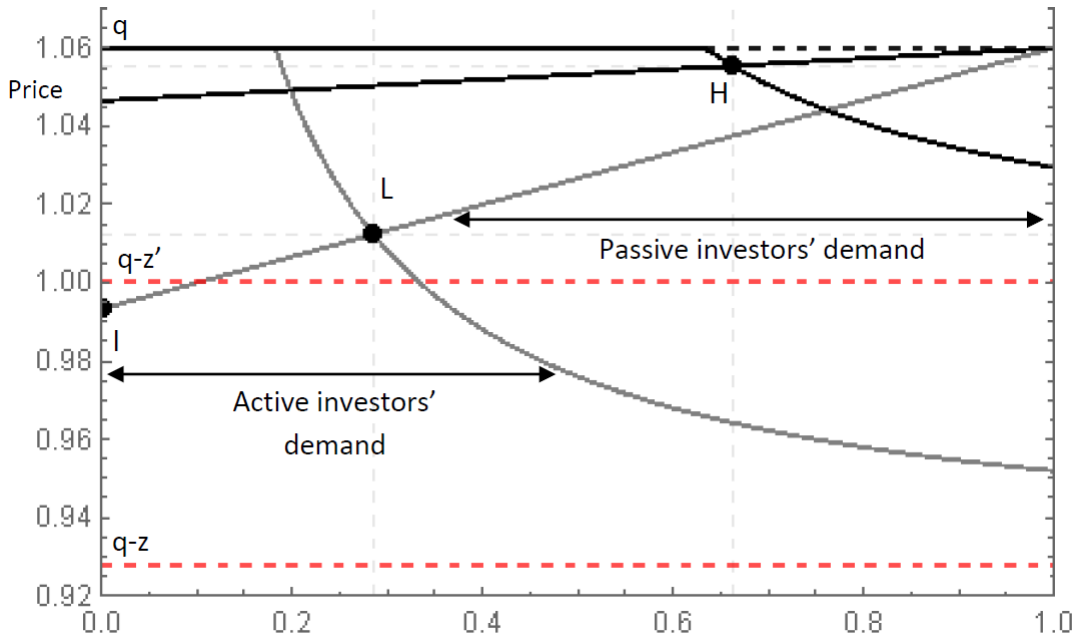


Figure 3.2: Market clearing price of risky assets: three cases

q	1.06
z	0.13
e	0.02
τ	0.09
Shocks	
dq	0.
dz	-0.07
Improvement in Asset Quality (dq-dz)	0.07

Table 3.1: Parameters used in calibration

	Initial Equilibrium	*Positive Shock*	*Shock Reversal*
Equilibrium			
Price	1.012	1.055	0.993
ya	0.285	0.663	0.
IBs Balance Sheet			
Asset	0.289	0.7	0.
Debt	0.265	0.663	0.
Equity	0.024	0.036	-0.005
Percent Change in Equity		51	-113.181
Leverage	11.978	19.212	0.

Table 3.2: Simulation results

Notwithstanding the absence of a risk premium, the boom equilibrium is distinctly fragile. Consider, for example, another possible shock - a write-down of the expected payoff, q, which ceteris paribus will shift both schedules vertically downwards. It might seem that a write-down of 6% is can be handled by the equity provisions made; but this is without taking account of the pecuniary externality – that the equity base of the banks is endogenous, and will fall as the price falls given 'mark to market' accounting. Allowing for the financial accelerator that this implies, which forces the highly-levered banks to contract their holdings as they sell risky assets into the market, it turns out that their initial equity can only stand a write-down of 4%.

Factors mentioned by Shin that might cause such parametric shocks to the mean return include – on the upside - a macroeconomic improvement lowering the probability that the borrowers would default on their loans; and a decline in the quality of mortgage borrowers as the market expands – on the downside. Miller and Zhang [2015] discuss the possibility that an initial probability upgrade may turn out to be mistaken; and Carlin and Soskice [2014] point explicitly to CRAs as a possible source of such mis-rating, with positive up-gradings later reversed. Danielsson et al. [2001], had earlier argued that "heavy reliance on CRAs is misguided as they have been shown to provide conflicting and inconsistent forecasts of individual clients' creditworthiness. They are unregulated and the quality of their risk estimates is largely unobservable." The conflict of interest that gave CRAs the incentive to issue erroneous ratings are considered in the next section; but meantime, as a test of robustness, consider a ratings upgrade that is later reversed.

A test of robustness: a reduction in downside risk, later reversed

To check the robustness of shadow banking in the Shin model, we first introduce a significant reduction in perceived risk; and then, when the equity base has expanded, a reversal of this 'good news'. The *unexpected increase in asset quality* that results in perceived risk reduction[10] could correspond to the highly favourable pre-crisis ratings given by rating agencies: while the *reversal of this good news* could reflect the subsequent sharp rise in the cost of insurance signalled by the ABX-HE indices, discussed below. The time-line of events is outlined below.

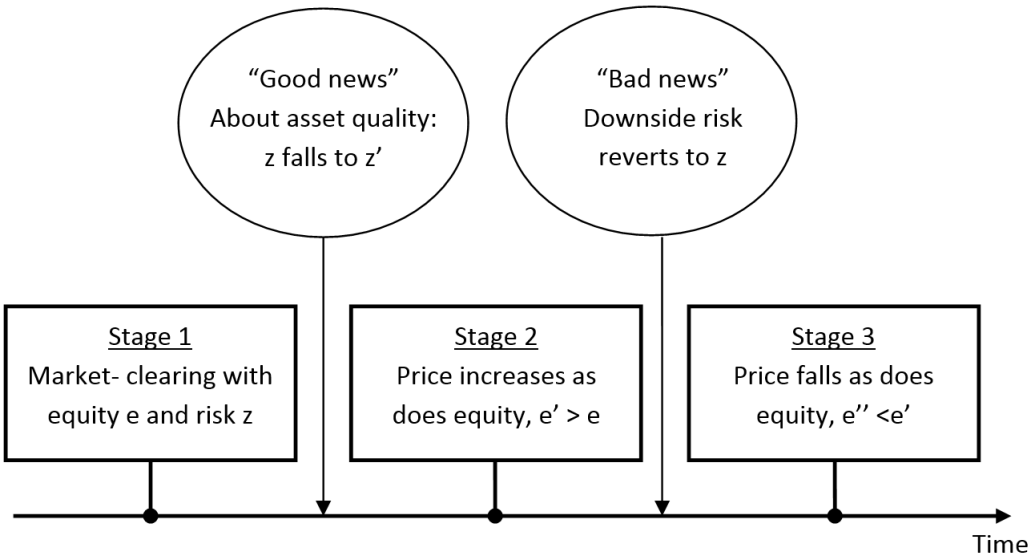


Figure 3.3: Timeline of events

In the first stage, with the downside risk parameter z and investment bank equity e, the equilibrium price p is determined along with y, the share of the risky asset held by active investors. This corresponds to point L in the figure. After markets have cleared on the assumption of an unchanging future distribution of asset returns, however, 'good news' on asset quality arrives: downside risk has fallen to $z' < z$. This unanticipated but welcome development leads to an increase in the price of risky assets; and the holdings of active investors also increase, as indicated by point H in Figure 3.2.

By marking assets to market at these higher prices, investors are effectively assuming no change in the future distribution of asset returns. They will, however, be disappointed, if 'bad news' arrives that downside risk z' has reverted back to

[10]Narrowing the 'downside', z, of risky asset's payoff, leaving the expected payoff unchanged at q as in Figure 3.2

what it was in the first stage, namely z.[11]

Given the original payoff distribution, but starting now from the higher equity base (e') achieved at stage 2 - with larger holdings financed by higher borrowing ($p'y' - e'$) - the question arises: how much will asset prices have to fall as active investors contract their balance sheets to meet the now-tighter VaR requirements; and will their equity be sufficient to take the hit?

Insolvency

Though news shocks that are reversed need not lead to insolvency, they can do so. The most obvious case is when the good news 'narrows the band' of downside risk enough to exclude the initial equilibrium price $p0$ (i.e. $q - z' > p_0$). If, for example, from an initial equilibrium at L in Figure 3.2 (with $z = 0.13$) the 'good news' was that the downside risk had fallen to $z = 0.06$ then a reversal will, of course, take equilibrium 'outside the band' involving losses larger than the maximum sustainable; so the entire equity of the banks will be wiped out by the 'bad news'.

In the context of a model with uniformly-distributed, bounded risk, this would be classified as a 'zero probability event', an outcome that takes prices lower than the worst the banks expect given the downside risk as perceived at H. Should it therefore be discounted? No, for two reasons. First because the design of the VaR regime was flawed in that externalities that could drive the system outside the bounds expected by individual banks were not checked by Pigouvian regulation; so the exaggerated impact of common shocks will be unexpected. Second because Shin's model may be expanded to allow for unbounded downside risk together with the availability of insurance to cover tail risk, as discussed above. In which case, moral hazard would play a role – and the insuring agency take a hit.

The example portrayed in Figure 3.2, however, is rather more subtle. It demonstrates that, even where a return to the initial equilibrium price would be sustainable (i.e. where $p_0 > q - z'$, a reversal of good news may trigger insolvency nonetheless. In the figure, reducing perceived down-side risk from 0.13 to 0.06 shifts equilibrium from L to H; and, as L remains within the 'narrow band' it might appear that a reversal is sustainable. But the asymmetry of capital gains (applied to initial holdings at L) and the capital losses (applied to expanded holdings at H) is sufficient to wipe out the equity of the active investors, leading to the equilibrium at point I where all risk assets are in the hands of mean-variance investors.

That such widespread insolvency was a possible outcome is supported by

[11] If the distribution of Q has unbounded support, we assume that the cost of CDS for insuring against the tail risks beyond q-z and q-z' are the same.

David Stockman's account of the financial crisis Stockman [2013, p. 543], where he asserts bluntly that, in the absence of Fed intervention: "Every single investment bank, including Goldman, Morgan Stanley, and the embedded hedge funds at J P Morgan, Citibank and Bank of America would have been rendered instantly insolvent and dismembered under court and FDIC protection."[12]

Catastrophic' behaviour

The tendency of the system to 'overshoot' its initial equilibrium (on the way down) when a quality upgrade is reversed depends on asset prices being 'marked to market'. This accounting practice makes endogenous the level of risk-taking by firms which keep their balance sheets at the limits set by VaR: but it operates asymmetrically. While the good news has a positive amplification effect applied to the initial level of equity at A, the rescinding of this good news has a negative amplification effect applied to the equity level at B, *boosted by the earlier good news.*[13]

Clearly accounting rules can have a marked effect on the dynamic response of the system to exogenous shocks; and, for large enough shocks, it appears that the price of risky assets can exhibit what Zeeman [1974] and Arnold [1984] refer to as 'catastrophic' behaviour – highly asymmetric responses to symmetric movements in exogenous forces. In the paper referenced, Christopher Zeeman sought to explain the gradual rise in equity prices in a boom followed by the sharp fall in the subsequent crash by the difference in behaviour between 'bulls and bears' – a psychological explanation that Arnold [1984] criticised as rather ad hoc. In the case we are discussing, however, the dynamics are derived explicitly from the 'rules of the game' – VaR rules sanctioned by Basel II to check moral hazard on the one hand; and market accounting regulations (FAS 157 in particular) designed to ensure fair asset pricing on the other.

The Shin model as analysed here appears to sustain the three charges made against BASEL II in Danielsson et al. [2001], namely that:

- VaR can destabilise and induce crashes when they would not otherwise occur
- Heavy reliance on CRAs is misguided
- Financial regulation is procyclical

[12]Something that he advocated, as he believed that the outcome would not have affected Main Street.

[13]If assets were not marked to market, however, the effects would be symmetric.

It should be noted that the results rely on the VaR constraint being continuously binding following the assumption that active investors are risk neutral. Indeed, the equity of active investors is just enough to cover losses in the case of the worst possible payoff of the risky asset. In other words, the active investors spend all they can afford to buy risky assets. If, however, they are willing to hold not only risky assets but also riskless bonds (i.e. the VaR constraint is not binding), the riskless bonds will be a buffer to withstand the asymmetric price adjustment given the reversal of shocks. Although the active investors can tolerate more if prices fall, insolvency is still possible if the shocks are large enough. The obvious weakness in the account so far provided is that it is driven by a sequence of unanticipated, exogenous quality shocks. Altering the common knowledge assumption, as in the next section, helps to make the sequence of shocks endogenous, however.

A different approach to making the shocks endogenous has been explored by Aymanns et al. [2016], in a Minsky-like extension of Shin's model. For them asset quality is judged, not by ratings, but from time-series estimates of downside risk made in a stochastic setting. As time moves on and the last crisis moves into the distant past, these assessments become progressively more rosy, and the system more risk prone – leading to another crisis. In fact, they derive an ever-repeated cycle of boom and bust which, they claim, is consistent with the operation of the BASEL II rules on prudential regulation.

3.2.2 Sourcing the shocks: asymmetric information with adverse selection

In Phishing for Phools, Akerlof and Shiller [2015] discuss how, with asymmetric information, markets may misallocate risk and resources; and claim they that structural flaws in US Investment Banking industry, and in the agencies that provided ratings for the products it dealt in, are a case in point. The switch from partnerships to limited liability, prior to the subprime crisis, gave the banks much greater willingness to take risks: but the degree of risk involved was grossly understated, as rating agencies – skilled in assessing repayment prospects for the debt of corporations and sovereigns – were paid by the banks to give favourable ratings to complex financial products whose properties defied conventional analysis.

Shin's Investment Banking model assumes common knowledge as to the quality of risk assets on the market; but the possibility of Investment Banks getting favourable ratings for assets known to be high-risk challenges this assumption. The reversal of 'good news' comes about when the mis-rating comes to light. As [Ak-

erlof and Shiller, 2015, p. 36] put it: 'The mortgage-backed securities may have been rated very highly; but they were largely backed by subprime loans with a high chance of default. When it was discovered that these loans were worth a lot less than previously thought, the investment banks were bankrupt.'

Financial developments over the course of 2008 seem to support this perspective. For, as summarised succinctly in [Sorkin, 2009, p. 529] and indicated in Table 3.3 below: 'Each of the former Big Five investment banks failed, was sold, or was converted into a bank holding company. Two mortgage giants and the world's largest insurer were placed under government control. And in early October, with the stroke of the president's pen, the Treasury – and by extension, American taxpayers – became part-owners in what were once the nation's proudest financial institutions.'

Adverse selection and the securitisation of subprime assets

As well as holding asset backed securities on their balance sheets, Investment Banks played a key role in the growth of securitisation that is portrayed above in Figure 3.1. The dual involvement of the banks contradicts the 'hot potato' – originate and distribute – version of events, as Shin [2010] argues. So what if the securitisation process made it difficult for investors to assess the quality of their investments?

To help analyse to the role of investment banks and the rating agencies in packaging and marketing MBS, we apply the adverse selection model of Akerlof [1970] under various assumptions about information as to quality. First we describe the inefficient low-trade equilibrium that Akerlof's analysis predicts given asymmetry of knowledge of quality as between buyers and sellers (but common knowledge as to the parameters of the quality distribution). This being so inefficient relative to the outcome with symmetric knowledge, the question posed is whether the credit rating agencies (CRAs) succeeded in restoring informational symmetry by delivering true quality ratings; or whether, as argued by Akerlof and Shiller, there was 'mining of reputation' by the CRAs who inflate the ratings so as to please the Investment Banks. In the latter case, we show how 'rating inflation' allows sellers to collect more than the assets are worth in a cheating equilibrium. However, if ratings lose all credibility when buyers discover evidence of mis-rating – and if buyers also lower their belief as to the lower bound of asset quality - the result could be market collapse, as indicated in Figure 3.1 above.

Asymmetric information

Let there be a pool of risky assets, each indexed by θ , a measure of 'quality'.[14] Assume that the price of risky assets is determined by risk-averse investors in a competitive market. With full information, we normalise the price of asset θ to be θ.

In what follows, we characterise pricing in competitive equilibrium under asymmetric information. The information structure is that the support and the distribution of is common knowledge to both the banks and the investors, but only the banks know the quality of any given risky asset. The risky assets are "packaged" and held or sold on by Investment Banks who assign reservation values to these assets denoted $r(\theta)$ where $r(\theta) < \theta$.

The pool of the risky assets available constitutes a set $[\underline{\theta}; \overline{\theta}]$, with the measure of quality below θ represented by a cumulative distribution function $F(\theta)$. Given the asymmetry of information as to quality, there will be a single price reflecting the average riskiness of assets made available at that price. Let this price be p.

As banks will only supply these assets if the market price covers their reservation value, $\Theta(p)$, the amount of risky assets supplied at any given price, is defined as:

$$\Theta(p) = \{\theta : r(\theta) \leq p\}$$

A competitive equilibrium is a price p^* and a set Θ^* of risky assets such that:

$$\Theta^*(p) = \{\theta : r(\theta) \leq p^*\}$$

and

$$p^* = E[\theta \mid \theta \in \Theta^*]$$

which together imply that the competitive price must satisfy:

$$p^* = E[\theta \mid r(\theta) \leq p^*]$$

i.e. that it matches the expected value of the assets which have reservation values less than the equilibrium price. (Those with higher value are withdrawn.)

An illustration

Let the pool of the risky assets be uniformly distributed in $[\underline{\theta}; \overline{\theta}]$, with reservation values of $r(\theta) = \alpha\theta < \theta$. The equilibrium price may be determined as follows.

[14] Thus if all risky assets have the same expected returns but differ in their standard deviations, the parameter θ would represent the inverse of the standard deviation.

Assume the equilibrium price to be p, then the set of risky assets offered by banks is:

$$\Theta(p) = \{\theta : r(\theta) \leq p\} = \{\theta : \alpha\theta \leq p\}$$

If this is the set of risky assets sold in the market, the conditional expectation of the quality of assets can be determined as:

$$E[\theta \mid \theta \in \Theta(p)] = \frac{\frac{p}{\alpha} + \underline{\theta}}{2}$$

The equilibrium is given by the condition requirement that, in a competitive equilibrium,

$$p^* = \frac{\underline{\theta}}{2 - \alpha^{-1}}$$

The equilibrium price is within the lowest and the average quality of asset if $\alpha \in \{\frac{1}{2}, 1\}$ and $\alpha \geq \frac{1}{2}(1 + \frac{\underline{\theta}}{\overline{\theta}})$.

With p^* as the equilibrium price, the highest quality asset sold in the market is $\theta_H = \frac{p^*}{\alpha}$. So the set of assets in equilibrium is $\Theta(p^*) = \{\theta : \underline{\theta} \leq \theta \leq \theta_H\}$. In the presence of asymmetric information, only lower quality assets are sold in equilibrium.

Note that when $\underline{\theta} = 0$ then $p^* = 0$ so only assets which have no reservation value are available in the market. Note also that the smaller is , the higher would be the equilibrium price. So decreasing increases the set of assets sold in equilibrium.

The competitive equilibrium in this case is illustrated in Figure 3.4. For prices falling between the lowest and highest 'reservation values' $r(\underline{\theta})$ and $r(\overline{\theta})$, the expected quality will lie on the schedule labelled BT running from the lower bound $\underline{\theta}$ at B to the mean $\tilde{\theta}$ at T. Equilibrium, where the price matches the expected quality, is at E, where BT crosses the 45 degree line. This is the 'rational expectations' equilibrium of Akerlof [1970] where the price is, *on average*, justified by quality. As only lower quality assets are put on the market, it is clearly inefficient relative to the *symmetric information* case, where price matches quality on each and every asset and all MBS will be on the market, as indicated by the dashed section of the 45 degree line between $\underline{\theta}$ and $\overline{\theta}$.

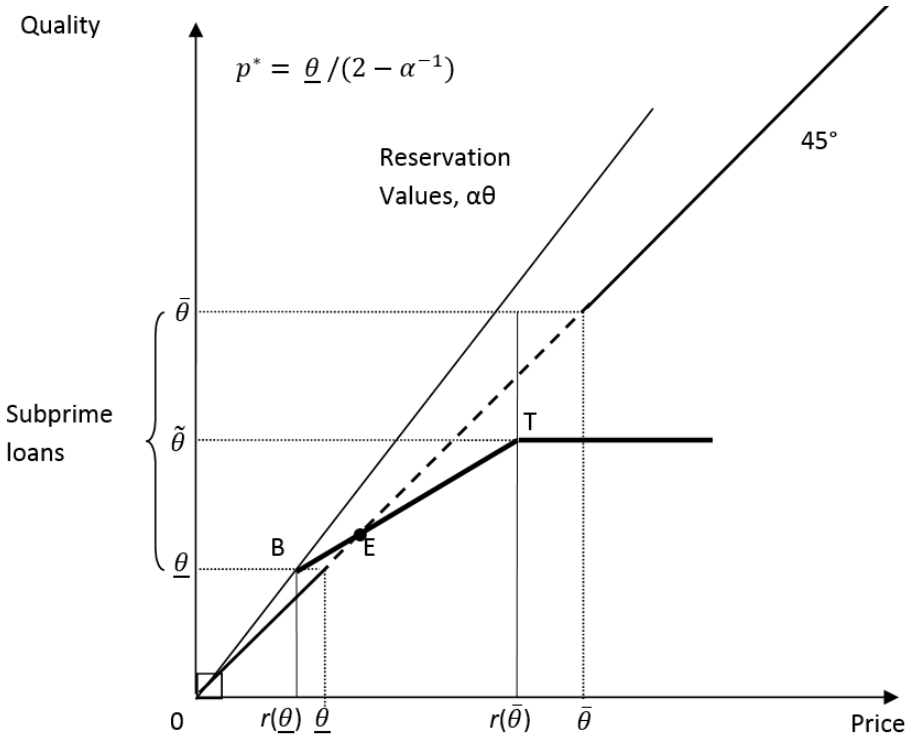


Figure 3.4: Market equilibria with asymmetric and aymmetric information

Faking the ratings

For sellers to bundle loans into 'buckets' of similar quality would seem to offer obvious efficiency gains. In the limit, if the grading is fine enough, Pareto efficient equilibrium might be achieved where all loans are traded and average quality rises to $\tilde{\theta}$.[15] Given the asymmetry of information, however, there is an obvious temptation for sellers to indulge in 'grade inflation'. Hence the case for third party authentication, by Credit Rating Agencies (CRAs) in particular.

With collusion between the sellers of MBSs and the CRAs – if the latter are prepared to upgrade quality ratings in order to retain business – then the grade inflation will not be checked. A sequence of events consistent with a rise in price of securitised assets followed by market collapse (as indicated earlier in Figure 3.1) is illustrated in what follows.

Given that the spread of quality is uniformly distributed in $[\underline{\theta}; \overline{\theta}]$, and equilibrium with adverse selection at E, correct authentication could add to the average value of MBS traded and, in principle, deliver mean quality of $\tilde{\theta}$. But, with collusion between the sellers and the CRAs, buyers can be misled as to the quality. Assume

[15] as in the symmetric information case just described.

for example that with 'grade inflation' the lower bound remains unchanged, but the upper bound apparently increases to $\overline{\theta}'$, where $\overline{\theta}' - \underline{\theta} = 2(\overline{\theta} - \underline{\theta})$ i.e. the spread has doubled, so the apparent quality range of authenticated assets on the market now has a mean value at $\overline{\theta}$, the high end of the actual distribution. The dashed line labelled MM in Figure 3.5 shows graphically how buyers are being misled, with the slope of less than 45 degrees indicating how the price/quality relationship is being distorted (as the overstating of product quality increasing as actual quality rises). If these distorted ratings are taken at face value, all assets will be traded but prices will systematically exceed actual quality (except at the very bottom). The average price paid will be $\tilde{p}$, as indicated on the horizontal axis, which will exceed the average quality shown as $\tilde{\theta}$ on the vertical axis, with 'overpayment' averaging $\tilde{p} - \tilde{\theta}$, as indicated by the bracket in the figure.

With buyers being systematically misled as to quality, this is no 'rational expectations' equilibrium. Differential information is actively being exploited to the advantage those who know the true quality of the MBS that they are misselling. In this in this respect it differs from models such as that of Tella [2017], where intermediaries have known incentives to 'steal' but markets adjust so that, in equilibrium, there is no stealing. In choosing between such different perspectives, subsequent legal findings can play a crucial role, as discussed below. What happens when the music stops and buyers discover that many of the loans are not, in fact, worth what they were led to expect? It seems self-evident that the ratings will lose credibility and buyers become more wary of subprime than before. Let us assume, specifically, that the ratings are totally disregarded, with prices determined as for equilibrium with adverse selection. Assume also that buyers also shift their beliefs to the detriment of MBS: while willing to credit that the upper support is $\overline{\theta}$, they now believe the lower support is zero. With sellers and buyers differing in respect of the parameters of the quality distribution, the equilibrium will not have the 'rational expectations' feature of Akerlof [1970].

What will the equilibrium be? Despite the quality being as originally specified, with bounds $\overline{\theta} > \underline{\theta} > 0$, the jaundiced beliefs of the buyers, with bounds $\overline{\theta} > \underline{\theta'} = 0$, now implies the schedule of expected quality (from the viewpoint of the buyers) is as shown as B'T' in Figure 3.6. As this lies below the 45 degree line showing actual quality except at the origin, sellers will find their asset quality systematically undervalued. So the market will collapse with no trade in assets of any quality in what is the no-trade equilibrium of Akerlof [1970], arrived at here by excessively pessimistic beliefs.

The behavioural phenomenon which Gennaioli et al. [2012] have dubbed 'ne-

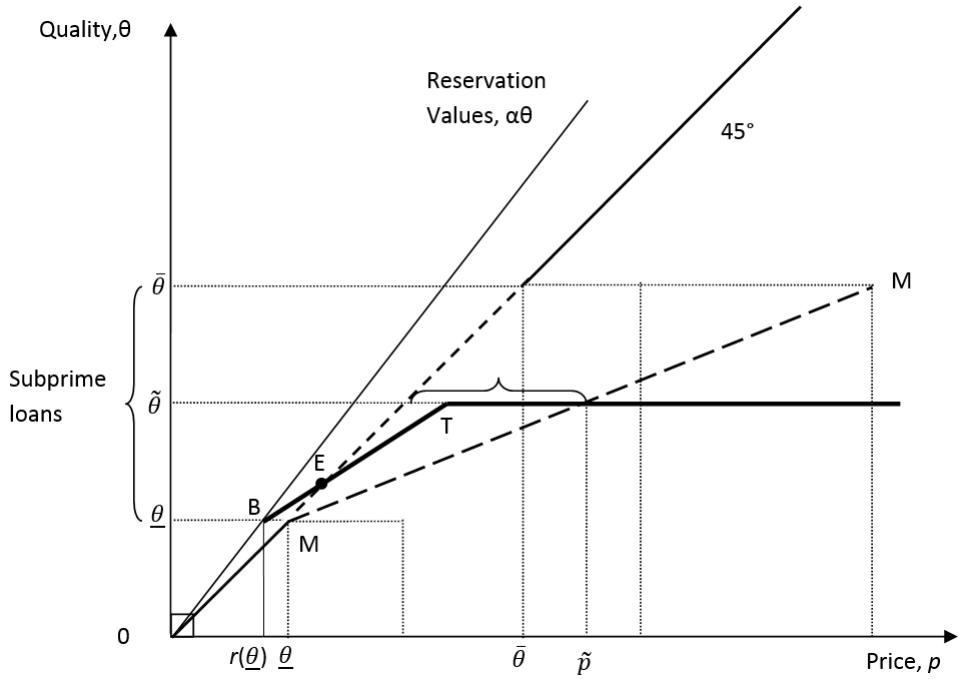


Figure 3.5: Cheating: market equilibrium with inflated ratings

glected risk' – the tendency of investors to ignore certain possible outcomes – also supports the 'phishing' perspective.

> "The key insight is that bankers will create securities that are vulnerable only to those neglected risks. ... For, example, if investors convince themselves that house prices throughout the country cannot fall by 10 percent or more, then bankers will create securities that retain their value in every scenario except when house prices throughout the country fall by 10 percent or more. Because these securities look riskless to investors, they will be produced in abundance. This large expansion in the supply of securities that look riskless will fuel an asset bubble by allowing optimists to buy even more expensive homes. When house prices do in fact fall by more than 10 per cent, the result is catastrophic." Mian and Sufi [2015, p. 13-14]

In addition, Foote et al. [2012] uses empirical evidence to argue that the decline in house price during the crisis might have been entirely neglected scenario rather than being considered with low probability ex-ante.

The account derived from the presence of asymmetric information – and its

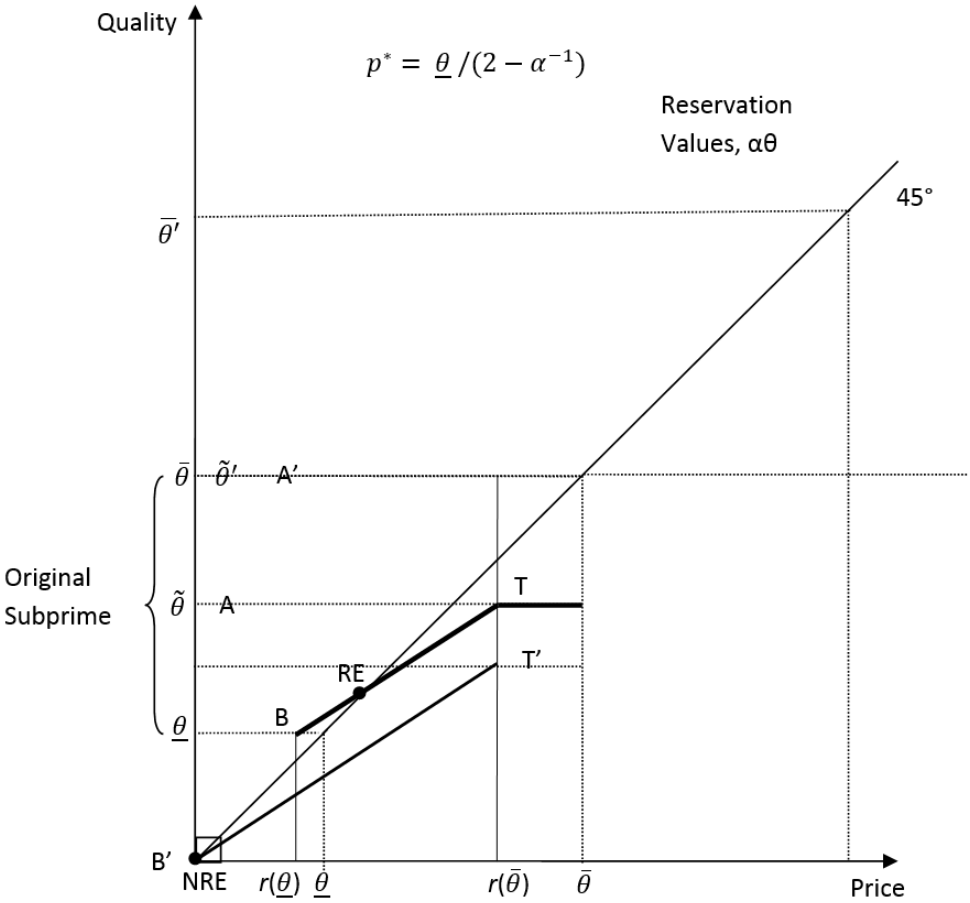


Figure 3.6: Despair: downside revision of sellers' distribution, leading to collapse.

exploitation by banks working in collusion with the rating agencies to whom prudential regulation had effectively been delegated – generates a boom/bust sequence much like that in the previous section. But the 'shocks' on asset quality becomes *endogenous*. This does not imply that the financial accelerator that Shin emphasises is irrelevant: the impact of developments, both positive and negative, on the equity base of the banks involved will amplify their effects on industry equilibrium, rendering implosion more likely.

If 'reality is one, though wise men speak of it variously', one may be tempted to ask whether – and how – these seemingly conflicting accounts might be combined. An ingenious exercise along these lines, Zhang [2017], involves applying the Shin model to determine Demand for risk assets by active and passive investors (based on common but less-than-complete knowledge as to quality), and Akerlof's approach to determine Supply. The latter will depend on the level of participation

by those securitising risk assets, who have full information about quality – where a *distribution* of downside risk provides the basis for distinguishing risk assets on the basis of quality. Adding the assumption of a given *quantity* distribution of risk assets (e.g. a quantity of each quality), generates the Supply curve. Aggregate demand by active and passive investors for bundles of MBS securities – who estimate quality by the unconditional mean of the distribution – gives the Demand for risky assets. So market clearing, where demand matches supply, provides one way of combining the two approaches.

Such an equilibrium – where the quality assumed on the demand side will exceed the mean quality on offer – has the virtue of showing that the rationality of equilibrium in Akerlof [1970] is not a necessary feature of a model with asymmetric information; and comparative statics will involve both amplification and endogenous supply. But evidence from subsequent legal prosecutions and fines indicates that simply conflating these two approaches omits a key aspect stressed by Akerlof and Shiller [2015], namely the incentive for those with superior information to turn it to their advantage. In practice, it seems, Suppliers turned to 'manipulation and deception', as discussed above.

If the subprime crisis merits the description of a perfect storm, it is because it involves so many contributory factors. The two models examined above highlight particular features – the challenges to financial market efficiency and stability coming from asymmetric information and from pecuniary externalities. How these may best be combined is left as unfinished business. For, like the elephant in the parable, reality is undeniably complicated. In the next section, we turn to another aspect – the idea of creditor panic, a bank run.

3.3 Illiquidity: mindless panic or realistic reassessment?

The 'insolvency' views discussed in Sections 3.2.1 and 3.2.2 focus on the poor quality of bank assets and the excessive leverage and risk-taking involved. The 'liquidity crisis' view by contrast emphasises:

> "excessive reliance on short-term borrowing and the resulting maturity mismatch, the weakness of 'mark to market' accounting rules...and the panic withdrawal of short-term funding that created wide-spread market illiquidity, resulting in undervaluation of assets and the dislocation of money markets where banks normally borrow short term." [Milne, 2009, p.18-19].

Indeed Gorton [2010], one of the leading proponents of this view, titled the paper on the subprime crisis presented to central bankers and academics at Jackson Hole, as 'The Panic of 2007'.[16]

The principal piece of evidence Gorton refers to is the cost of insuring against losses on subprime mortgages, as measured by the ABX-HE indices. From January 2006 onwards these indices were constructed to price traded insurance contracts, each contract providing cover on repayments of a bundle of Mortgage Backed Securities for a period of five years.[17] Figure 3.7 shows the movements in the BBB and AA versions of this ABX index, reflecting the cost of purchasing investment-grade tranches of twenty major MBS products.

While both indices initially stood at par, the relatively riskier ABX-BBB index began to fall at the beginning of 2007; and both indices began to fall sharply after August 2007 - the date the Panic began, according to Gorton. Continued precipitous decline took the BBB down to about 5c in late 2008; by which time even the less risky ABX-AA index was down to 20c, implying up-front insurance costs of 80c in the dollar.

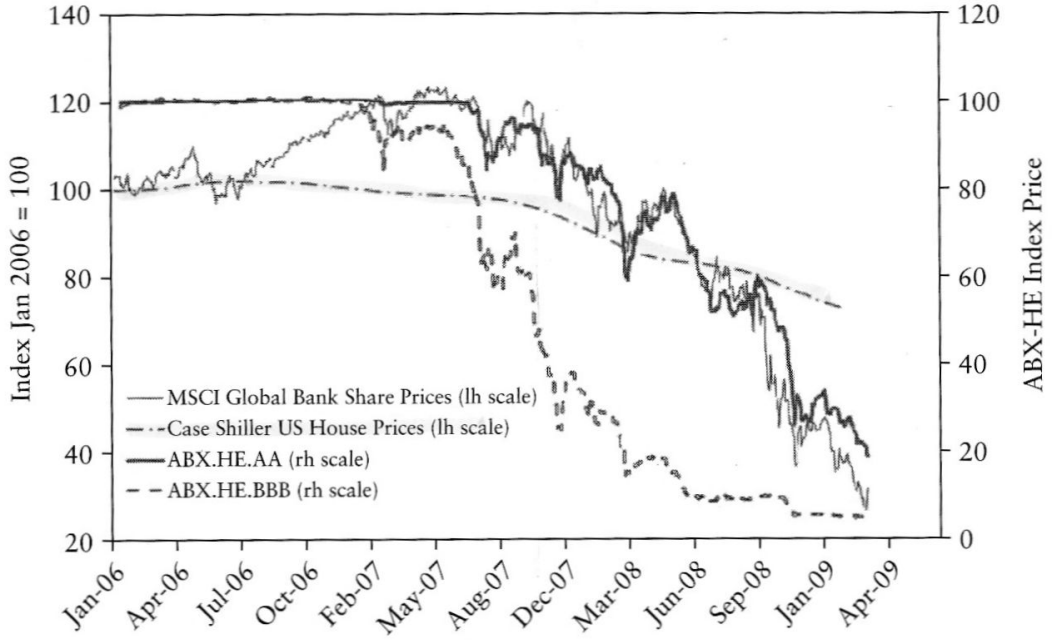


Figure 3.7: US House prices, ABX indices, and share prices of global banks.
Source: [Milne, 2009, p.201].

[16]This paper was later incorporated in his monograph on the subprime market, Gorton [2010].

[17]Thus a price of 80 for a particular AAA contract on a given date means that the protection buyer must pay 20% of the par value of the AAA index to get protection for the next five years.

Reasons for panic: opacity or product design?

For [Gorton, 2010, p.209]the main reason for panic was the 'loss of information' involved in securitisation and the consequent 'opacity' of MBS securities in terms of their asset backing. He states specifically that 'House price declines and foreclosures do not explain the Panic'. But, in commenting on Gorton's paper, [Holmstrom, 2008, p.201] argued to the contrary:

> The problem with sub-prime related securities was not the lack of transparency as such... the real problem was the sensitivity of the MBSs to a fall in the average house price. ... The dynamic credit enhancement model only worked as long as house prices were rising, a point that seems obvious in retrospect.

For there was a catch to the 'dynamic credit enhancement' on offer: the finance provided when house prices were rising would cease when house prices stopped rising, or began to fall. When those who had been lent the funds found no refinancing was available, they would be unable to avoid the scheduled step-up in rates (possibly doubling); and, if house prices were falling, they would need to post more collateral or pay down the loan: otherwise, they could become homeless as their homes were repossessed.

That the banks involved were aware of this emerged from subsequent legal investigation. An email from Angelo Mozilo, co-founder of Countrywide to other Countrywide bank executives, dated August 1, 2005, warned explicitly that:

> "when the loan resets in five years there will be enormous payment shock and the borrower is not sufficiently sophisticated to truly understand the consequences, then the bank will be dealing with foreclosure in potentially a deflated real-estate market. This would be both a financial and reputational catastrophe."

In circumstances when house prices were already high by historical standards[18], aggressive marketing of such loans looks to have two undesirable consequences (a) to push house prices yet higher; (b) to leave as homeless those who were unaware of how and when the finance would effectively be withdrawn when the bubble burst. In the process, homeowners may well have been deceived by the mortgage lenders. [Mian and Sufi, 2015, p.149] note that: 'Home owners mistakenly

[18] As [Gorton, 2010, p. 202] himself notes: 'it was widely understood that house prices were likely a bubble'.

believed that house prices would rise forever. Perhaps this was a silly belief, but the image of a sophisticated home owner gaming lenders and the government is wrong. If anything, sophisticated lenders may have taken advantage of naïve home owners by convincing them that house prices would continue to rise.'[19]

As the Case-Shiller index of House Prices plotted in Figure 3.7 above indicates, property prices in main US cities peaked in the third quarter of 2006, and went on to decline by about 30% over the next two and a half years. This – the timing of house price declines – supports Holmstrom's analysis.

The evolution of house prices in the US over the long run also supports the idea that house prices had been experiencing a 'bubble' in the years when subprime lending had widened access to house purchase (with Private Label Securitisation lowering credit standards). Figure 3.8 shows the index constructed by Robert Shiller giving U.S. house prices in real terms since 1880, where the spike that developed in the early years of this century is clearly visible; and the other series shown suggest that it was not related to underlying fundamentals.

Contagion

Holders of the 'liquidity crisis' view argue that, by retaining only super-senior tranches on their books, investment banks were immune from insolvency risks. But, as [Shin, 2010, Chapter 8] points out, the prevalence of interbank lending and borrowing provides a channel for contagion: the liquidity shock suffered by a bank with good assets may be the consequence of withdrawals by another bank suffering equity losses from poor asset quality (causing it to reduce its balance sheet). In other words, one bank's liquidity shock could reflect another's solvency shock.

[19] In fact, as indicated below when discussing the fines on investment banks, 'part of the settlement requires Bank of America to pay down mortgages for certain home owners; reduce tax payments for others'.

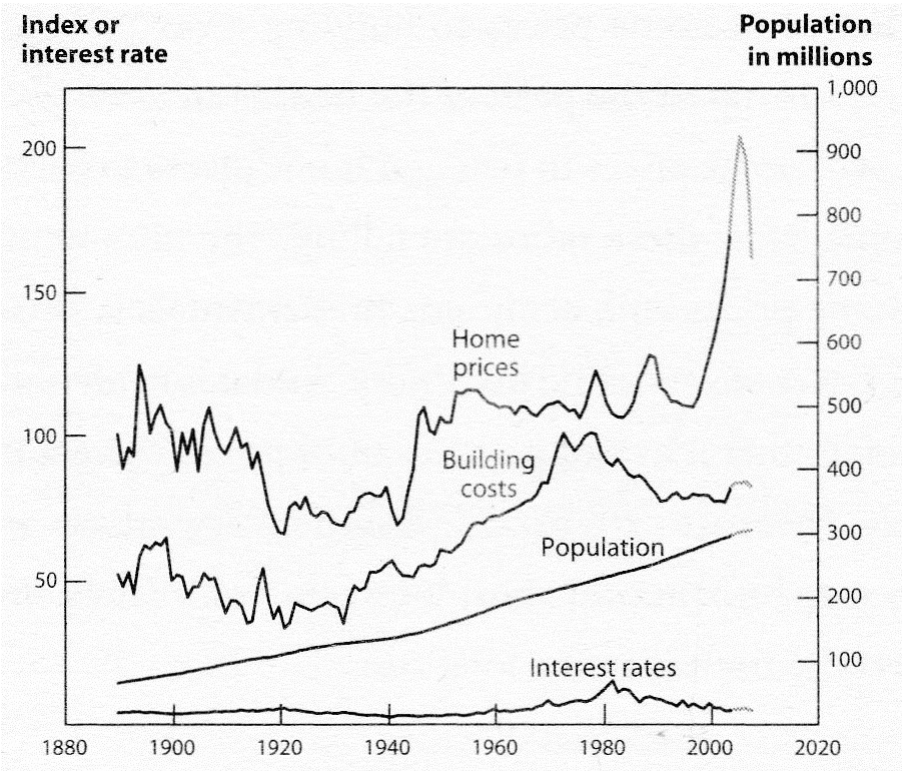


Figure 3.8: U.S. Real home prices in the long run.
Source: R. Shiller(2008, p.33).

3.4 Policy actions and legal evidence

Liquidity provision by the Fed

For Gorton, the opacity of the products created to securitise loans to subprime households leading to creditor panic in 2007 was a key factor in the financial crisis. There is, of course, no question that the banks were exposed to liquidity risk: 'the use of overnight repos became so prevalent that, at its peak, Wall Street investment banks were rolling over a quarter of their balance sheets every night', [Shin, 2010, p. 156].

Action was, moreover, taken by the Fed to help provide liquidity. Thus in March 2008 the Fed created a Primary Dealer Credit Facility making it easier to lend to security firms by widening the range of eligible collateral. Further, when Morgan Stanley and Goldman Sachs – both "enthusiastic practitioners of the new Wall Street model that combined sky-high leverage with heavy reliance on short-term borrowing" – faced a debilitating loss of credit in September 2008, they were

granted the status of Bank Holding Companies" thereby pulling the two beleaguered companies inside the Fed's safety net. That stopped the runs." [Blinder, 2013, p. 153-154].

But this action to extend the safety net was taken in what Blinder calls 'The Panic of 2008', when the structural problems of subprime lending described by Holmstrom were already apparent; in other words, it was a liquidity crisis driven by bad fundamentals. This confirms that the mercurial nature of investment bank liabilities left them prone to creditor panic – so called 'silent bank runs' where creditors fail to rollover their investments; but it hardly supports Gorton's thesis - that a pure liquidity crisis based on opacity , 'The Panic of 2007', was the primary driver of the financial crisis. In the words of Yellen [2017], "the deterioration from early 2007 until early September 2008 was a slow trickle compared to the tidal wave that nearly wiped out the financial sector that September".

Capital injections by the US Treasury: official purchase of preference shares

For Shin, the wholesale take-up of low-quality subprime assets by highly-leveraged banks at a time when measured risks seemed low was the key factor, leaving them exposed to insolvency as and when 'bad news' arrived.

> "As balance sheets expand, new borrowers must be found. Someone has to be on the receiving end of the new loans. When all prime borrowers have a mortgage, but balance sheets still need to expand, then banks have to lower their lending standards in order to lend to subprime borrowers. When the downturn arrives, the bad loans are either sitting on the balance sheets of the large financial intermediaries or they are on special purpose vehicles that are sponsored by them. This is so, since the bad loans were taken on precisely in order to utilise the slack in their balance sheets caused by the apparent lull in measured risks. Although final investors such as pension funds and insurance companies will suffer losses, too, the large financial intermediaries are more exposed in the sense that they face the danger of seeing their capital wiped out." [Shin, 2010, p. 156-157].

This perspective, that the 'fair weather' expansion strategy posed the risk of insolvency when storm clouds appeared, finds support in the action taken by the US Treasury in October 2008. As Blinder notes: 'most banks were presumably under-capitalised on a mark-to-market basis at the time. They needed capital desperately,

and most of them could not raise it on the dire circumstances of October 2008. ... Equity injections would improve banks' capital positions directly'.

Alongside losses and write-downs totalling $344b incurred in 2007/8, the table provides details of the principal capital injections[20] made by the US Treasury using TARP funds, running to a total of almost $100b for the banks in the table. For a pure liquidity crisis, where the investments of the banks are not in question, such capital support is not necessary. But in this case, when house prices were already falling, subprime insurance had become prohibitively expensive, the MBS market had essentially closed down and losses amounted to a third of a trillion dollars, such solvency support was considered essential.

Phishing for Phools?

As noted above, the asymmetric information account of Akerlof and Shiller stands in sharp contrast to the 'common knowledge' perspective of Shin, where active banks compete with patient lenders to supply funds to risky borrowers, maximising profits in a regulatory regime facing exogenous 'news shocks' amplified by VaR-based financial accelerators. What is offered, instead, is an analysis based explicitly on 'the economics of manipulation and deception'. Investment banks can make high returns by taking on 'tail risk', concealing this by getting the investments rated as first class; they can further increase their profits by selling the magic elixir to others, while hedging their own exposure by purchasing CDS. On this account, the risks to which the shadow banking system is exposed are the consequence of exploiting information asymmetry for profit.

Although Akerlof and Shiller make only passing reference to legal measures, substantial support for their perspective comes from legal decisions subsequent to the crisis, as indicated in Table 3.3.

[20] Enforced purchases of preference shares.

The 'Big Five' US Investment Banks (as of early 2008)	Assets, Leverage, and equity end 2007	Fate after crisis	'Big Eight' Banks (Current Survivors)	Credit losses and write downs 2007-8	Capital injections October 2008	Subsequent fines for Mis-selling of MBS
Goldman Sachs	$1,120b (26; $43b)*	Became a Bank H Co in Sep 2008	Goldman Sachs	$10b (0.7)**	$10b	$5b
Morgan Stanley	$1,045b (33; $32b)	Became a Bank H Co in Sep 2008	Morgan Stanley	$19b (2.1)	$10b	$3b
Merrill Lynch	$1,020 (32; $32b)	T/O by Bank of America, Sep, 2008	Bank of America	ML: $73b (7.5) BoA: $57b (1.8)	$25b	$17b (+$37b set aside)
Lehman Bros	$691b (31; $22b)	Liquidation, Sep 2008	_	$30b (5.0)	_	_
Bear Sterns	$396b (33; $12b)	T/O by J P Morgan, Mar 2008	J P Morgan	$41b (2.8)	$25b	$13b
			Citigroup	$114b (4.0)	$25b	$7b
Totals	$4,272b IBs only			**$344b**	**$95b**	**$45b**

Notes: *Figures in brackets are leverage, Assets / Equity, followed by Equity. ** Figure in brackets shows ratio of losses and write downs to 2006 pre-tax earnings.

Table 3.3: Big five investment banks and survivors of the big eight: losses, capital injection, and fines.

Sources: Losses: [Milne, 2009, p. 249]; Injections: [Sorkin, 2009, p. 524]; Fines: (DoJ web reports).

Mis-selling of MBS

First there are 'fines' on the Investment Banks themselves - settlements agreed to with Federal and/or State prosecutors for having misled other investors as to the quality of the MBS they sold. The sums paid by investment banks and the big commercial banks such as Bank of America, J P Morgan and Citigroup amount to $45b, as shown in the Table (of which $8b were levied on the two surviving investment banks, and $20b on the big banks that had taken over Bear Sterns and Merill Lynch).

The largest fines – and some of the most chilling evidence – comes from the case against Bank of America which, in addition to acquiring Merill Lynch, had earlier taken over Countrywide Financial, the largest lender of subprime mortgages in the US. At a press conference where the settlement against Bank of America was announced, Eric Holder, the U.S. attorney general, is on record as saying:

> "These financial institutions knowingly, routinely, falsely, and fraudulently marked and sold these loans as sound and reliable investments. Worse still, on multiple occasions – when confronted with concerns about their reckless practices – bankers at these institutions continued to mislead investors about their own standards and to securitise loans with fundamental credit, compliance, and legal defects."

The fines were for misleading investors as to the quality of the mortgages that had been securitised and sold on. But what about the homeowners who had been persuaded to take out loans which might well fail? In the case of Bank of America 'part of the settlement requires Bank of America to pay down mortgages for certain home owners; reduce tax payments for others; and pay to demolish abandoned homes in certain neighborhoods to reduce urban blight'. In addition, it appears, 'the bank has also set aside $37.3 billion to buy back bad mortgages from investors'. (Guardian newspaper report).[21]

Collusion with CRAs?

The allegation of collusion between Credit Rating Agencies and investment banks has also been the subject of court proceedings; with fines imposed on the two major agencies as follows. In February, 2015 SP settled for a fine of $1.5b – and it was

[21] It would be interesting to see how these redress procedures compare with those set by the Financial Conduct Authority for UK banks who had mis-sold interest rate swaps to small and medium enterprises, as described on www.the-fca.org.uk/consumers/interest-rate-hedging-products.

reported that 'SP executives admitted that they made decisions about testing and rating CDOs based at least partly on the effect they might have on relationships with the banks issuing them'. In January of 2017, Moody's settled for a sum of $0.9b. Both credit rating agencies have thus agreed to pay substantial settlements for mis-rating; with SP admitting what Akerlof and Shiller allege, namely that the ratings were influenced by the incentives to retain the business.

For Gorton the main reason for panic was the 'opacity' of MBS securities in terms of their asset backing leading to a collective 'loss of confidence' hitting the investment banks when the housing bubble burst. But the legal proceedings and fines imposed indicate that the investments did not deserve confidence that they had enjoyed. According to evidence from the courts, mortgage originators had sold financial products to low-income households without explaining the downside risk; and investment banks had mis-sold bundles of the resulting mortgages to other financial institutions, keeping super-senior tranches on their balance sheets. So creditor panic could be understood as a reaction to discovering the business model of subprime lending was seriously flawed.

Summary overview

In previous sections we have presented various insights offered by seasoned observers on reasons for crisis – be it the powerful momentum effects of VaR-based banking, the irresistible temptation to conceal risks faced by many financial operatives, or the quick-silver nature of the liabilities financing much of the risky lending. Each of these factors, it seems, could alone lead to financial disaster. But the actions taken by those holding the levers of monetary and fiscal policy at the time, together with the legal determinations made subsequently by prosecutors in US courts of law, show that, in reality, the crisis had multiple contributory causes.

How evidence from policy action and the law courts supports such a multi-faceted perspective is indicated in Figure 3.9

That the investment banks were hit with liquidity shocks is clear enough. For Gorton [2010], who endorses the soundness of their business model, there was a pointless panic in 2007, stemming- from a 'lack of transparency', see the area labelled (1). His view was challenged by Holmstrom, who pointed the finger at fundamentals; and it was in 2008, after Lehman's went bankrupt, that the Fed gave banking status to Goldman Sachs and Morgan Stanley – the only remaining independent Investment Banks – and supplied them with plentiful liquidity, pending subsequent support from the Treasury. So there was a liquidity shock, but not without good reason.

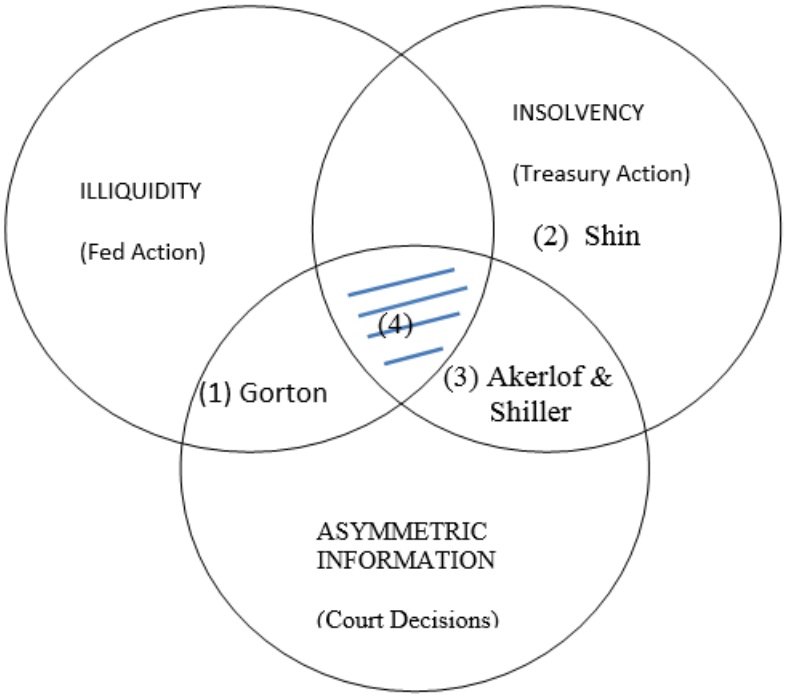


Figure 3.9: Illiquidity, insolvency and asymmetric information, as indicated by policy action and legal proceedings

For banks, of course, risk of insolvency is a key driver of creditor sentiment: and for investment banks who had invested in subprime, solvency was surely at risk: why else did the US Treasury have to step in to provide equity for them with TARP funds? Shin's iconic analogy of the Millenium Bridge suggests that the threat of insolvency could be attributed to some unanticipated, exogenous 'bad news' greatly amplified by 'pecunuiary externality'. In this case, no illiquidity nor asymmetric information need be involved, as for the area labelled (2) in the Figure.

The evidence of legal decisions, however, confirms that world-leading investment banks and rating agencies exploited asymmetric information to market subprime assets as high-quality investments for private profit, as alleged by Akerlof and Shilller. Insolvency involving asymmetric information could lie in the areas labelled (3) and (4), depending on the impact of liquidity shocks.

From policy actions to support the institutions involved – and the fines subsequently imposed upon them – one is led to conclude that the shaded area (4) in the Figure – 'triangulated' by the three views just discussed – gives the best impression of the nature of the crisis. More discussion about the figure with regard to

the interaction between illiquidity and insolvency is also provided in Chapter 1. Its heart of the crisis lay in the betrayal of trust by institutions that were the very pillars of the financial system. That revelations of this could trigger price effects that could have wiped out their equity base is hardly surprising. For these were hardly exogenous 'bad news' shocks; more like evidence - to be borne out subsequently in courts of law - that key players had their hands in the till!

In her speech a decade after the crisis began, the current chair of the Federal Reserve endorsed a multi-faceted approach, noting that: "the vulnerabilities within the financial system in the mid-2000s were numerous and, in hindsight, familiar from past financial panics". As she went on to observe:

> "In response, policymakers around the world have put in place measures to limit a future build up of similar vulnerabilities, ...Preeminent among these domestic and global efforts have been steps to increase the loss-absorbing capacity of banks, regulations to limit both maturity transformation in short-term funding markets and liquidity mismatches within banks, and new authorities to facilitate the resolution of large financial institutions and to subject systemically important firms to more stringent prudential regulation." Yellen [2017]

Legal evidence – why "deferred prosecution agreements"?

To help establish asymmetry of information, we have appealed to the 'fines' imposed on banks and rating agencies for mis-selling and mis-rating. But the courts have been criticised on the grounds that, in contrast to what happened in previous crises – that of savings-and-loan associations in the 1980s and the accounting frauds of the 1990s, for example – 'not a single high-level executive has been successfully prosecuted in connection with the recent financial crisis', Rakoff [2014]. In the article cited, retired Judge Rakoff notes that there has, in fact, been a shift from prosecuting high-level individuals to prosecuting companies. In order to change 'corporate culture', the policy pursued is to secure "deferred prosecution agreements" (DPAs) in which the company, under threat of criminal prosecution, agrees to pay a fine and to take remedial measures to prevent future wrong-doing.

As the Department of Justice has argued, it is indeed difficult to prove fraudulent intent on the part of high-level management of the banks and companies in cases that involve the mis-selling of innovative financial products.[22] While the product itself may provide welfare improvement from better risk-sharing, its complexity

[22]Making successful criminal prosecution is a great deal more costly than securing a DPA, an important factor if, as some allege, the relevant Federal agencies were being starved of funds.

not only gives room for sellers to manipulate the quality being offered to investors, but also proves to be a 'grey area' for court cases. The very complexity allows bank executives to argue that they have the same beliefs about the quality of MBS as they describe to investors, for example.[23] As Kay [2017] notes, however, 'the very prevalence of such [DPA] settlements is an indication that their deterrent effect is small. Senior executives appear not to mind paying out large amounts of shareholders' money to escape any personal liability for their actions, or the actions of those whom they ostensibly supervise'.

Improving the prospects for prosecution may involve reducing the 'grey areas' arising from the complexity of innovative financial products. This will not only assist in pursuit of fraud, but, by reducing product opacity, should restore confidence in the market. It will also require more funding for the purpose.

Interestingly enough, a recent proposal by [Mian and Sufi, 2015, Chapter 12] is that appropriate financial innovation may be the key to preventing a recurrence. They argue that standard mortgage contract, which 'forces the borrower to bear the full burden of a decline in house prices until his equity is completely wiped out, be replaced by Shared-Responsibility Mortgages (SRM). An SRM has two important differences: (1) the lender offers downside protection to the borrower; and (2) the borrower gives up 5 percent capital gain to the lender on the upside'. The risk-sharing involved in such contracts keeps the loan-to-value ratio stable even when house prices fall: it should also make lenders more cautious about lending into the boom – and so limit house price volatility.[24]

Insofar as the subprime experiment was designed to give marginal borrowers access to housing, it must be judged a spectacular failure. Far better to provide explicit subsidies; e.g. government matching of down-payments by new homebuyers as [Calomiris, 2009, p. 29] suggests; or to promote the introduction of SRM contracts as Mian and Sufi and David Miles recommend.

Banking regulation after the Global Financial Crisis of 2008

In responses to the Global Financial Crisis of 2008, the Basel Committee on Banking Supervision (BCBS) has introduced the Basel III involving the revision of banking capital adequacy and liquidity. On bank capital, one of the key focuses is to ensure that banking sector take into account the macro-financial en-

[23]In the circumstances, rather than individual banks 'phishing for phools' it might be more appropriate to talk of 'market phishing', to borrow the terminology of Akerlof and Shiller [2015].

[24]The UK experiment with a 20% 'shared equity loan' available under the Government's 'Help to buy' programme launched in 2013 is analysed in Miles [2015], who argues that even such limited risk-sharing could have substantially mitigated the Great Recession in the UK.

vironment associating with its business. As a result, the countercyclical buffer has been introduced to achieve the broader macroprudential objective in ensuring safe and soundness of banking sector from excessive credit growth period that usually supports the build-up of system-wide risks. In fact, by forcing banking sectors to internalise the externalities they may pose to the economy, the countercyclical nature of the measure also helps in containing the amplification effect of financial systems in exacerbating financial cycles and, thus, in alleviating the risk of squeeze in the supply of credit during the downturn.

On liquidity, the main components in Basel III include the introduction of the net stable funding ratio (NSFR) and the liquidity coverage ratio (LCR). Both measures, to ensure the soundness of banking sector, emphasise the need to set limits on maturity transformation, which lies at the heart of banking business given its role in reconciling the preferences of savers and investors. However, by imposing fixed requirements over the cycle, these new liquidity measures could compromise the macroprudential objective of the Basel III due to the interaction between maturity transformation and leverage of banking system. In fact, leverage would not be profitable without maturity transformation. In the environment where the funding is cheap and the market-wide perception of risk is low, inexpensive maturity transformation encourages banks to leverage more and to expand their balance sheets. However, when the downturn arrives, a liquidity problem on the liabilities side of financial institutions' balance sheet could lead to fire sale on the asset side as discussed in Chapter 1. In this sense, excessive maturity transformation could pose risks and make the system more vulnerable to a shock. As a result, there are good reasons to regulate system-wide maturity transformation by taking into account its cyclical variation in a similar manner as countercyclical capital buffer.

3.5 Conclusion

Given mark-to-market accounting and the usual VaR conventions, highly-leveraged investment banks could, it seems, face insolvency due solely to exogenous common shocks to fundamentals: a simple reversal of 'good news' on the perceived quality of risky assets could be sufficient. We have argued, however, that the shocks were in practice endogenous – due to the mis-selling of subprime assets by investment banks, assisted by excessively favourable assessments on the part of rating agencies. Rather than some 'rational expectations' equilibrium with common knowledge, the legal evidence is of temporary cheating equilibrium leading to crisis when the truth emerges.

That financial institutions who nightly need to roll over a quarter of their balance sheets are exposed to creditor panic is uncontroversial. What is controversial is to maintain that the subprime crisis was caused by unreasoning panic based on product opacity rather than fear based on bad news about the business model.

The three views we have focussed on provide plausible threats to financial stability coming from differing factors – from externalities, cheating, or balance sheet fragility. So one may be tempted to ask: which one is correct? As the parable of the elephant suggests, however, each may provide a partial perspective of a complex reality that involves all of them. The actions of policy-makers and the courts supports this conclusion, as does the ex post assessment of the current chair of the US Federal Reserve.

That these threats to stability should be complements and not substitutes is more than a point of technical interest. What is being described is how radically the US financial system was exposed to failure. The emergency actions of policy-makers was, however, to provide unprecedented liquidity and capital support to world-leading banking institutions soon to be found guilty of serious malfeasance. Rating agencies, upon whom the Basel Committee had seen fit to rely for ensuring the quality of bank risk assets, were likewise found by top prosecutors to be complicit in deception.

Acemoglu and Robinson [2013] have famously argued that a nation's prospects for successful long-run growth depend essentially on the quality of its institutions: and the US is typically cited as an example of best practice. That the US financial system and its key institutional pillars should be so injury-prone must throw some doubt on this assessment, particularly if the law is being hobbled in its pursuit of those responsible.

Chapter 4

Heterogeneous Beliefs, Endogenous Risk, and Crash

4.1 Introduction

Three key features lie at the heart of many financial crises, including the global financial crisis of 2009: optimism, leverage, and default. Bhattacharya et al. [2015] use a formal model to show that agents become more optimistic about the future prospects for an economy after a prolonged period of good news because they update their expectations over time according to previously realised good outcomes. This contributes to the severity of the crisis. Indeed, after a number of good expectations have been realised, new expectations grow; agents increase their leverage, and find it more profitable to shift their portfolios to projects that are on average riskier, thus making their portfolios more vulnerable to shocks. As a result, when a downturn occurs, the ensuing financial crisis turns out to be more severe. The study also highlights default as a source of pecuniary externality when investors fail to internalise the subsequent impact of their decision to take risk and the prospect of default into account.

Indeed, the default occur when the borrower's promised payment exceeds the value of collateral, and the lender receives the collateral. The fragility of asset or collateral value is usually built during the good time. From this perspective, the recent global financial crisis also highlights the importance of different beliefs of investors in pushing asset prices up during the boom times, especially when optimists in the economy can leverage their view by buying assets, that are then vulnerable to subsequent collapses in asset prices.

The build-up and subsequent crash in asset prices could be attributed to the

introduction of financial innovations, as shown in Fostel and Geanakoplos [2012]. Their study suggested that even without asymmetric information, leverage and financial innovations can have major effects on asset prices if markets are incomplete. In particular, the sequential introduction of these financial innovations was, in and of itself, enough to cause boom and bust. The study shows that, in a general equilibrium model of competitive markets, where investors have heterogeneous beliefs, leveraged lenders can raise the value of securitised assets; moreover, if the risky asset is 'tranched', given the availability of covered credit default swap (CDS) contract backed by a risky asset, this will lead to an asset bubble. The innovation that can potentially trigger collapse is the introduction of naked CDS contracts backed by a crash in asset prices. Instead of being an insurance against failure, naked CDS act as a vehicle for optimists and pessimists to leverage their views. As a result, the availability of such naked CDS contracts is enough to burst the bubble. (For further detail and discussion, see *Appendix D.1.*)

In fact, even without the sequential introduction of financial innovations, the crash in asset prices could also be explained by the leverage cycle, as suggested by Geanakoplos [2010], which studies the binomial model with endogenous leverage using the general equilibrium analysis of collateralised lending and asset prices. Although the baseline model in his analysis shows the crash in asset prices when some optimists whose equity is wiped out have to leave the market, it exhibits no default; loans are collateralised based on the worst payoff of assets and the promise rules out default in equilibrium. This idea of no default in the binomial framework is formalised and emphasised in Fostel and Geanakoplos [2015], which proposes a 'Binomial No-Default Theorem' suggesting that with a single financial asset serving as collateral in a static binomial model, we can assume without loss of generality that there is no default.

To allow for risky loans and default in equilibrium, Simsek [2013] uses a continuum of states to analyse the effect of disagreements of belief in fundamentals between two types of traders involving assets and financial contracts. This study (extending Geanakoplos [2003], which is closely related to Geanakoplos [2010]) suggests that belief disagreement affects asset prices and leverage altogether, and that investors disagree on asset prices to a greater extent than their level of disagreement on fundamentals.

In contrast to Simsek [2013], Yan [2017] shows that belief heterogeneity affects the relationship between asset prices and leverage. Indeed, changes in market average beliefs, rather than belief disagreement itself, determine the co-movement of leverage and the price of an underlying asset that serves as collateral for loans.

The study also shows that the asset prices are above their fundamental value when collateral constraints are binding. However, the default in this model arises when the realisation of collateral value is less than the face value of debt.

This chapter studies reversibility in a general equilibrium model of competitive markets with heterogeneous beliefs. The analysis in this chapter is related to the literature on heterogeneous beliefs (as discussed above), and on default (as summarised in Chapter 1). The main contribution of this chapter is to show that a simple reversal of shocks in terms of unanticipated increases in the quality of risky assets,[1] could lead to widespread insolvency and collapse; this is because the shocks are amplified by the interaction between asset prices and heterogeneous beliefs, resulting in asymmetric adjustment of risky asset prices.

The model setup in this chapter is akin to the binomial framework as in Fostel and Geanakoplos [2012]. However, there are two key modifications: first, investors are risk averse, and they only value consumption in the last period. This ensures that investors prefer strictly positive consumption in both states of the world as long as their beliefs on the probability attached to both states are not zero. This implies that the borrowing constraints of those investors are not binding. Indeed, the key determination of investors' borrowing is their wealth, which greatly relies on asset prices, rather than on borrowing constraints. Second, to study the adjustment of risky asset prices in equilibrium, the model introduces the reversal of an unanticipated increase in the quality of a risky asset. [One of the main shortcomings in this setting is that, in contrast to Bhattacharya et al. [2015], the probability agents attach to the states of nature are fixed. In addition, investors could be interpreted as perfectly myopic agents, who totally neglect changes in the quality of a risky asset. These could be the key improving areas of the model.]

There are three consequences of these modifications: first, the shock is amplified by an interaction between asset prices and heterogeneous beliefs due to the allocational effect among optimists and pessimists through asset prices. Thus an unanticipated improvement as to asset quality raises the equilibrium price of the asset. As a result, the optimists, who currently hold the asset, become wealthier with capital gains that stems from the increase in asset price. With such a gain, this group of optimists, who highly value the asset, are willing to increase borrowing to obtain more assets, pushing the asset price to increase further. The system is prone to subsequent crash even the positive shock is reversed to the initial level because

[1] The foundation of unanticipated news shock is motivated by 'neglected risk', a behavioural phenomenon mentioned in Gennaioli et al. [2012] as discussed in Chapter 1. The approach in applying unanticipated shocks is similar to the one used in Chapter 3 in analysing the role of externalities in Shin's model.

the asset holdings are concentrated in a small group of highly leveraged optimists. Second, the default occurs if, given the reversal of positive shock, the asset price drops below the worst-possible payoff of risky asset as perceived before the shock. The sharp drop in asset price could be attributed to low-valuation buyers when the optimists choose to reduce their holdings of the asset in response to the reversal of the shock, while the pessimists find this an opportunity for investment. Third, this default is systemic when some optimists, who provide high valuation on the risky asset, default at the same time, and those optimists must exit the market.

Section 4.2 describes the model setting of a two-period economy with heterogeneous beliefs. Collateral and financial contracts are also specified in this section. Section 4.3 characterises the basic equilibrium without leverage, provides numerical examples, and gives analysis and discussion. Section 4.4 characterises the equilibrium, where leverage is allowed for all agents, and analyses the possibility of default with the use of numerical examples. Finally, Section 4.5 sums up all analyses, and concludes the paper.

4.2 Two-period economy with heterogeneous beliefs

In this general equilibrium model with collateral, there are two dates, $t = 0, 1$. Uncertainty is represented by two states of nature, at $t = 1$, $S_1 \in \{U, D\}$ (a tree). There are two types of assets, which produce consumption goods, at $t = 1$. The safe or risk-free asset, W, generates consumption goods of 1 in both states, U and D. The risky asset, Y, generates consumption goods of 1 and R, where $R < 1$, in states U and D, respectively. The payoffs of both assets are described in Figure 4.1.

There is a continuum of risk-averse agents, $h \in [0, 1]$, who are not impatient and only value consumption at $t = 1$. Agents are different due to heterogeneous beliefs in the probability attached to the states of the economy, q_U^h and q_D^h. The height of optimism of the agents is thus denoted by their indexes h. Their utility is characterised by a neoclassical utility function (increasing, strictly concave, twice continuously differentiable) of consumption goods, C at $t = 1$. The agents maximise their expected utility:

$$U^h(C_U, C_D) = q_U^h C_U + q_D^h C_D \tag{4.1}$$

where, q_U^h is continuous and strictly monotonically increasing in h. For numerical purposes, let us assume $U(C) = ln(C)$, $q_U^h = h$, and $q_U^D = 1 - h$. In addition, each agent has been endowed with one unit of each asset, at $t = 0$, and nothing else.

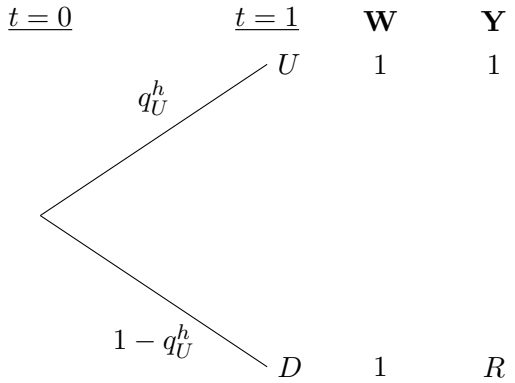


Figure 4.1: Risk-free and risky asset payoffs

4.2.1 Collateral and financial contracts

Suppose, agents can trade financial contracts at $t = 0$. The financial contracts is a non-state contingent, and backed by collateral, which is the sole enforcement mechanism for repayment. In other words, this 'no recourse' loan contract allows the lender to seize the collateral, but, in the case of default, the lender cannot receive any further compensation. Suppose that only the risky asset Y is eligible as a collateral. As a result, in order to make a promise, the borrower must hold the risky asset Y, at $t = 0$.

In the usual theory, the supply-equals-demand equation determines the equilibrium interest rate in the loan market. However, when taking into account the collateral, it might be unclear which equation should determine the equilibrium level of collateral until Geanakoplos [2010] sheds light in solving this endogenous collateral problem. The financial contract is specified by pairing promise and collateral. Thus, the same promise with different collateral should be considered a different market, and should correspond to different interests rates. In other words, each level of interest rate corresponds to a specific contract, a pair of (promise, collateral). The borrowing amount b (or we can call it the price of the contract) thus has a one-to-one relationship with gross interest rate r determined in equilibrium. Then, b (or r) together with the risky asset price p determines the loan-to-value (LTV) ratio and the margin requirement defined as $1 - LTV$. The reciprocal of the margin requirement represents equilibrium leverage. In sum, the equilibrium with endogenous collateral could be characterized by equilibrium leverage. In equilibrium, the LTV ratio, margin requirement, and leverage are thus simultaneously determined by asset price p and the price of financial contract b (or interest rate r).

With a single financial asset serving as collateral in static binomial model,

we can assume without a loss of generality using the Binomial No-Default Theorem (Fostel and Geanakoplos [2015]) that the maximum-minimum debt contract is the only debt contract actively traded, and there is no default. The maximum-minimum contract is a non-state-contingent promise that corresponds to the worst-case payoff of the collateral R. As a result, there is no actual default. The borrowing constraint and agent's 'no default' condition becomes $b \leq Ry^h$, where, R and y represent a promise and collateral, respectively. Here, and in what follows, it is clear that the maximum possible amount of b in the economy is $R < 1$. This guarantees that the supply of loans is always greater than demand, given (as we will see) that an agent will not borrow until the safe asset is used up. Taking into account that agents are not impatient, and that these agents do not value consumption before the end of the period, the interest rate r become zero.

Note that $b > 0$ implies that an agent is borrowing by selling a promise, while $b < 0$ means that an agent is lending by buying a promise. Lending will reduce the budget of an agent at $t = 0$, but enable him to consume more at $t = 1$. Actually, a lending agent will need to be indifferent between holding a safe asset and buying a promise given r equals zero.

4.2.2 Budget Set

The risk-free asset W can be considered as a numeraire in this model by normalising the price of the risk-free asset at $t = 0$.The price of consumption in each state U and D equals one. Thus, we can possibly consider W as analogous to cash.

Now, suppose, agent h, who is currently holding (endowed with) a portfolio containing w_e^h and y_e^h units of risk-free and risky assets, respectively. The current borrowing of the agent is b_e^h. A positive value of b^h implies borrowing, while a negative value implies lending. Given the risky asset price p, and zero interest rate, each agent h optimises his portfolio by choosing, at $t = 0$, his cash holding w^h, holding of risky asset y^h, and the amount of loan b^h, in order to maximise the expected utility in Eq(4.1). The budget constraint is then:

$$w^h + py^h - b^h \leq w_e^h + py_e^h - b_e^h$$

Notice that the right-hand side of the budget constraint represents the net worth of agent h at the beginning of the period, while the left-hand side shows his net worth given his portfolio choice w^h, y^h, and b^h. Since the interest rate is zero, the problem can be simplified by defining net borrowing as $\omega^h = w^h - b^h$.. The agent is also subject to the borrowing constraint, $b^h \leq y^h R$, and short-selling constraints

w^h and $y^h \geq 0$. The consumption plan of each agent (c_U^h, c_D^h) could be described as:

$$c_U = w^h + y^h - b^h$$
$$c_D = w^h + Ry^h - b^h$$

4.2.3 Equilibrium

There exist a risky asset market and a loan market for agents at $t = 0$. As mentioned earlier, in this zero-interest rate economy without default, the price of a loan is equal to the promise. Thus, the equilibrium is described by the price of risky asset p, the cash holdings of each agent w^h, the asset purchase of each agent y^h, financial contract trade b^h, and the consumption plans of all agents,$((p), (w^h, y^h, b^h, c_U^h, c_D^h))_{h\in(0,1)} \in (R_+) \times (R_+ \times R_+ \times R \times R_+ \times R_+)^H$, such that all markets are clear:

$$\int_0^1 w^h dh = 1$$
$$\int_0^1 y^h dh = 1$$
$$\int_0^1 b^h dh = 0$$

4.3 No-leverage economy

This section considers the reversibility of shocks to asset quality in the economy when borrowing and lending are not allowed, and no promise can be made. First, the initial equilibrium is determined, followed by the equilibrium with good news, where agents re-optimise their allocations given unanticipated positive shocks. This resembles the situation during the subprime boom when there was a series of credit upgrades in subprime-related securities. However, the impressive ratings were downgraded afterward following the introduction of the ABX Index when the fundamental price of these securities seems to have been discovered. The subsequent legal evidence on fines of banks and credit-rating agencies (shown previously) clearly depicts the collusion between investment banks and rating agencies in rating manipulation. Finally, the equilibrium when the good news shock is reversed is described at the end of this section. Figure 4.2 illustrates the timeline of events. It is also important to note that, in this model, the 'rational expectations' assumption is relaxed because the shocks are unanticipated. This is consistent with the behavioural perspective, or so-called 'neglected risk', as in Gennaioli et al. [2012]. The neglected risk phe-

nomenon indicates the tendency of investors to ignore certain unlikely outcomes. As mentioned in Chapter 1, the phenomenon could partly explain why many crises surprise the market.

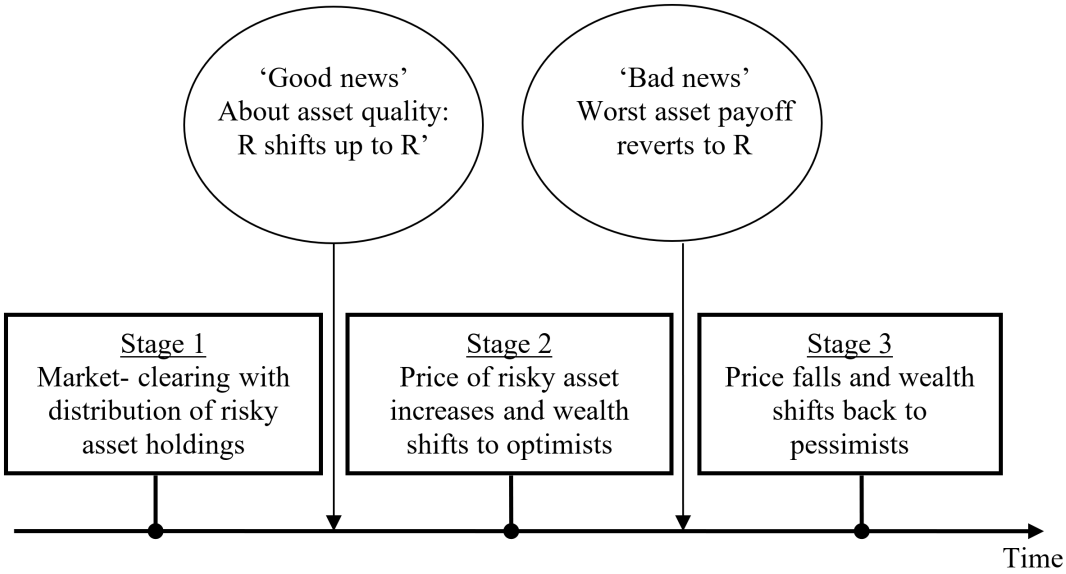


Figure 4.2: Timeline of events

In the absence of leverage, agents can only trade risk-free asset W and risky asset Y. They cannot use those assets as collateral for borrowing, and no promise can be made. In addition, we also assume that short selling is not allowed in this economy. An agent will maximise the expected utility subject to the budget constraint and short-selling constraints. Suppose, at the beginning of the period, agent h is endowed w_e^h and y_e^h units of risk-free and risky assets, respectively. The agents optimise their portfolios by choosing their holdings of risk-free asset w^h and risky asset y^h.

Appendix D.2.1 shows the private decision of an agent. In this economy, there will be three types of agents: (1) extreme optimists with $h \geq h_b = (\frac{p-R}{1-R})(\frac{1}{p})$, who provide higher valuation to the risky asset than others in the economy, will hold only Y; (2) cautious optimists with $\frac{p-R}{1-R} = h_s \leq h < h_b = (\frac{p-R}{1-R})(\frac{1}{p})$ who will have portfolios containing a combination of W and Y; and (3) pessimists with $h \leq h_s = (\frac{p-R}{1-R})$ who prefer fully insured consumption and hold only W. Figure 4.3 illustrates the types of agents in an economy.

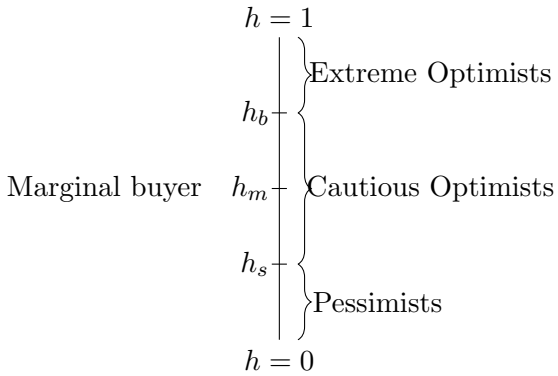


Figure 4.3: Agent types in no-leverage equilibrium

Given the private decision, we can identify the marginal optimist h_b, who is indifferent in holding W (he is also the most optimistic agent that is indifferent between W and Y), and the marginal pessimist h_s, who will be indifferent in holding Y (he is also the least optimistic agent that is indifferent between W and Y). The demand function for risky asset Y given unanticipated news could be described by:

$$y^h(w_e^h, y_e^h) = \begin{cases} \frac{w_e^h + p y_e^h}{p} & if \quad h \geq h_b \\ \frac{(h(1-R)-(p-R))}{(1-p)(p-R)}(\omega_e^h + p y_e^h) & if \quad h_b > h \geq h_s \\ 0 & if \quad h < h_s \end{cases} \tag{4.2}$$

The holdings of safe assets could be described by:

$$w^h(w_e^h, y_e^h) = \begin{cases} 0 & if \quad h \geq h_b \\ \frac{((p-R)-hp(1-R))}{(1-p)(p-R)}(\omega_e^h + p y_e^h) & if \quad h_b > h \geq h_s \\ w_e^h + p y_e^h & if \quad h < h_s \end{cases} \tag{4.3}$$

where

$$\begin{aligned} h_s &= \left(\frac{p-R}{1-R}\right) \\ h_b &= \left(\frac{p-R}{1-R}\right)\frac{1}{p} \end{aligned} \tag{4.4}$$

4.3.1 Initial equilibrium

The initial endowment for all agents is $w_e^h = 1$ and $y_e^h = 1$. Each agent's demand function for risky asset is y^h followed by Eq(4.2). The characterisation of equilibrium can then be presented below. This is followed by the numerical example at the end.

Due to strictly monotonicity and continuity of interior solution (cautious optimists) of y^h in h and the connected set of heterogeneous agents $h \in [0, 1]$, there will be a unique marginal buyer h_m, who will be indifferent between buying and selling Y, and who will not participate in the asset market because he finds his endowment optimal; that is, $y_m^h = y_e^h$. In equilibrium, extreme optimists with $h > h_b$, who, among all the agents give the highest valuation to Y, will buy all they can afford of Y(i.e. selling all W). Cautious optimists $h_b > h > h_s$ will hold a combination of W and Y. Pessimists $h < h_s$ prefer perfectly insured consumption, and will sell all their endowment of Y for durable consumption good W.

To describe the system of equations that characterises equilibrium, I determine the initial equilibrium as the system of two equations below. Eq(4.5) demonstrates the market-clearing condition, which equates supply and demand for risky asset Y. The marginal optimist h_b and marginal pessimist h_s are described in Eq(4.4). The right-hand side of Eq(4.5) shows aggregate selling revenue by agents who lie below the marginal buyer h_m. The first term represents the revenue of cautious optimists, who prefer to reduce their holdings of Y. The second term represents the revenue of pessimists, who sell all their endowment of Y. The left-hand side shows aggregate buying expenditure by agents who lie above the marginal buyer h_m. The first term represents the expenditures by cautious optimists, who prefer to acquire more Y. The second term represents the expenditures by extreme optimists, who spend all of their W endowment for Y. For the equilibrium price of Y, note also that no arbitrage condition implies $R \leq p \leq 1$.

$$\int_{h_m}^{h_b} p(y^h - 1)dh + \int_{h_b}^{1} dh = \int_{h_s}^{h_m} p(1 - y^h)dh + \int_{0}^{h_s} pdh \tag{4.5}$$

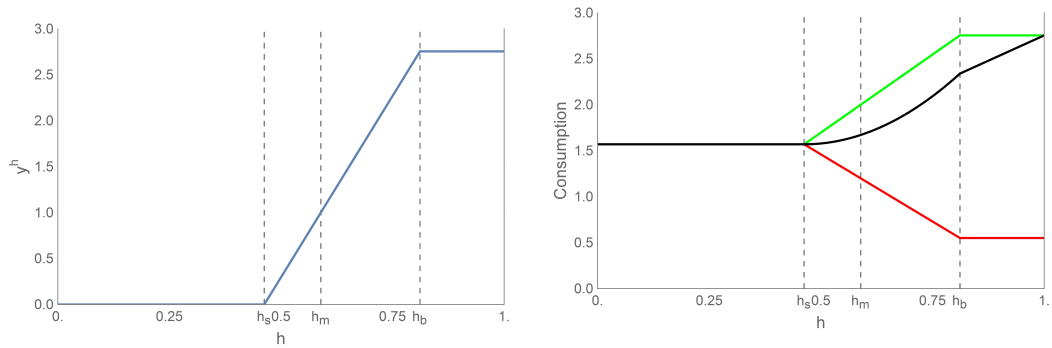
Eq(4.6) represents the valuation of unique marginal buyer h_m, a cautious optimist who is indifferent between buying and selling Y. The left-hand side of the equation shows demand for Y of the marginal buyer; this demand needs to be equal to his endowment of 1.

$$\frac{((1 - R)h_m - (p - R))(1 + p)}{(1 - p)(p - R)} = 1 \tag{4.6}$$

From Eq(4.6), we reach the following:

$$h_m = \left(\frac{p-R}{1-R}\right)\left(\frac{2}{1+p}\right)$$

With a system of equations above with two equations and two unknowns, we can solve for the risky asset price p and the marginal buyer h_m. We can then determine the marginal optimist h_b and the marginal pessimist h_s, accordingly. For $R = 0.2$, we get, $p = 0.570$, $h_m = 0.589$, $h_b = 0.811$, and $h_s = 0.462$. Figure 4.4 describes the allocation of Y in the economy as well as the consumption plan of each agent.[2]



(a) Distribution of risky asset holdings

(b) Consumption plans in U(Green) and D(Red), and expected consumption (Black)

Figure 4.4: No leverage economy - an initial equilibrium

4.3.2 Equilibrium with unanticipated good news

Suppose, there is an unanticipated positive shock to asset quality in which the common belief of risky asset payoff in state D is shifted from R to R', where $R < R' < 1$. This situation could be thought of as an unanticipated credit upgrade on subprime mortgage securities. All agents re-optimise their allocations given this new information arriving the market. However, the initial equilibrium provides heterogeneity in cash holdings w^h, and asset holding y^h as shown in Figure 4.4a. Given the unanticipated good news, the demand functions for the risky asset of the agents, y'^h, could be explained by Eq(4.2) where $w_e^h = w^h$ and $y_e^h = y^h$. We can then show the characterisation of equilibrium below; this is followed by a numerical example at the end.

The good news results in an increase in the risky asset price to p' as, in equilibrium, wealth is being transferred to optimistic agents $h > h_m$ who previously

[2]A detailed simulation result is provided in *Appendix D.2.3*.

bought Y and then, given good news, increase demand for acquiring more Y. The demand from this group of optimists increases for two reasons: first, the worst risky asset is believed to be improved to R' by all the agents. Second, with 'mark to market' accounting, capital gains from the price rise for risky assets on their balance sheets will raise their equity and wealth. In addition, the price will be raised further simply because these optimistic agents, who become richer, value the risky asset more than others do.

Given that y'^h is increasing in h, there exists the unique marginal buyer h'_m, who is indifferent between buying and selling Y; that is, $y'^{h'_m} = y^{h'_m}$. From Eq(4.2), we can easily observe that agents' wealth after re-distribution, $w^h + p'y^h$, is increasing with h as y^h is increasing in h for $h \in (h_s, h_b)$. Figure 4.6 provides a numerical example of the re-distribution of wealth. Indeed, the marginal buyer will shift up to h'_m as the asset price increases because, given that he is indifferent between buying and selling Y, it's clear that the marginal buyer will determine the asset price in equilibrium.

Unanticipated good news also raises the marginal pessimist to h'_s *(see proofs in Appendix D.2.2)*. This implies that there will be a larger group of pessimists because the price of the risky asset is now too expensive from their point of view. With an increasing number of pessimists, there will be more disagreement in the economy. This also implies that the marginal buyer, h'_m will be agent who was previously a cautious optimist $h \in (h_s, h_b)$, given that former extreme optimists $h > h_b$ cannot acquire more Y because they have already spent all their wealth on Y in the initial equilibrium. However, the marginal optimists h'_b can be above or below h_b depending on the price of Y. If the price is too high, some of the former extreme optimists could become cautious optimists by reducing their holdings of Y. From their point of view, these assets are too expensive. In this scenario, the marginal optimists h'_b will shift up to be above h_b. However, if the price is not so expensive relative to the improvement in R, some cautious optimists will evolve into extreme optimists by spending all their wealth on Y. In this scenario, the marginal optimist will lie below h_b, and there will be more extreme optimists in the economy.

The equilibrium is, thus, determined by the system of two equations, Eq(4.7) and Eq(4.9), below. Eq(4.7) demonstrates the market-clearing condition, which equates supply and demand for a risky asset. The right-hand side shows aggregate revenue. The first term represents the revenues of cautious optimists $h'_s < h < h'_m$, who are reducing their holding of Y as, in their views, it is too expensive relative to fundamental improvement. The second term represents the revenues for pessimists $h_s < h < h'_s$, who previously were cautions optimists, but turn out to be

pessimists by selling all their holdings of Y . The left-hand side shows the aggregate expenditure. The first term represents the expenditure of cautious optimists $h'_m < h < h'_b$, who choose to increase their holding of Y . The second term represents the expenditures of new extreme optimists who will spend all they can afford for Y .

$$\int_{h'_m}^{h'_b} p'(y'^h - y^h)dh + \int_{h'_b}^{h_b} dh = \int_{h'_s}^{h'_m} p(y^h - y'^h)dh + \int_{h_s}^{h'_s} p' y^h dh \tag{4.7}$$

where,

$$\begin{aligned} h'_b &= \left(\frac{p' - R'}{1 - R'}\right)\frac{1}{p'} \\ h'_s &= \frac{p' - R'}{1 - R'} \end{aligned} \tag{4.8}$$

Eq(4.9) represents marginal buyer h'_m who is a cautious optimist indifferent between buying and selling of Y, and who will not participate in the market because he finds his current holding optimal; that is, $y'^{h'_m} = y^{h'_m}$. The left-hand side represents $y'^{h'_m}$, while the right-hand side represents$y^{h'_m}$.

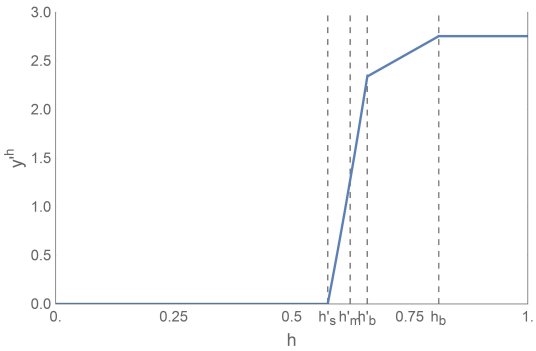
$$\frac{(h'_m(1 - R') - (p' - R'))}{(1 - p')(p' - R')}(w^{h'_m} + p' y^{h'_m}) = \frac{(h'_m(1 - R) - (p - R))}{(1 - p)(p - R)}(1 + p) \tag{4.9}$$

Plugging in $w^{h'_m}$ and $y^{h'_m}$ given by Eq(4.2) into Eq(4.9) gives:

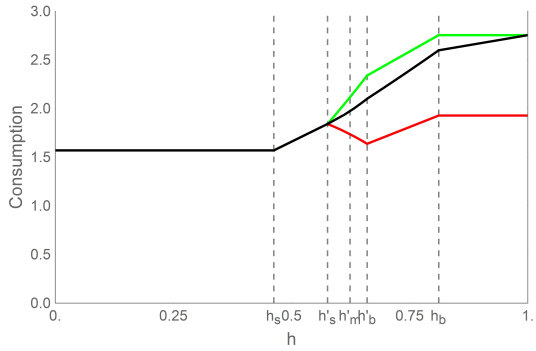
$$h'_m = \left(\frac{p' - R'}{1 - R'}\right)\left(\frac{2}{1 + p'}\right)$$

From a system of equations above with two equations and two unknowns, we can solve for the risky asset price, p' and the marginal buyer h'_m. We can then determine the marginal optimist, h'_b, and the marginal pessimist, h'_s, accordingly. For $R' = 0.7$, we get $p' = 0.873$, $h'_m = 0.624$, $h'_b = 0.660$, and $h'_s = 0.576$. Figure 4.5a shows the allocation of Y across agents. Figure 4.5b shows consumption plans in U and D of all agents. It is also interesting to see the re-distribution of wealth after the good news, as shown in Figure 4.6.[3]

[3] A detailed simulation result is provided in *Appendix D.2.3.*



(a) Distribution of risky asset holdings



(b) Consumption plans in U(Green) and D(Red), and expected consumption (Black)

Figure 4.5: No leverage economy - a good news equilibrium

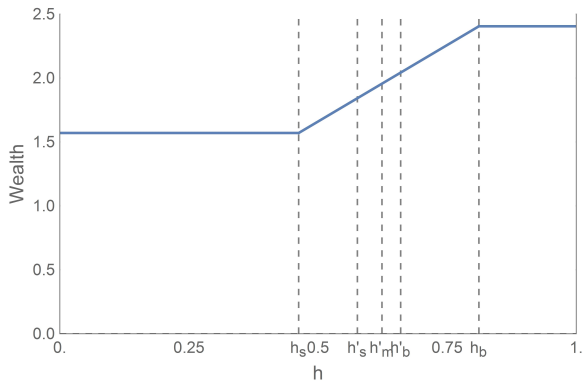


Figure 4.6: No leverage economy - wealth re-distribution with good news

4.3.3 Equilibrium with the reversal of good news

Suppose, the common belief about the risky asset quality is later discovered, and that its worst payoff is equal to R'' where $R'' < R'$. This is, to some extent, similar to the situation when the credit default swap (CDS) agreements were introduced into the market, and the market prices of mortgages backed the securities being discovered. Observing the sharp drop in the ABX Index in 2007, investors started to lower their valuation of a mortgage-backed security (MBS). Initially, as the investors had little information about a given MBS quality, they needed to rely on ratings provided by CRAs. The ratings were then shown to have been inflated, and, subsequently, these were revised down. As a rating usually indicates the level of risk attached to a given security, we can think of the rating in terms of the worst possible payoff of a risky asset, which could represent R in this model.

As shown in Figure 4.5a, the 'good news' equilibrium re-allocates the holdings of cash and of risky assets to be w'^h and y'^h, respectively. As a result, the demand functions for risky assets, given the arrival of bad news (or the reversal of good news), y''^h, could be explained by Eq(4.2), where, $w_e^h = w'^h$ and $y_e^h = y'^h$. We can then show the characterisation of equilibrium below; this is followed by numerical example at the end.

Bad news results in a drop in the risky asset price to p'' as, in equilibrium, wealth is being transferred to less optimistic agents, $h < h'_m$, who previously sold Y given good news. In response to a drop in asset quality and a lower risky asset price, there is a group of optimists who want to reduce their holdings of risky assets, and to supply Y into the market, while a group of less-optimistic agents, who preserve cash given good news ,are willing to acquire more Y. This less-optimistic group finds a drop in price of Y to be an opportunity for investment. The low valuation provided by these less-optimistic buyers exacerbates risky asset prices further in equilibrium.

There exists the unique marginal buyer h''_m, who is indifferent between buying and selling Y, that is, $y''^{h''_m} = y'^{h''_m}$. From Eq(4.2), it can be shown that the agents' wealth after re-distribution $w'^h + p''y'^h$ is decreasing with h for $h \in (h'_s, h'_b)$. This is because the risky asset demand (and supply) given good news $y'^h - y^h$ is increasing (and decreasing) with h. The numerical example of re-distribution of wealth is provided in Figure 4.8. In this case, although the price of Y drops, the marginal buyers shift up further to h''_m because the agents below h''_m will be buying Y, while the agents above h''_m will be reducing their holdings of Y. The agent $h'_m < h < h''_m$ will be acquiring additional Y.

A similar logical argument, as in *Appendix D.2.2*, can be applied to show that unanticipated bad news will lower the marginal pessimist to $h''_s < h'_s$. Intuitively, lower marginal pessimist implies that less-optimistic agents are more willing to hold Y as, from their point of view, the risky asset is relatively cheap compared to the drop in its worst payoff. However, the marginal optimists h''_b can be above or below h'_b depending on p''.

The equilibrium is determined by the system of two equations, Eq(4.10) and Eq(4.11) below. Eq(4.10) demonstrates the market-clearing condition, which equates supply and demand for risky assets. The left-hand side shows aggregate revenue. The first terms represent revenues of agents who were previously extreme optimists $h'_b < h < h''_b$,, who become cautious optimists by selling parts of Y. The second term represent cautious optimists $h''_m < h < h'_b$ who are reducing their holding of Y as, in their views, its price is too expensive. The right-hand side shows aggregate expenditure. The first term represents expenditure of cautious optimists

$h'_s < h < h''_m$, who are increasing their holdings of Y. The second terms represent the expenditures by new cautious optimists $h''_s < h < h'_s$, who see bad news as an opportunity to acquire additional Y.

$$\int_{h'_b}^{h''_b} p''(y' - y'')dh + \int_{h''_m}^{h'_b} p''(y' - y'')dh = \int_{h'_s}^{h''_m} p''(y'' - y')dh + \int_{h''_s}^{h'_s} p''(y'' - y')dh \tag{4.10}$$

where,

$$h''_b = (\frac{p'' - R''}{1 - R''})(\frac{1}{p''})$$
$$h''_s = \frac{p'' - R''}{1 - R''}$$

Eq(4.11) represents marginal buyer h''_m, who is indifferent between buying and selling Y, and who will not participate in the market because he finds his current holding optimal that is $y''^{h''_m} = y'^{h''_m}$. The left-hand side of the equation represents $y''^{h''_m}$, which is given by applying Eq(4.2) to this equilibrium where the marginal buyer's wealth is represented by $w'^{h''_m} + p''y'^{h''_m}$. The right-hand side represents $y'^{h''_m}$.

$$\frac{(h''_m(1 - R'') - (p'' - R''))}{(1 - p'')(p'' - R'')}(w'^{h''_m} + p''y'^{h''_m}) = \frac{(h''_m(1 - R') - (p' - R'))}{(1 - p')(p' - R')}(w^{h''_m} + p'y^{h''_m}) \tag{4.11}$$

which implies

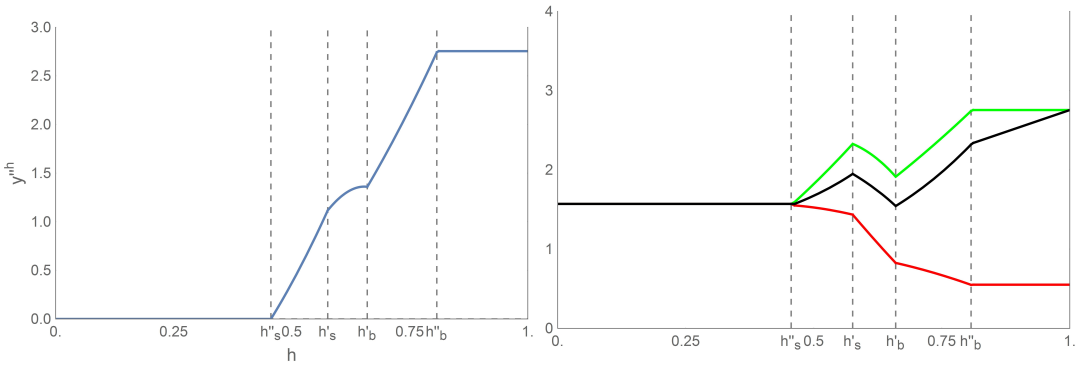
$$h''_m = \frac{-R' - p''(1 + R'(-2 + R'')) + R'' + p'(1 - p''R' + (-2 + p'' + R')R'')}{(p' - p'')(-1 + R')(-1 + R'')}$$

From a system of equations above with two equations and two unknowns, we can solve for the risky asset price p'' and the marginal buyer h''_m. We can then determine the marginal optimist h''_b and the marginal pessimist h''_s, accordingly. For $R'' = R = 0.2$, we get $p'' = 0.565$, $h''_m = 0.626$, $h''_b = 0.807$, and $h''_s = 0.456$. Figure 4.7a shows allocation of Y across agents. Figure 4.7b shows consumption plans in U and D of all agents. It is also interesting to see the re-distribution of wealth after the good news, as shown in Figure 4.8. Figure 4.9 summarises the distribution of agents in all equilibria of the non-leverage economy.[4]

The reversal of good news leads to new allocations of Y across agents, and to a drop in the price of Y. The impact on wealth will be perceived most by the agents who bought the risky asset given good news. Indeed, the price of Y becomes even

[4] A detailed simulation result is provided in *Appendix D.2.3.*

lower than the price in the initial equilibrium, though the asset quality is reversed to the same quality, as in the initial equilibrium. This amplification mechanism arises mainly from the risk-averse assumption as the downward sloping supply curve is exhibited by the optimists who need to reduce their holdings of Y, even though a lower price will affect their wealth and lessen their incentive to hold Y.



(a) Distribution of risky asset holdings

(b) Consumption plans in U(Green) and D(Red), and expected consumption (Black)

Figure 4.7: No-leverage economy - a 'good news' reversal equilibrium

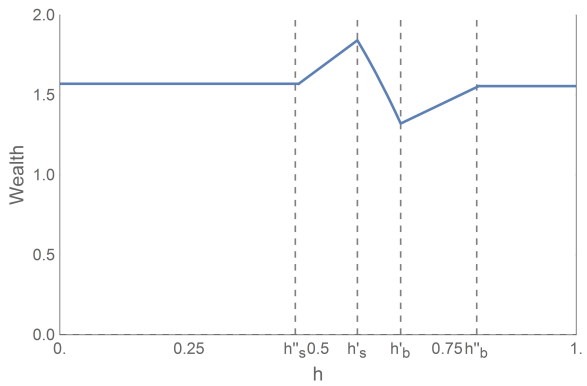


Figure 4.8: No-leverage economy - wealth re-distribution with 'good news' reversal

4.4 Leverage economy

In this economy, agents can borrow to finance their investment of risky assets. As a result, the agents who were extreme optimists in the previous economy without leverage, who would like to consume more in U, but were constrained by their endowment, can now shift their consumption from D to U by borrowing to buy

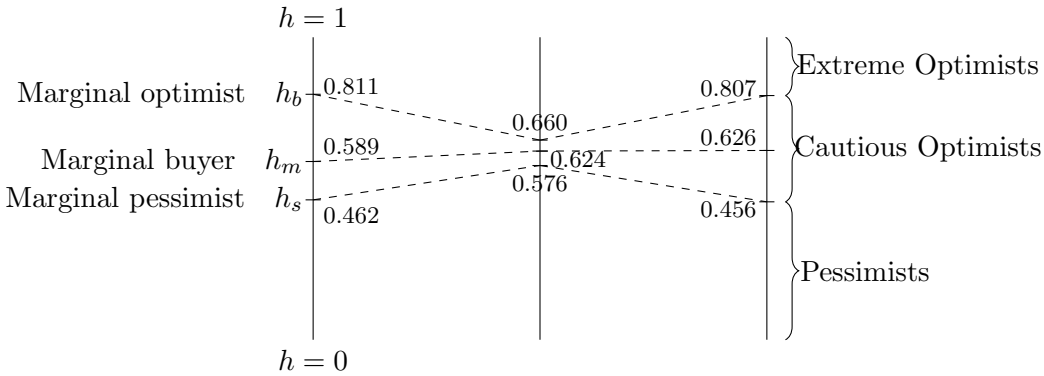


Figure 4.9: No-leverage economy - agent types in three equilibria
Given: R=0.2; R'=0.7; R"=0.2

additional Y. In particular, borrowing expands the interior set of agents, who can equalise their marginal utility of cash and of risky asset. As we shall see later, all agents except the pessimist can obtain their optimal solution with equalised marginal utility because (1) risk-averse agents always prefer strictly positive consumption in D as long as their belief on the probability attached to D is not zero, and (2) borrowing with constraint $b^h \leq y^h R$ allows them to vary their consumption in D down to zero when the constraint is binding (all Y payoff in D will be paid as a promise). The implication here is that the the borrowing constraint for agents will not be binding except for the top optimist $h = 1$, who believes with certainty that the risky asset payoff will be 1, and who will discard all consumption in D. Since borrowing provides more cash in hand for optimists, who give high valuation to the risky asset, the price will increase more than had been the case in the economy without leverage.

In terms of the welfare implication, the market is still incomplete and inefficient. Agents can only move their consumption from D to U by borrowing that is backed by risky asset Y. However, they cannot transfer their consumption from U to D. Thus, only optimists can price in (though imperfectly) their willingness to buy risky assets, while pessimists have no mechanism to shift their consumption to D, and cannot price in their willingness to lower consumption further in D. As a result, the price is inflated due to market incompleteness. Fostel and Geanakoplos [2012] show that financial innovations to complete the market could lead to a crash once the economy shifts from a 'tranching' economy to an Arrow-Debreu economy.

Suppose, agent h is currently holding (endowed with) a portfolio containing w_e^h and y_e^h units of risk-free W, and risky asset Y, respectively. The current borrowing of the agent is supposed to be b_e^h. A positive value of b^h implies borrowing,

while a negative value implies lending. The agent h maximises his expected utility by choosing a portfolio containing risk-free asset w^h and risky asset y^h, and by borrowing amount b^h. The problem can be simplified by defining net borrowing, $\omega^h = w^h - b^h$. For simplicity, we further assume that an agent will not borrow until a safe asset is used up; that is, an agent will borrow ($b^h > 0$) iff $\omega^h < 0$.

Appendix D.3.1 shows a private decision of an agent. In this leverage economy, there will be three types of agents: (1) extreme optimists with $h \geq (\frac{p-R}{1-R})(\frac{1}{p})$, who will buy all they can afford for Y, and sell a promise to (borrow from) the other groups of agents; (2) cautious optimists with $\frac{p-R}{1-R} \leq h < (\frac{p-R}{1-R})(\frac{1}{p})$, who will hold a combination of Y and W, and will not participate in borrowing; and (3) pessimists with $h \leq (\frac{p-R}{1-R})$, who prefer fully insured consumption and will sell all their holdings of Y. Only durable consumption good W will be in the pessimists' portfolio.

Now, we can identify the marginal optimist h_b, who will be indifferent in holding W and in borrowing. (It should be noted that cash and borrowing are perfect substitutions given an interest rate equal to zero.) The marginal pessimist h_s will be indifferent in holding Y. The demand function for risky assets given unanticipated news could be described by:

$$y^h(\omega_e^h, y_e^h) = \begin{cases} \frac{(h(1-R)-(p-R))}{(1-p)(p-R)}(\omega_e^h + py_e^h) & if \quad h \geq h_s \\ \\ 0 & if \quad h < h_s \end{cases} \tag{4.12}$$

The ω could be described by Eq(4.13) below. It should be noted that for extreme optimists, $h > h_b$, $-y^h R \leq \omega^h < 0$, as they hold only Y, and also sell a promise in equilibrium. For cautious optimists and for pessimists, $h < h_b$, $\omega^h > 0$, as they will hold positive amounts of W in their portfolios.

$$\omega^h(\omega_e^h, y_e^h) = \begin{cases} \frac{((p-R)-hp(1-R))}{(1-p)(p-R)}(\omega_e^h + py_e^h) & if \quad h \geq h_s \\ \\ \omega_e^h + py_e^h & if \quad h < h_s \end{cases} \tag{4.13}$$

where,

$$\begin{aligned} h_s &= \left(\frac{p-R}{1-R}\right) \\ h_b &= \left(\frac{p-R}{1-R}\right)\frac{1}{p} \end{aligned} \tag{4.14}$$

4.4.1 Initial equilibrium

This section describes initial equilibrium in the leverage economy, in which agents are allowed to borrow by selling a promise. The initial endowment for all agents is $w_e^h = 1$ and $y_e^h = 1$. Each agent's demand functions for risky asset y^h follow Eq(4.12). The characterisation of equilibrium can then be presented below; this is followed by a numerical example at the end.

Given the endowment and demand function as described in Eq(4.12), we can easily see that there will be a unique marginal buyer h_m, who will be indifferent between buying and selling risky assets, and who will not participate in the market because he finds his endowment optimal; that is, $y_m^h = y_e^h = 1$. In equilibrium, all agents $h < h_s$ prefer perfectly insured consumption, and will sell all their endowment of Y for durable consumption good W. All agents $h > h_s$ will hold a positive amount of risky asset y^h, which is increasing with h. In addition, there is a marginal buyer h, who is indifferent in holding W and in borrowing, where, $h_b = \frac{p-R}{1-R}$. All agents $h_b > h > h_s$ will hold a combination of W and Y without borrowing. Agents with $h > h_b$, who among all agents give highest valuation to Y, will buy all that they can afford, including borrowing to buy Y.

The initial equilibrium is thus characterised by the system of two equations below. Eq(4.15) equates supply and demand for risky asset Y, and demonstrates the market-clearing condition. The marginal optimist h_b and the marginal pessimist h_s are described in Eq(4.14). The right-hand side shows aggregate selling revenue by agents who lie below the marginal buyer h_m. The first term represents the revenue of cautious optimists, who prefer to reduce their holdings of Y. The second term represents the revenue of pessimists, who sell all their endowment of Y. The left-hand side shows aggregate buying expenditure by cautious optimists who lie above the marginal buyer h_m.

$$\int_{h_m}^{1} p(y^h - 1)dh = \int_{h_s}^{h_m} p(1 - y^h)dh + \int_{0}^{h_s} pdh \tag{4.15}$$

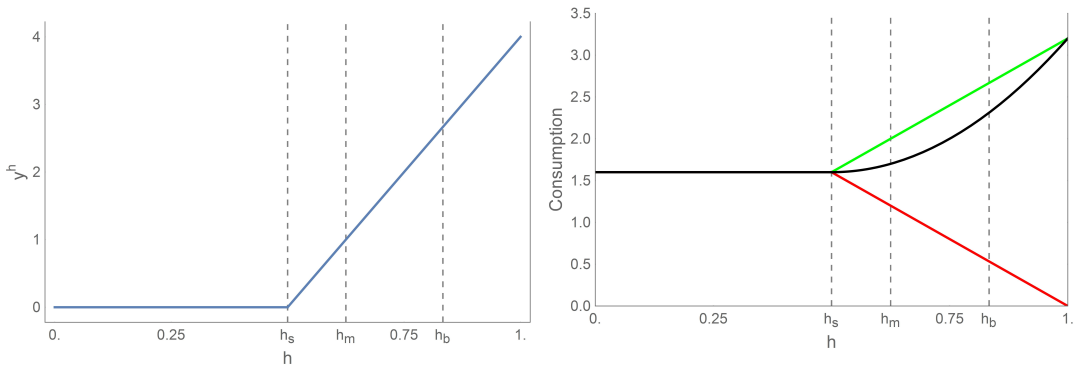
Eq(4.16) represents unique marginal buyer h_m, who is indifferent between buying and selling Y. The left-hand side of the equation shows demand for Y of the marginal buyer, which needs to be equal to his endowment of 1. We can also see that h_m is increasing in p. Compared to the non-leverage equilibrium, the leverage economy should evidence higher h_m as p is higher.

$$\frac{((1-R)h_m - (p-R))(1+p)}{(1-p)(p-R)} = 1 \tag{4.16}$$

With a system of equations above with two equations and two unknowns, we can solve for the risky asset price p and the marginal buyer h_m. In equilibrium, the market-clearing price $p \in (1, R)$ is increasing with R and could be described by Eq(4.17). The marginal buyer h_m could also be specified by Eq(4.18). We can then determine the marginal optimist h_b, and the marginal pessimist h_s, accordingly. For $R = 0.2$, we get $p = 0.600$, $h_m = 0.625$, $h_b = 0.833$, and $h'_s = 0.5$. Figure 4.10a shows the allocation of Y across agents. Figure 4.10b shows the consumption plans in U and D of all agents.[5] Compared to the non-leverage economy, the leverage economy evidences higher h_m, meaning that risky asset holding is more concentrated in group of top optimists, and h_m corresponds to a higher p.

$$p = -1 + R + \sqrt{2 - R^2} \tag{4.17}$$

$$h_m = \left(\frac{p - R}{1 - R}\right)\left(\frac{2}{1 + p}\right) \tag{4.18}$$



(a) Distribution of risky asset holdings

(b) Consumption plans in U(Green) and D(Red), and expected consumption (Black)

Figure 4.10: Leverage economy - an initial equilibrium

4.4.2 Equilibrium with unanticipated good news

The initial equilibrium has allocational effects, and provides heterogeneity in cash holdings w^h and asset holding y^h. The allocation of Y after the initial equilibrium is shown in Figure 4.10a. The demand functions for risky assets of the agents y'^h could thus be explained by Eq(4.12), where $w_e^h = w^h$ and $y_e^h = y^h$. We can then show the characterisation of equilibrium below, followed by a numerical example at the end.

[5] A detailed simulation result is provided in *Appendix D.3.3*.

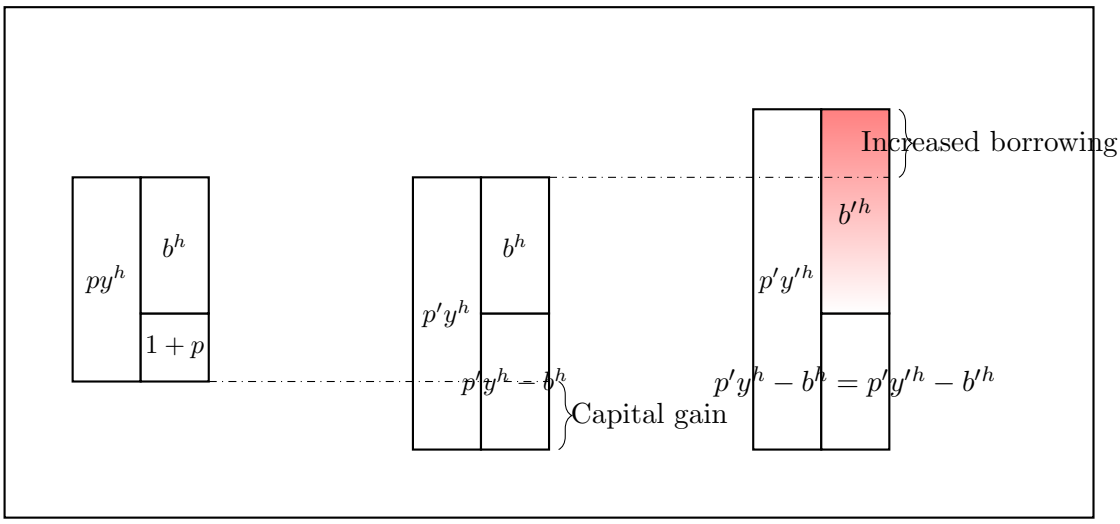


Figure 4.11: Leverage economy - extreme optimists' balance sheet with good news

With an unanticipated positive shock to asset quality, where R is shifted to R' and $R < R' < 1$, it is obvious that risky asset prices will rise in response to good news about asset quality. Suppose, the price rises to p'. This improves the wealth of the optimists with $h > h_m$, who previously bought Y in initial equilibrium, and who are willing to acquire more Y given good news. Again, the demand from these group of optimists increases for two reasons: first, the worst risky asset payoff is believed by all agents to be improved to R'. Second, with 'mark to market' accounting, capital gains from the price rise for risky assets on their balance sheets will raise their equity value. The extreme optimists will also borrow more to buy the risky assets. Figure 4.11 illustrates the increase in borrowing that leads to balance-sheet expansion. In addition, the price will rise further simply because these optimistic agents, who become richer, and who value risky assets more than others do. However, with leverage allowing an agent to borrow more to buy risky assets, the price could be so high that the number of extreme optimists will be squeezed, and the number of marginal optimists will grow. This scenario is shown later in the numerical examples. The system is prone to default because there are only few extreme optimists with extremely high leverage; this makes the price skyrocket.

Given y'^h is increasing in h, there exists the unique marginal buyer y'^{h_m}, who is indifferent between buying and selling Y; that is, $y'^{h_m} = y^{h_m}$. From Eq(4.12), we can easily observe that the agents' wealth after re-distribution, $w^h + p'y^h$, is increasing with h as y^h is increasing in h for $h \in (h_s, h_b)$. The numerical example of a re-distribution of wealth is provided in Figure 4.13. Indeed, the marginal buyer will shift upwards to h'_m as price increases because, given that the marginal buyer

is indifferent between buying and selling Y, it's clear that the marginal buyer will determine the asset price in equilibrium.

Unanticipated good news also raises the marginal pessimist to h'_s. The proof is shown in *Appendix D.2.2*. With an increasing number of pessimists, there will be more disagreement in the economy. The marginal optimist h'_b can be above or below h_b depending on the price of Y. If the price is too high, some formerly extreme optimists could become cautious optimists by reducing their holdings of Y for W. In the leverage economy, it is also possible that the marginal buyer h'_m lies above h'_b, where the demand for Y comes only from formerly extreme optimists who are wealthier and borrowing more in order to acquire additional Y. In this scenario, agent $h \in (h'_b, h_b)$ will sell some Y in order to reduce the borrowing, thus, increasing consumption in D. In equilibrium, $h > h'_m$ will buy additional risky assets. Agents $h_s < h < h'_m$ will sell part of their risky assets.

The equilibrium is thus determined by the system of two equations, Eq(4.19) and Eq(4.21), below. Eq(4.19) equates supply and demand for risky assets and demonstrate the market-clearing condition. The right-hand side shows the aggregate revenue, which comes from the pessimists and the cautious optimists who lie below the marginal buyer h'_m. It should be noted that as h_s increases, there are agents $h \in (h_s, h'_s)$, who were previously cautious optimists, but become pessimists as the asset price rises to a level that is too high for them. Also, it is also possible that h'_m lies above h'_b, implying that the extreme optimists with $h'_b < h < h'_m$ still borrow to buy Y, but just decrease their borrowing and holdings of Y. The left-hand side shows aggregate expenditure, which could come purely from the extreme optimists given $h'_m > h'_b$.

$$\int_{h'_m}^{1} p'(y'^h - y^h)dh = \int_{h'_s}^{h'_m} p(y^h - y'^h)dh + \int_{h_s}^{h'_s} p'y^h dh \tag{4.19}$$

where,

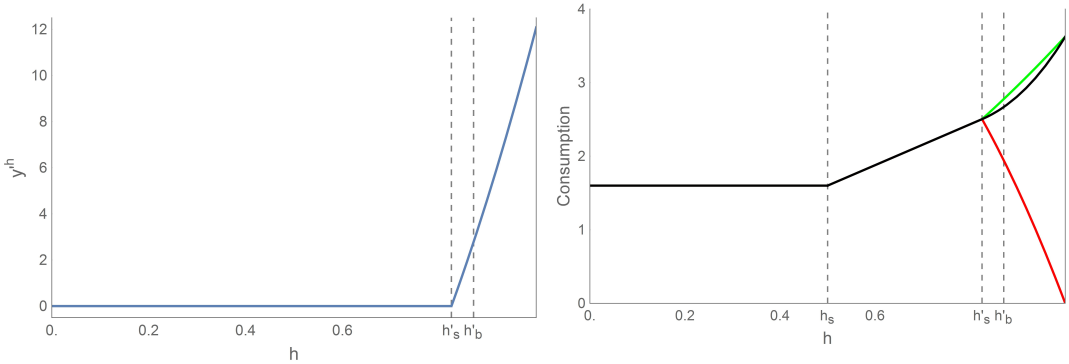
$$h'_s = \frac{p' - R'}{1 - R'} \tag{4.20}$$

Eq(4.21) represents a unique marginal buyer who is indifferent between buying and selling risky assets. As explained in the previous section, we should expect much higher marginal buyer when compared to non-leverage economy due to borrowing and h_m is increasing in p.

$$\frac{(h'_m(1 - R') - (p' - R'))}{(1 - p')(p' - R')}(\omega^{h'_m} + p'y^{h'_m}) = \frac{(h'_m(1 - R) - (p - R))}{(1 - p)(p - R)}(1 + p) \tag{4.21}$$

From a system of equations above with two equations and two unknowns, we can solve for the risky asset price p' and the marginal buyer h'_m. We can then determine the marginal optimist h'_b and the marginal pessimist h'_s, accordingly. For $R' = 0.7$, we get $p' = 0.948$, $h'_m = 0.874$, $h'_b = 0.871$, and $h'_s = 825$.[6]

Note that, in this case, $h'_m > h'_b > h_b$. Figure 4.12a shows allocation of Y across agents. Figure 4.12b shows consumption plan in U and D of all agents. It is also interesting to see the re-distribution of wealth after the good news as shown in Figure 4.13.



(a) Distribution of risky asset holdings

(b) Consumption plans in U(Green) and D(Red), and expected consumption (Black)

Figure 4.12: Leverage economy - a 'good news' equilibrium

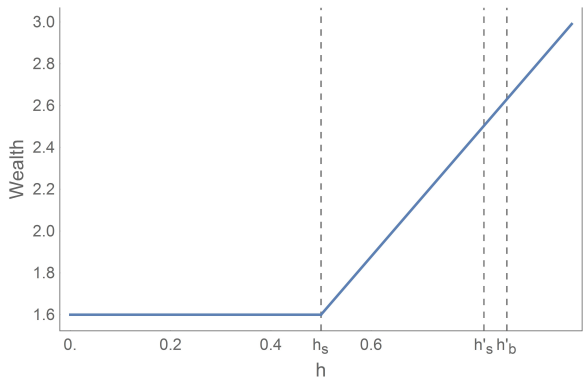


Figure 4.13: Leverage economy - wealth re-distribution with 'good news'

4.4.3 Equilibrium with the reversal of good news

In this section, we consider the situation in which the common belief on risky asset quality is revised down to from R' to R'', where $R'' < R'$. As previously shown

[6] A detailed simulation result is provided in *Appendix D.3.3*.

in Figure 4.12a, the 'good news' equilibrium re-allocates the holdings of cash and of risky assets to be w'^h and y'^h, respectively. As a result, given bad news (or the reversal of good news), the demand functions for risky assets of the agents y''^h could be explained by Eq(4.12), where, $w_e^h = w'^h$, and $y_e^h = y'^h$. We can then show the characterisation of equilibrium below; this is followed by a numerical example at the end.

Bad news results in a drop in the risky asset's price to p'' as, in equilibrium, wealth is being transferred to less-optimistic agents $h < h'_m$, who previously sold Y given good news. Similar to the situation a non-leverage economy, in response to a drop in the asset quality and the lower risky asset price given bad news, there is a group of optimists who want to reduce their holdings of risky asset, and to supply Y into the market, while a group of less-optimistic agents who preserve cash given good news are willing to acquire more of Y. These groups of agents see the drop in the price of Y as an opportunity for investment. Low valuation provided by these less- optimistic buyers exacerbates the risky asset price further in equilibrium. However, we should expect a lower equilibrium price in a leverage economy because the balance sheet mechanism is allowed to work.

There exists the unique marginal buyer y''^{h_m}, who is indifferent between buying and selling of Y, that is, $y''^{h_m} = y'^{h_m}$. From Eq(4.12), it can be shown that the agents' wealth after re-distribution, $w'^h + p''y'^h$, is decreasing with h, for $h > h'_s$. This is because, given a good news, the risky asset demand (supply) by these group of agents, $y'^h - y^h$, is increasing (decreasing) with h. As a result, loss is increasing in h while equity is decreasing in h, for $h > h'_s$, as leverage allows these agents to aggressively buy Y given good news. Thus, summarising the above statements, the extent of loss and equity corresponds to the amount of Y they bought given good news. The numerical example of re-distribution of wealth is provided in Figure (4.15). In this case, although the price of Y drops, the number of marginal buyers shifts up further to h''_m because agents below h''_m will be buying Y while the agents above h''_m will be reducing their holdings of Y. The agents $h'_m < h < h''_m$ will be acquiring additional Y.

A logical argument, similar to the one applied in *Appendix D.2.2*, can again be applied to show that the unanticipated bad news will lower the number of marginal pessimists to $h''_s < h'_s$. Intuitively, the lower number of marginal pessimists, given the news reversal, implies that some agents who were formerly pessimists will become cautious optimists who are willing to hold Y because, from their point of view, the risky asset is now cheap, even when there is a drop in its quality. However, the number of marginal optimists h''_b can be above or below h'_b, depending on p''.

In contrast to the situation in a non-leverage economy, in a leverage economy it is possible that the reversal of risky asset quality back to the initial quality R could potentially lead to bankruptcy and default. The default will occur if and only if $p'' < R'$. The default threshold could be defined as (see proof in *Appendix D.3.2*).

$$h_d = Min\left[\frac{(1-p'')(p'-R')}{(p'-p'')(1-R')}\,,1\right] \tag{4.22}$$

The equilibrium is thus characterised by the system of two equations below. Now, agents $h > h''_m$ will be selling Y while $h''_s < h < h''_m$ will be instead of buying Y because the price is so attractive for them. Eq(4.23) equates supply and demand for risky assets, and demonstrates the market-clearing condition. The left-hand side shows aggregate revenue, which comes from default optimists, whose risky assets will be seized and liquidated, and from some agents who were previously extreme optimists, who find that they need to reduce their holdings because of drop in asset quality and their wealth. These groups of agents lie above the marginal buyer h''_m. It should be noted that as the number of marginal buyer is increasing further, there are agents $h \in (h''_m, h''_d)$, who were previously extreme optimists, but become cautious optimists by selling some Y. The right-hand side shows aggregate expenditure by all agents $h''_s < h < h''_m$, who have safe asset and see investment opportunities in acquiring more risky asset.

$$\int_{h''_d}^{1} p''y'dh + \int_{h''_m}^{h''_d} p''(y'-y'')dh = \int_{h''_s}^{h''_m} p''(y''-y')dh \tag{4.23}$$

where,

$$h''_s = \frac{p''-R''}{1-R''} \tag{4.24}$$

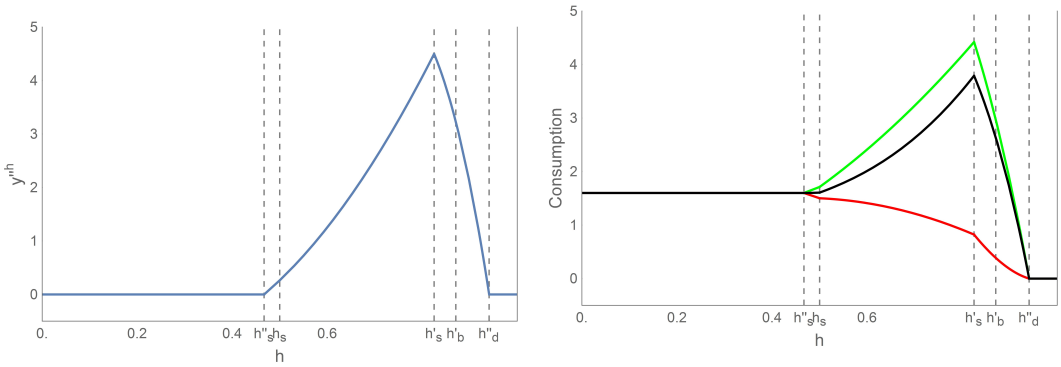
Eq(4.25) represents marginal buyer h''_m, who is indifferent between buying and selling Y and will not participate in the market because he finds his current holding optimal, that is, $y''^{h''_m} = y'^{h''_m}$. Based on Eq(4.12), the left-hand side represents $y^{h''_m}$ while the right-hand side represents $y^{h'_m}$.

$$\frac{(h''_m(1-R'')-(p''-R''))}{(1-p'')(p''-R'')}(w'^{h''_m}+p''y'^{h''_m}) = \frac{(h''_m(1-R')-(p'-R'))}{(1-p')(p'-R')}(w^{h''_m}+p'y^{h''_m}) \tag{4.25}$$

which implies

$$h''_m = \frac{-R'-p''(1+R'(-2+R''))+R''+p'(1-p''R'+(-2+p''+R')R'')}{(p'-p'')(-1+R')(-1+R'')}$$

From a system of equations above with two equations and two unknowns, we can solve for the risky asset price p'' and the marginal buyer h''_m. We can then determine the marginal optimist h''_b, the marginal pessimist h''_s, and the default threshold h''_d, accordingly. For $R'' = R = 0.2$, we get $p'' = 0.574$, $h''_m = 0.875$, $h''_b = 0.814$, $h''_s = 0.467$, and $h''_d = 0.941$.[7] Figure 4.14a shows allocation of Y across agents. Figure 4.14b shows consumption plans in U and D of all agents. It is also interesting to see the re-distribution of wealth after the good news, as shown in Figure 4.15. Figure 4.16 summarizes distribution of agents in all equilibria of the leverage economy.



(a) Distribution of risky asset holdings

(b) Consumption plans in U(Green) and D(Red), and expected consumption (Black)

Figure 4.14: Leverage economy - a 'good news' reversal equilibrium

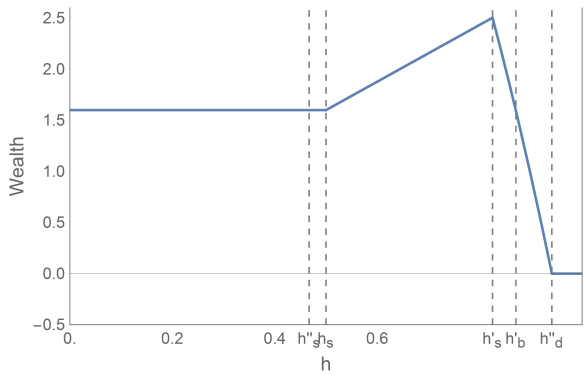


Figure 4.15: Leverage economy - wealth re-distribution with 'good news' reversal

[7] A detailed simulation result is provided in *Appendix D.3.3*.

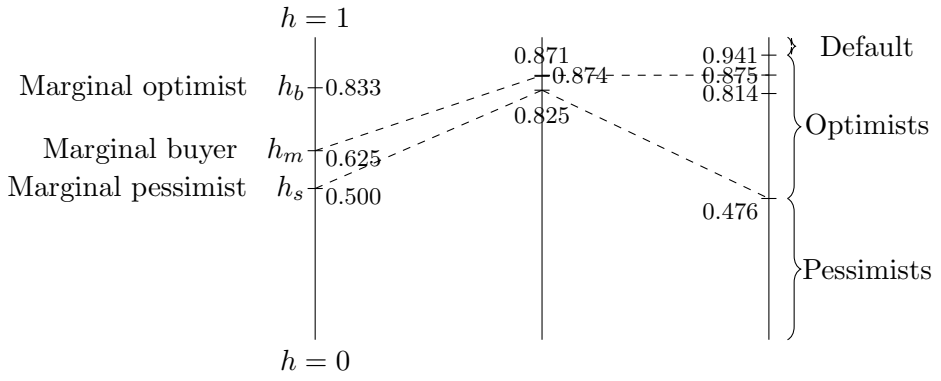


Figure 4.16: Leverage economy - agent types in three equilibria
Given: R=0.2; R'=0.7; R"=0.2

4.5 Discussion and conclusions

This paper's analysis shows insolvency and defaults in a heterogeneous beliefs framework. In particular, we introduced unanticipated 'good news' by raising the worst payoff of the risky asset, which results in increasing the asset price and in redistributing wealth from the pessimists to the optimists. Expansions in the equity base of the optimists encourage them to take more risk by buying more assets for their portfolios. However, the unexpected reversal of good news shows the asymmetry of capital gains and losses, which may trigger insolvency and defaults. Several key takeaways from this paper merit emphasis:

First, in contrast to Geanakoplos [2010], and to Fostel and Geanakoplos [2012], this paper's assumption related to risk aversion provides interesting insights. First, the assumption allows the effect of pecuniary externalities to work, leading to the *amplification* of the reversal shocks, as well as the *asymmetry of capital gain and loss*. Equity-based expansions of the ranks of the optimists encourage them to take on excessive risk by acquiring additional risky assets for their portfolios. By contrast, the pessimists are, however, risk averse and, so their demand is price sensitive (not perfectly elastic). This also leads to fragile disagreement (or 'scary good news') when the risks are highly concentrated in the portfolio of the optimists.[8] The numerical examples provided in Section 4.3.2 regarding the no leverage economy, in which borrowing is not allowed indicate that the top 43 percent of the agents hold all the risky assets in the economy. When leverage is possible in the leverage economy, as shown in the example provided in Section 4.4.2, only 18 percent of the agents in the economy hold all the risky assets. However, the reversal of good news leads optimists

[8]Geanakoplos [2010] emphasized the role of scary bad news in causing higher asset volatility.

to adjust their portfolios by reducing their asset holdings, while the pessimists who preserve cash in the 'good news' equilibrium see low asset prices as opportunities for investment.

Second, if the reversal of good news takes the price lower than the worst payoff that agents had expected when good news arrived $p'' < R'$, then *the default is unavoidable*. The defaults involve negative equity of the optimistic defaulting borrowers.[9] This is in contrast to the Binomial No-Default Theorem (as in Fostel and Geanakoplos [2015]), which suggests the 'no default' outcome in the economy with a single financial asset serving as collateral in a static binomial model. The assumption of unanticipated shocks as a 'zero probability event', supported by the 'neglect risk' phenomenon (as in Gennaioli et al. [2012]), plays a role for the defaults in this model.

Third, this paper provides the default outcomes that *endogenise the extent of systemic defaults*. The size of a group of agents who default in this model is endogenously determined by the asset price in responses to the reversal of good news. In particular, the numerical results show that the number of systemic defaults is larger with bigger shocks. For example, the shock of $R' - R = 0.6 - 0.2 = 0.4$, results in default for top 2.7 percent of agents. If, however, the shock increases to $R' - R = 0.7 - 0.2 = 0.5$, as provided in the numerical results above, the top 6 percent of agents will default. These examples could reflect the credit rating agencies' highly favourable pre-crisis ratings, which were subsequently revised, as signaled by the ABX-HE indices. The same size of systemic defaults in the pre-crisis US subprime mortgage securitisation markets could have been amplified further with various spillover mechanisms within the financial system; this situation contributed to the Great Recession.

[9] Although Geanakoplos [2010] also gave the results of the disappearance of optimists given bad news, such disappearance doesn't directly affect the lenders. This is because, in these circumstances, the collateral will be seized and the equity of those optimists are just wiped out as the loan traded in equilibrium involves a promise that rules out default outcomes.

Appendix A

Illiquidity and Insolvency: Background

A.1 Amplification mechanism and endogenous risks

There are various mechanisms within the financial system that could amplify a small shock into a major event. Such amplification mechanisms could lead to boom-bust cycles. In good times, excessive risk taking usually manifests itself as an increase in asset prices, allowing institutions to take additional risks thanks either to an endogenous increase in equity and net worth, or to relaxed asset price-dependent constraints imposed either by regulators (who may set capital requirements) or by the institutions themselves (who may set borrowing constraints). With heightened ability to borrow, such institutions will increase leverage and expand their balance sheets.[1]

Additional risks taken during good times leads, in bad times, to a sharp drop in asset prices when the deleveraging process begins, especially in an economy characterised by limited access to external finance and an incomplete insurance market. A powerful amplification mechanism of shocks is usually propelled by 'feedback loops between prices and distress selling of assets'. As mentioned in Shim and Von Peter [2007], this distress selling could be triggered either by insolvency shocks as a result of excessive risk taking, or by liquidity shocks. Distress selling emphasises the role of pecuniary externalities as an engine of amplification, in which the effects of correlated actions on price are not internalised in the private decision within a

[1] During the expansionary phase of the financial cycle, financial intermediaries choose to have common exposure on credit and liquidity risks (or to correlate their risks) due to 'strategic complementarities'; that is, their payoff from holding common risks increases with the number of other agents taking the same strategies.

competitive market framework.

On the seller's side, various mechanisms could generate asset fire sales, including forced selling due to liquidity needs, idiosyncratic financial constraints, maintenance costs (e.g., Moore [2013]), or even the agent's preference to liquidate assets for consumption (e.g., Davila and Korinek [2017]). A survey of the substantial literature on amplification mechanisms is provided in the survey of financial frictions by Brunnermeier et al. [2012]. Another fast-growing branch of the literature, lying between detailed partial equilibrium models of banking and 'institution-free' general equilibrium, focuses on adding 'financial frictions' to dynamic stochastic general equilibrium (DSGE) models cast in the Gali/Woodford tradition of modern macroeconomics.[2]

On the buyer's side, there are many ways to generate the usual downward-sloping demand curve for risky assets. In an economy with uncertainty, the introduction of a risk-averse buyer who demands a greater risk premium to hold more risky assets, could result in a sharp drop in asset prices (e.g., Shin [2010]). In a model with uncertainty and heterogeneous beliefs, agents who have different valuations of risky assets can also generate a downward sloping curve, even if the agents are risk neutral. The disappearance of the most optimistic agents when given bad news results in an equilibrium price determined by investors whose valuation of the asset is low (e.g., Geanakoplos [2010]).[3] In an economy without uncertainty, the imperfectly elastic demand curve could be generated simply by buyers' access to a concave technology (e.g., Davila and Korinek [2017]). While there is an extensive corpus of literature on amplification and asset fire sales, the key early studies contributing to the area include Bernanke and Gertler [1990], Shleifer and Vishny [1992], and Kiyotaki and Moore [1997].

With these amplification mechanisms, it would appear that the price of risky assets can exhibit what Zeeman [1974] and Arnold [1984] refer to as 'catastrophic' behaviour: highly asymmetric responses to symmetric movements in exogenous forces. Zeeman sought to explain the gradual rise in equity prices in a boom, followed by the sharp fall in the subsequent crash by the differences in behaviours between 'bulls and bears'.

[2]We make no attempt to analyse these contributions; but, for a good illustration, with a concise summary of other papers in this burgeoning field, the reader may be referred to Coimbra and Rey [2017].

[3]In Geanakoplos [2010]), most optimistic agents disappear because their equity is wiped out in a bad state of the world, as these agents prefer zero consumption and borrow to buy risky assets by promising all their wealth in the bad state. However, there is no default in this model, as lenders still receive the full promise from these optimistic agents.

A.2 Pecuniary externalities and inefficiency

Although the amplification of shocks as a result of pecuniary externalities and distress selling manifests itself in bad times, the endogeneity of asset prices also applies in normal or good times. This is because agents take prices as a given and do not internalise either collective effects or joint behaviour. Market outcomes are therefore not welfare efficient, as asset fire sales normally result in indirect wealth transfer via price mechanisms from more productive agents to less productive agents. Geanakoplos and Polemarchakis [1986] discuss the constrained inefficiency of an economy with incomplete markets; this has led to a recently growing body of literature in this area.

In particular, Davila and Korinek [2017] illustrate two different mechanisms of pecuniary externalities leading to constrained inefficient allocations. The first mechanism, the so-called 'distributive externality', emerges when agents have different marginal rates of substitution (MRS), which could be a result of binding constraints or incomplete markets in which the privately optimal allocation is not possible, given the availability of asset span in the economy. The second mechanism, the so-called 'collateral externality', emerges when the price-dependent financial constraints (e.g., borrowing constraints, regulatory capital requirements, Value at Risk, and incentive-compatible constraints) are binding. While the distributive externality mechanism can exhibit either over- or under-investment depending on differences in MRS across agents, the collateral externality mechanism always shows over-investment. For both mechanisms, the ex-post constrained Pareto improvements could be achieved by ex-ante redistribution of resources among agents to allow ex-post wealth transfer through price changes, thus providing a room for government intervention with a view to achieve welfare improvement.

Various studies emphasise the role of distributive externality in leading to constrained inefficiency. Lorenzoni [2008] demonstrates that, in an economy with uncertainty, competitive financial contracts with limited levels of commitment by lenders can result in excessive borrowing ex-ante and excessive volatility ex-post due to incomplete insurance contracts. In particular, excessive borrowing causes excessive contraction in investment and risky asset pricing. Subsequently, Moore [2013] points out that uncertainty and financial contracts are not necessary for generating constrained inefficiency. The need to maintain capital by paying maintenance costs using proceeds from selling assets can also trigger fire sales and can lead to market failure. In both examples, a simple ex-ante transfer between two groups of agents can bring about a Pareto improvement. Caballero and Krishnamurthy [2003], Ko-

rinek [2010], and Stein [2012] are among the other studies that show over-borrowing can lead to excessive leverage, resulting in constrained inefficiency.

A further strand of literature emphasises the role of collateral externalities. Greenwald and Stiglitz [1986] show that the competitive equilibrium of an economy with asymmetric information also exhibits constrained inefficiency in a setting in which agents are subject to price-dependent constraints. Bianchi [2011] shows over-borrowing as a result of collateral externalities, occurring when borrowers are subject to collateral constraints that depends on an endogenous price. It could be that these two mechanisms of externalities interact with each other, as shown in Korinek and Simsek [2016].

As a key driver of market failures, pecuniary externalities create a role for government intervention aimed at welfare improvement. Moreover, the potential for market failure also justifies macro-prudential regulation, a subject widely discussed after the subprime crisis. While micro-prudential regulation is necessary to ensure individual stability, or to lessen the extent of the principal agent problem, it is not sufficient for the maintenance of system-wide stability, and it may have unintended consequences. The key channels of such unintended consequences work through pecuniary externalities. Various policy tools aiming at lessening the extent of amplification and buoyant asset prices include debt-to-income (DTI) ratios, loan-to-value (LTV) ratios, countercyclical capital buffers, and capital surcharges based on systemic risk contribution. In particular, Goodhart et al. [2013] discuss various prudential regulations that can improve welfare by mitigating the effects of asset fire sales.

Appendix B

Pecuniary Externalities in 'Bank Run' Models

B.1 Equilibrium without aggregate uncertainty

In this appendix, we assume that each consumer can have access to both short-term and long-term assets.

Efficient Solution

Let us assume that the planner has complete information about the economy. In particular, the planner knows exactly who is an early consumer, and who is a late consumer. The planner chooses the allocation between short-term and long-term assets at date 0, and is only subject to the feasible constraint. The planner's problem is

$$\max_{y} \quad \lambda U(C_1) + (1-\lambda)U(C_2) \tag{B.1}$$

subject to:

$$\begin{aligned} \lambda C_1 &\leq y \\ (1-\lambda)C_2 &\leq (1-y)R + (y - \lambda C_1) \end{aligned}$$

As it is never optimal to carry over any of short-term assets from date 1 to date 2, the constraints are binding and become $\lambda C_1 = y$ and $(1-\lambda)C_2 = (1-y)R$. That is $(1-\lambda)C_2 = (1-\lambda C_1)R$. The solution to this problem is:

$$\frac{U'(C_1)}{U'(C_2)} = R$$

Where $C_1 = \frac{y}{\lambda}$ and $C_2 = \frac{(1-y)R}{1-\lambda}$

Given that $R > 1$ and $U''(C) < 0$, the optimality condition implies $C_2 > C_1$. An individual's consumption will be higher or lower, depending on whether he is early or late consumer. This also allows us to relax the assumption that the planner knows exactly who is an early consumer, and who is a late consumer because there is no incentive for a late consumer to withdraw early. The solution is incentive compatible.

Another interesting insight is that λ disappears from the optimality condition. This is because, from the planner's point of view, the effect of an increase in λ on the marginal rate of substitution, in which the planner will place greater value on an early consumer's consumption, is offset by the cost of satisfying early consumer by giving up the return from holding long-term assets resulting in lower consumption for a late consumer. The left term of the equation below shows the marginal rate of substitution of the planner, and the right term shows the ratio of liquidity price from the planner's point of view.

$$\frac{\partial E[U]/\partial C_1}{\partial E[U]/\partial C_2} = \frac{\lambda R}{1-\lambda}$$
$$\frac{\lambda U'(C_1)}{(1-\lambda)U'(C_2} = \frac{\lambda R}{(1-\lambda)}$$

Another interpretation of the disappearance of λ in the optimality condition is that the increase in λ results in the increasing marginal utility of liquidity, which is cancelled out by the increasing marginal cost of liquidity. Since both constraints are binding, the planner will always provide return from short-term assets to early consumers, and returns from long-term assets to late consumers. At optimality, the marginal utility of liquidity (or short-term assets) needs to be equal to the marginal cost of liquidity which is forgone return that results in less consumption for late consumers. The left term of the equation below refers to the marginal benefit of liquidity (or early consumption), while the right term refers to the marginal cost of liquidity (or early consumption).

$$\frac{\partial E[U]}{\partial C_1} = -\frac{\partial E[U]}{\partial C_2}\frac{\partial C_2}{\partial C_1}$$
$$\lambda U'(C_1) = \lambda R U'(C_2)$$

Note that, when $U(C) = ln(C)$, R plays no role in determining the optimal liquidity because binding constraints imply that the optimal allocation doesn't re-

quire each asset's returns to be shared between both types of consumers. In other words, returns from short-term assets will be delivered to early consumers, and the returns from long-term assets will be delivered to late consumers. The left-hand side of the equation below shows that R obviously plays no role in the marginal benefit of liquidity as long as $R > 1$. In addition, for the marginal cost of liquidity, the right-hand side points out that the effect of R is also cancelled out. This is because while an increase in R raises the opportunity cost of liquidity when a late consumer has to forgo the payoff of long-term assets, it increases the marginal utility of late consumption when the planner values late consumer more.

$$\begin{aligned}
\frac{\partial E[U]}{\partial C_1} &= -\frac{\partial E[U]}{\partial C_2}\frac{\partial C_2}{\partial C_1} \\
\lambda(\frac{\lambda}{y}) &= -((1-\lambda)\frac{(1-\lambda)}{(1-y)R})\ \ (\frac{-\lambda R}{1-\lambda}) \\
y &= \lambda
\end{aligned}$$

Autarky

As each consumer is unable to trade assets, and has to consume the returns generated by his portfolio, a consumer's problem is:

$$\max_{y} \quad \lambda U(C_1) + (1-\lambda)U(C_2) \tag{B.2}$$

subject to:

$$\begin{aligned}
C_1 &\leq y \\
C_2 &\leq y + (1-y)R
\end{aligned}$$

Both constraints are binding because there is no incentive for a consumer to hold a short-term asset from date 1 to date 2, and because a consumer consumes all the return of long-term asset at date 2 due to the assumption of the neoclassical utility function. The solution to this problem is:

$$\frac{U'(C_1)}{U'(C_2)} = \frac{(1-\lambda)(R-1)}{\lambda}$$

If we assume the risk aversion is equal to 1 in the standard CRRA function, then $U(C) = ln(C)$. The solution is $(y, C_1, C_2) = (\frac{\lambda R}{R-1}, \frac{\lambda R}{R-1}, (1-\lambda)R)$

Market

A consumer now can trade at date 1. Once the uncertainty is revealed, the market allows an early consumer to sell long-term assets in exchange for consumption at date 1, and it allows a late consumer to buy long-term assets in exchange for consumption at date 2. Let P denotes the price of a long-term asset at date 1. A consumer's problem is:

$$\max_{y} \quad \lambda U(C_1) + (1-\lambda)U(C_2) \tag{B.3}$$

subject to:

$$C_1 \leq y + P(1-y)$$
$$C_2 \leq ((1-y) + \frac{y}{P})R$$

It is obvious that both constraints are binding. In equilibrium, $P = 1$ because if $P > 1$, the long-term asset dominates the short-term asset at date 0, and no one wants to hold short-term assets. As a result, $P = 0$ which contradicts $P > 1$. If $P < 1$, the short-term asset dominates long-term asset at date 0, and no one wants to hold long-term assets at date 0. As a result, consumers will bid price up to P = R. This again cannot be an equilibrium as $P = R$ contradicts $P < 1$.

At this price, a consumer is indifferent between long- and short-term assets at date 1, and the consumer's portfolio choice is irrelevant. The consumer's consumption will be $(C_1, C_2) = (1, R)$. There are multiple equilibria as they can be anything, as long as the aggregate liquidity or short-term asset in the economy is equal to λ (eg. equlibrium when the λ fraction of consumers hold only short-term assets in their portfolios, and the $(1-\lambda)$ fraction of consumers hold only long-term assets; or equilibrium when each consumer has the same portfolio choice, which is the λ fraction of his endowment invested in short-term assets and the $(1-\lambda)$ fraction invested in long-term assets.

Bank

Given that there is only uncertainty about the time preference of consumption of the consumer, the banking solution is efficient and similar to the planner's solution.

B.2 Proof of efficient allocation

The efficient allocation is characterised by considering the problem of a planner who seeks to maximise utility of the representative consumer by making investment decision on short- and long-term assets at date0, and to distribute the proceeds from investments to consumers who value consumptions at date 1 and date 2. The investment decision of the planner is given y and $1-y$, which represent units of short-term assets and long-term assets, respectively, invested per a unit of good that is the endowment of consumer at date 0. Consumption bundle of representative consumer is described by $C = \{C_L, C_{2L}, C_{1H}, C_{2H}\}$. The planner's problem is described by:

$$\max_{y, C_{1L}} \pi(\lambda_H U(C_{1H}) + (1-\lambda_H)U(C_{2H})) + (1-\pi)(\lambda_L U(C_{1L}) + (1-\lambda_L)U(C_{2L})) \quad \text{(B.4)}$$

subject to:

$$\lambda_H C_{1H} = y \qquad (1-\lambda_H)C_{2H} = (1-y)R$$
$$\lambda_L C_{1L} \leq y \qquad (1-\lambda_L)C_{2L} = (1-y)R + (y - \lambda_L C_{1L})$$

Note that the constraint at date 2 is always binding because $U(.)$ is strictly concave; therefore, all proceeds from investment will be used up. In addition, the constraint at date 1 in state H is always binding. Let me assume that the constraints in both states are binding. Given the optimal holding of liquidity, y, the efficient allocation ***across time*** provides $C_{1H} < C_{1L}$ and $C_{2H} > C_{2L}$ as $\lambda_H > \lambda_L$. Then, if $C_{1L} < C_{2L}$, the consumption bundle becomes $C_{1H} < C_{1L} < C_{2L} < C_{2H}$. If $C_{1L} \geq C_{2L}$, the consumption bundle becomes $C_{1H} < C_{1L} = C_{2L} < C_{2H}$ as the consumer will be better off by postponing consumption from date1 to date2 until $C_{1L} = C_{2L}$. In the latter case, the constraints is not binding in state L, and consumers receive a fully insured consumption bundle regardless of consumers' types equal to the expected consumption $y + (1-y)R$. Therefore, the efficient allocation across time ensures binding constraints in state H, while whether the constraint in state L is or is not binding depends on the optimal C_{1L} and C_{2L}. The result is intuitive because it has never been optimal for the planner to carry forward liquidity from date 1 to date 2 in both states; the planner can do better by allocating more goods into long-term assets. and thereby can enjoy higher returns and higher expected consumption. Also, the efficient solution, where $C_{2S} \geq C_{1S}$, provides an incentive-compatible outcome in the sense that there is no incentive for a late consumer to pretend to be early consumer. Therefore, we can relax the assumption that the planner needs to know investors' types. The solution can be summarised as follows:

First, let me consider the efficient allocation *across time* for a given value of y. The Lagrangian is;

$$\mathscr{L} \;=\; \pi(\lambda_H U(C_{1H})+(1-\lambda_H)U(C_{2H}))+(1-\pi)(\lambda_L U(C_{1L})+(1-\lambda_L)U(C_{2L}))+\mu(\frac{y}{\lambda_L}-C_{1L})$$

$$\begin{array}{llll} \text{FOC w.r.t.} & C_{1L}; & (1-\pi)\lambda_L(U'(C_{1L})-U'(C_{2L}))-\mu=0 & ,\ C_{1L}>0 \\ \text{FOC w.r.t.} & \mu; & y-\lambda_L C_{1L}\geq 0 & ,\ \mu\geq 0 \end{array}$$

$$\begin{array}{lll} \textit{Case1}\ \mu=0,\ y-\lambda_L C_{1L}>0 & U'(C_{1L})-U'(C_{2L}=0 \quad \text{and} & y-\lambda_L C_{1L}>0 \\ \textit{Case2}\ \mu>0,\ y-\lambda_L C_{1L}=0 & (1-\pi)\lambda_L(U'(C_{1L})-U'(C_{2L})=\mu \quad \text{and} & y-\lambda_L C_{1L}=0 \end{array}$$

Therefore, the first-order conditions to the problem that are necessary and sufficient can be written as:

$$U'(C_{1L})-U'(C_{2L}) \;\geq\; 0$$

with the complementary slackness condition:

$$[U'(C_{1L})-U'(C_{2L})][y-\lambda_L C_{1L}] \;=\; 0$$

Note that $U'(C_{1L}) \geq U'(C_{2L})$ implies $C_{1L} \leq C_{2L}$. As a result, the incentive constraint is automatically satisfied. In particular, case1 implies that $C_{1L}=C_{2L}=y+(1-y)R$, and that the budget constraint at date 1 in state L is not binding. Case2 implies that $C_{1L}<C_{2L}$ and $C_{1L}=\frac{y}{\lambda_L}$ and $C_{2L}=\frac{(1-y)R}{(1-\lambda_L)}$, and that the budget constraint is binding. Therefore, the optimality of the consumption bundle across this state given y can be written as:

$$\begin{array}{ll} C_{1H}=\dfrac{y}{\lambda_H} & C_{2H}=\dfrac{(1-y)R}{(1-\lambda_H)} \\ C_{1L}=\min\{\dfrac{y}{\lambda_L},\ y+(1-y)R\} & C_{2L}=\max\{\dfrac{(1-y)R}{1-\lambda_L},\ y+(1-y)R\} \end{array}$$

Second, let us consider the efficient allocation ***across state*** to obtain the optimal holding of liquidity. The planner's problem now becomes:

$$\begin{aligned} \max_{y} \quad & \pi(\lambda_H U(\frac{y}{\lambda_H})+(1-\lambda_H)U(\frac{(1-y)R}{(1-\lambda_H)})) \\ & +\ (1-\pi)(\lambda_L U(\min\{\frac{y}{\lambda_L},\ y+(1-y)R\})\ +\ (1-\lambda_L)U(\max\{\frac{(1-y)R}{1-\lambda_L},\ y+(1-y)R\})) \end{aligned}$$

Since the objective function in continuous and $U(.)$ is strictly concave, we can obtain y^*, and it is unique. The solution to the planner's problem consists of a portfolio choice of y^* and consumption bundle $C^* = \{C^*_L, C^*_{2L}, C^*_{1H}, C^*_{2H}\}$ such that optimal holding of liquidity y^* is given by:

$$\frac{\pi U'(C_{1H}) + (1-\pi)U'(C_{1L})}{\pi U'(C_{2H}) + (1-\pi)U'(C_{2L})} = R$$

where the consumption bundle satisfies:

$$C^*_{1H} = \frac{y}{\lambda_H} \qquad C^*_{2H} = \frac{(1-y)R}{(1-\lambda_H)}$$

$$C^*_{1L} = \min\{\frac{y}{\lambda_L},\ y + (1-y)R\} \qquad C^*_{2L} = \max\{\frac{(1-y)R}{1-\lambda_L},\ y + (1-y)R\}$$

Suppose $U(C) = ln(C)$, the equilibrium is characterised by:

$$y* = \begin{cases} \frac{((R-1)(1+\pi\lambda_H)+\pi) - \sqrt{(R-1)(1+\pi\lambda_H)+\pi)^2 - 4\pi\lambda_H R(R-1)}}{2(R-1)} & \text{if} \quad R \leq \hat{R} \\ \pi\lambda_H + (1-\pi)\lambda_L & \text{if} \quad R > \hat{R} \end{cases}$$

where

$$\hat{R} \equiv \frac{(1-\lambda_L)(\pi\lambda_H + (1-\pi)\lambda_L)}{\lambda_L(1 - (\pi\lambda_H + (1-\pi)\lambda_L))}$$

represents the critical long-term asset return such that if R is lower than the critical value, $C_{1L} = C2L$, and $C_{1L} \leq C_{2L}$ otherwise. In the latter case, the optimal holding of liquidity in the economy is equal to the expected probability of being an early consumer. Again, the consumption bundle is described by:

$$C^*_{1H} = \frac{y^*}{\lambda_H} \qquad C^*_{2H} = \frac{(1-y^*)R}{(1-\lambda_H)}$$

$$C^*_{1L} = \begin{cases} y^* + (1-y^*)R & if \quad R \leq \hat{R} \\ \frac{y^*}{\lambda_L} & if \quad R > \hat{R} \end{cases} \qquad C^*_{2L} = \begin{cases} y^* + (1-y^*)R & if \quad R \leq \hat{R} \\ \frac{(1-y^*)R}{(1-\lambda_L)} & if \quad R > \hat{R} \end{cases}$$

B.3 Comparative statics in equilibrium without run (for $R = 2;\ \pi = 0.35;\ \lambda_H = 0.85;\ \lambda_L = 0.80;\ U(C) = ln(C)$)

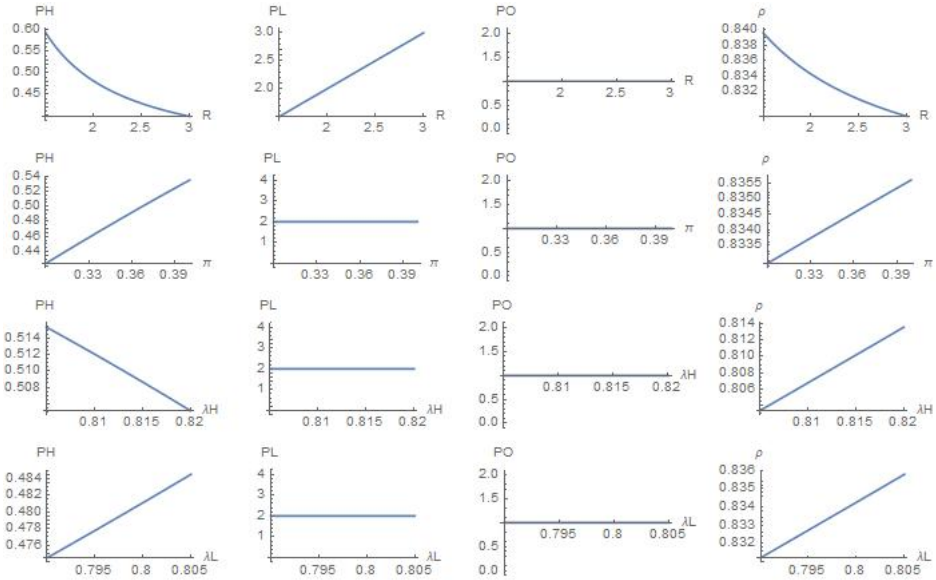


Figure B.1: Illustration of relationship between the characterisation of equilibrium without run (P_H, P_L, P_0, ρ) and parameters $(R, \pi, \lambda_H, \lambda_L)$

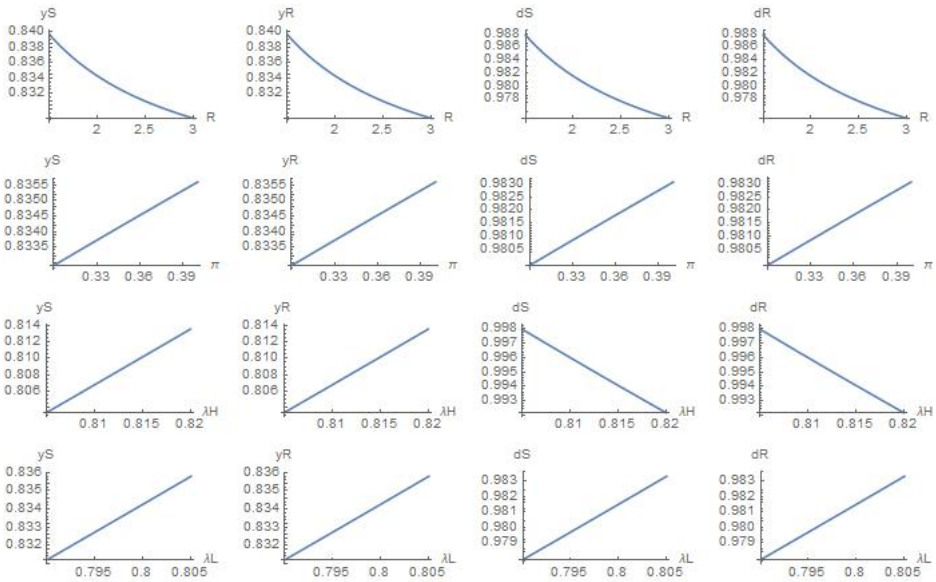


Figure B.2: Illustration of relationship between bank's choices (y^S, y^R, d^S, d^R) and parameters $(R, \pi, \lambda_H, \lambda_L)$

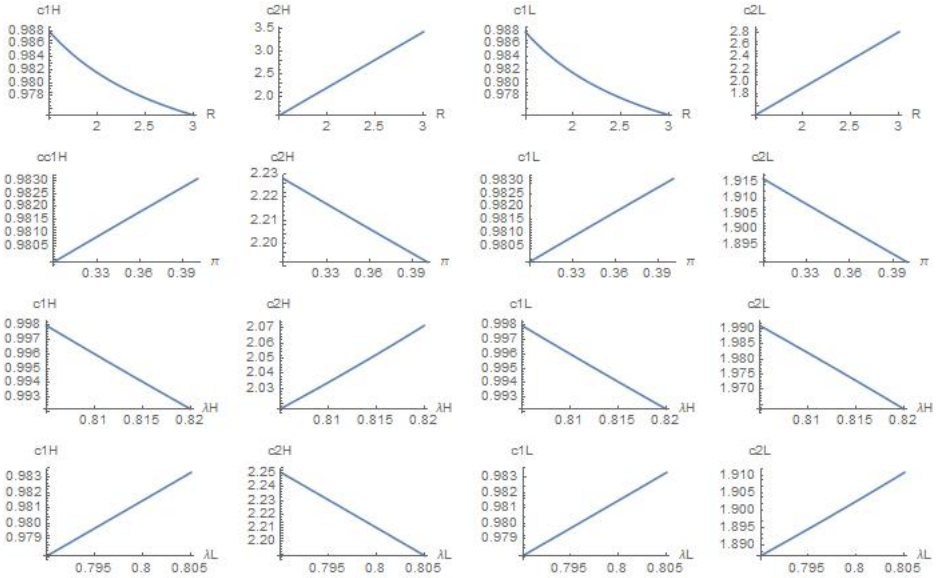


Figure B.3: Illustration of relationship between safe bank's consumption plan $(C^S_{1H}, C^S_{2H}, C^S_{1L}, C^S_{2L})$ and parameters $(R, \pi, \lambda_H, \lambda_L)$

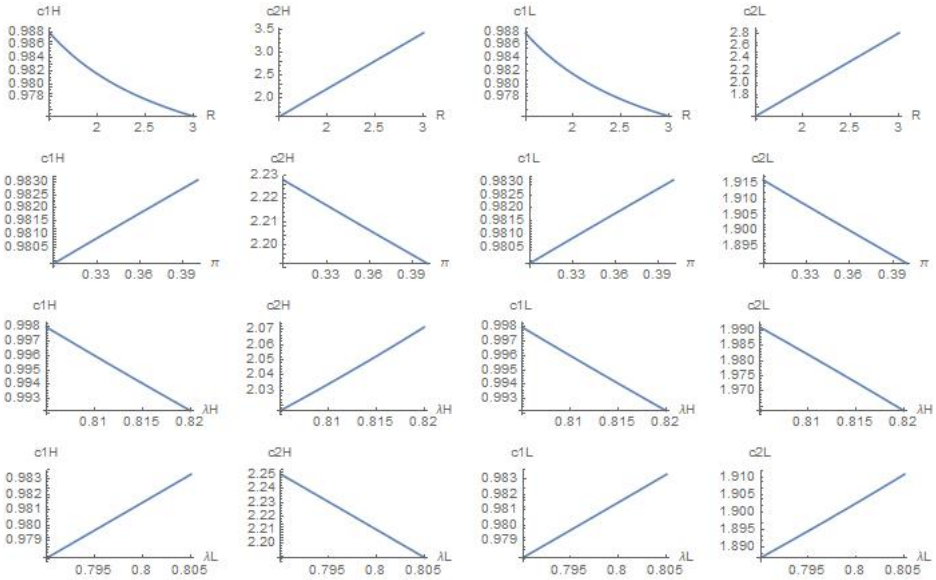


Figure B.4: Illustration of relationship between risky bank's consumption plan $(C^R_{1H}, C^R_{2H}, C^R_{1L}, C^R_{2L})$ and parameters $(R, \pi, \lambda_H, \lambda_L)$

B.4 Comparative statics in equilibrium with run (for $R = 2;\ \pi = 0.35;\ \lambda_H = 0.85;\ \lambda_L = 0.80;\ U(C) = ln(C)$)

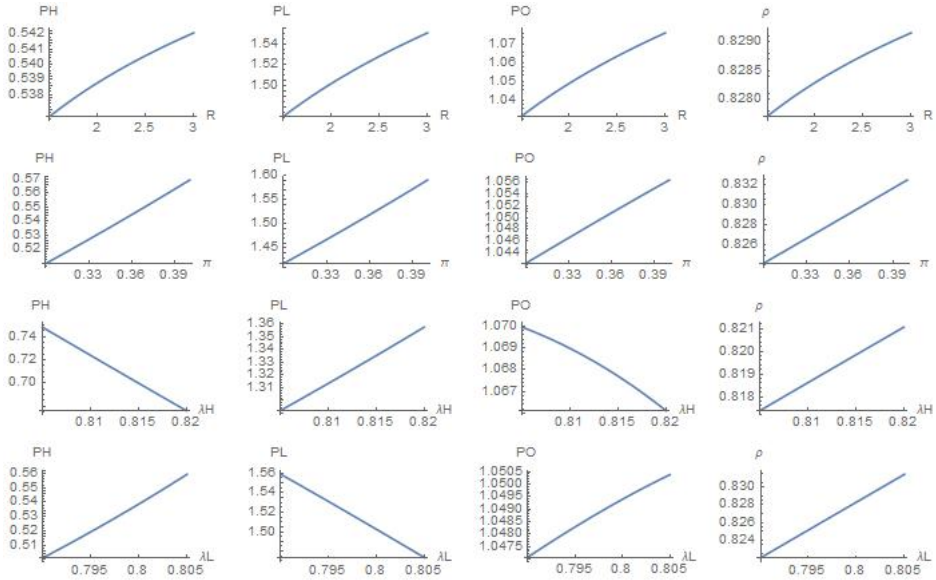


Figure B.5: Illustration of relationship between the characterisation of equilibrium with run (P_H, P_L, P_0, ρ) and parameters $(R, \pi, \lambda_H, \lambda_L)$

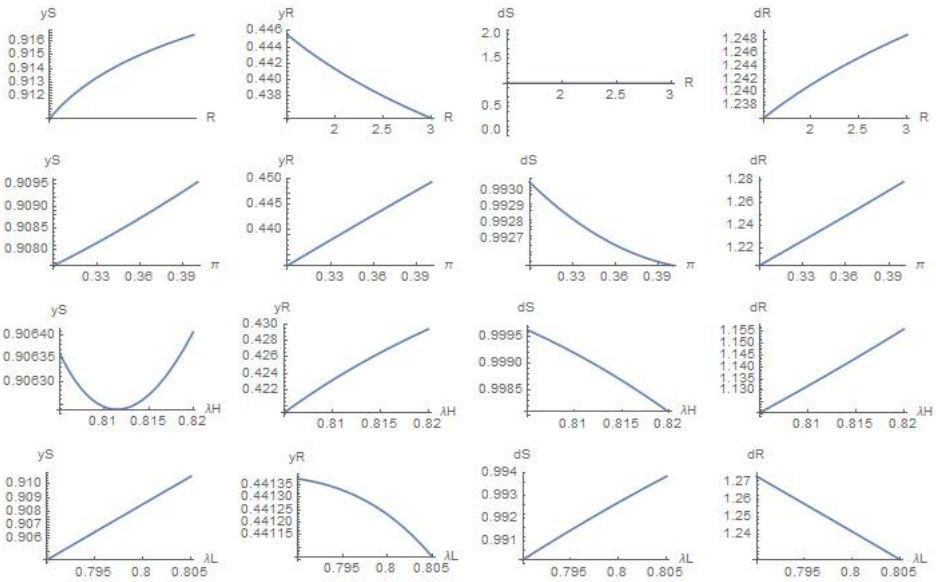


Figure B.6: Illustration of relationship between bank's choices (y^S, y^R, d^S, d^R) and parameters $(R, \pi, \lambda_H, \lambda_L)$

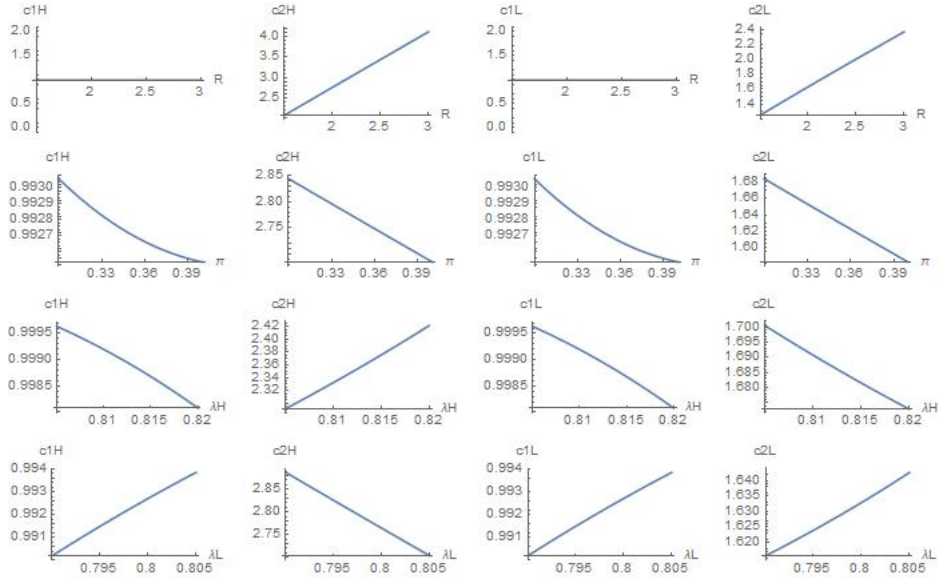


Figure B.7: Illustration of relationship between safe bank's consumption plan $(C^S_{1H}, C^S_{2H}, C^S_{1L}, C^S_{2L})$ and parameters $(R, \pi, \lambda_H, \lambda_L)$

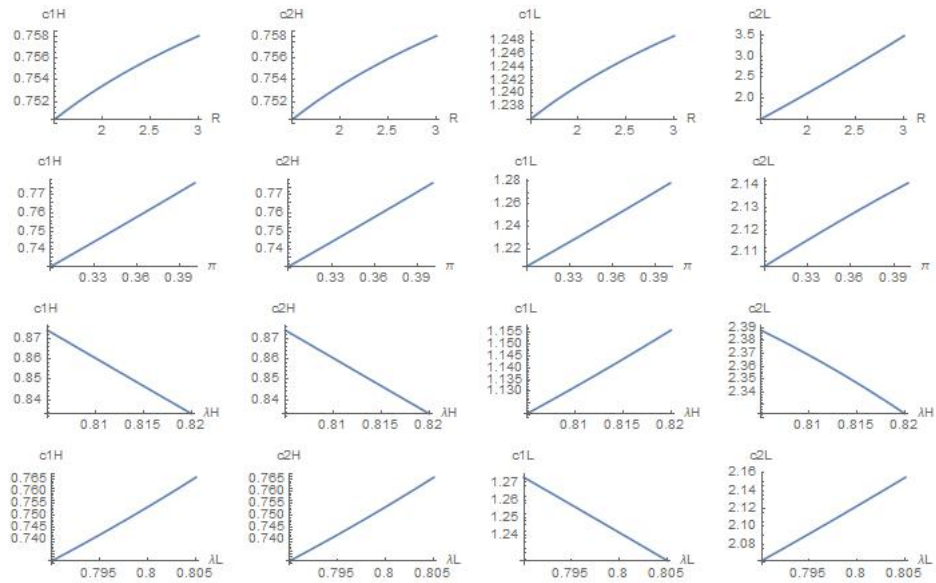


Figure B.8: Illustration of relationship between risky bank's consumption plan $(C^R_{1H}, C^R_{2H}, C^R_{1L}, C^R_{2L})$ and parameters $(R, \pi, \lambda_H, \lambda_L)$

Appendix C

Subprime Assets and Financial Crisis: Theory, Policy and the Law

C.1 Effect of exogenous shocks raising asset quality

How to determine the ultimate effect taking these externalities into account? To derive this formally, note first how, with mark-to-market gains following an improvement in asset quality, the previously binding VaR constraint is relaxed and the new equity level of active investors is given by:

$$e' = p'y - (q - z)y \tag{C.1}$$

where $p'y$ denotes assets revalued at new prices and $(q - z)y$ is pre-existing level of borrowing. The increased equity value allows active investors to take more risky assets onto their balance sheets. These expand until VaR constraint is again binding, so:

$$e' = p'y' - (q' - z')y' \tag{C.2}$$

where y' denotes the new optimal holdings of risky assets held by active investors, and the improved asset quality is indicated by $q' > q, z' < z$.

For the holding of risky assets by active portfolio managers, equations Eq(C.1) and Eq(C.2) imply the expanded level of asset holdings following such favourable shocks is:

$$y'_A = y_A \left(1 + \frac{(q' - q) - (z' - z)}{p' - q' + z'}\right)$$

or

$$y'_A - y_A = y_A \left(\frac{(q' - q) - (z' - z)}{z' - \frac{z'^2}{3\tau}(1 - y'_A)}\right) \tag{C.3}$$

given the market clearing condition, $y'_P + y'_A = 1$.

Appendix D

Heterogeneous Beliefs, Endogenous Risk, and Crash

D.1 The sequence of financial innovation and the Great Recession

Fostel and Geanakoplos [2012], hereafter F&G, put forward the possibility that the mortgage boom and bust crisis of 2007-2009 might have been caused by financial innovation. They suggest that that 'the astounding rise in subprime and Alt A leverage ... together with the remarkable growth of securitization and tranching... raised the prices of underlying assets such as houses and mortgage bonds ... [and] that the introduction of Credit Default Swaps(CDS) in 2005, and 2006 brought those prices crashing down.'

By contrast, in this section of our paper, we use the Foster and Geanakoplos heterogeneous beliefs approach to provide a general equilibrium perspective, where the 'good news' corresponds to an increase in optimism, and the 'bad news' corresponds to the introduction of CDS swaps which allow for shorting of the risky asset.

F&G assumed an endowment economy populated by risk-neutral agents with identical endowments, but heterogeneous beliefs, who can trade their initial (equal) holdings of a safe and a risky asset (whose payoffs in the 'up' state and 'down' states are specified as (1,1) and (1,$R < 1$), respectively) in a competitive general equilibrium. The focus is on how financial innovations that affect trading possibilities will alter the price of the risky asset.

Given their heterogeneous prior beliefs about the probability q of the good outcome for the risky asset, participants can be ordered in terms of increasing op-

timism, where $0 \leq h \leq 1$ denotes the 'height' of optimism, so $q = h$; for example, or perhaps, $q = 1 - (1-h)^2$. Without leverage, the solution for the market-clearing price, p, of the risky asset, conditional on any given degree of optimism, is given by equating the revenue from sales of the asset to the aggregate expenditure by the more optimistic agents:

$$p = (1-h)(1+p) \quad \text{so} \quad p = (1-h)/h$$

as shown by the schedule NN in Figure D.1. The reason that the market-clearing price is negatively related with h comes from the fact that only agents to the right of the NN schedule hold the risky asset. So, the same amount of risky assets has to be held by a smaller fraction of the population as h increases.

However, the assumed heterogeneity of beliefs implies a positive relationship between the beliefs of the marginal investor and the market price. Specifically, assuming that the market price of the risky asset, $p = q + (1-q)R$, reflects the beliefs of the marginal investor, which defines the schedule labelled as Marginal Buyers Valuation in the Figure. If, for example, $q = h$, then the schedule is simply:

$$p = (1-h) + hR$$

As shown in Figure D.1, The intersection of this schedule with NN at EN identifies the marginal holder of the risky asset, and determines the equilibrium price for the no-leverage economy.

Thus, the story in F&G begins in an equilibrium without leverage (EN) when the pessimists sell all the risky assets to the optimists and hold all cash, with total payoffs in both 'up' and 'down' states being shared by the optimists and the pessimists. But then, with leverage, the pessimists sell their risky assets to optimists and hold all cash, with the optimists consuming more in the 'up' state.

Formally, with leverage, 'expenditure' on the risky asset by the more-optimistic agents only has to cover the downside risk, so the market-clearing equation becomes:

$$p - R = (1-h)(1+p)$$

This increases the price conditional on any given degree of optimism, as indicated by the schedule LL; and in the new equilibrium EL the price of the risky assets will increase as a smaller number of optimists will be able to hold them all.

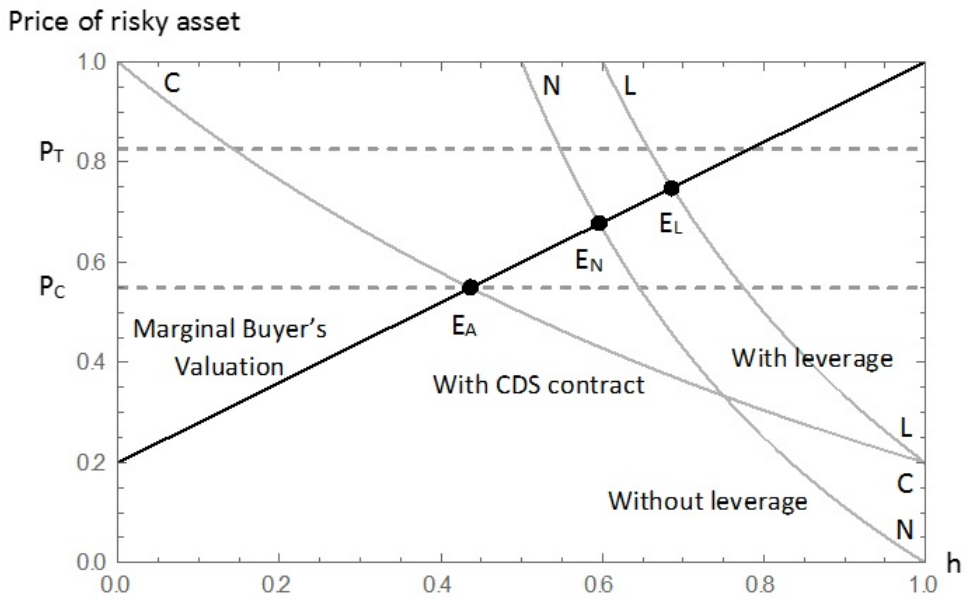


Figure D.1: Four equilibria in Fostel-Geanakoplos model, assuming $q = h$

F&G went on to determine the impact that 'tranching' the risky asset will have. They argued that giving the risky asset holders the ability to sell the 'down tranche' effectively gives the most-optimistic participants access to an Arrow security for the 'up' state (where the risk assets have high return), while the pessimists who buy it gain access to an Arrow security for the 'down state' (where the risky assets have low return R).[1] The effect is to raise the price of the risky asset to p_T, as shown by the upper dashed line in the figure. Tranching, they argued, enriches the set of securities on offer; but with access tied to asset ownership, as it distorts the asset price.

By the way of contrast, the price of Arrow securities themselves, are assumed to be freely available and enforceable without collateral, which implies the much lower 'complete markets' price pC for the risky asset, which is indicated by the lower dashed line in the figure.[2] Indeed – and this was the nub of their argument – F&G argued that, where tranching is already available, introducing CDS contracts (which can be used to insure against losses when the risky asset 'fails') had the effect of completing the market.[3] The curve CC in Figure D.1 shows market clearing for

[1]Note that, for F&G, 'tranching' refers to the ex-ante creation of Arrow securities, while in normal parlance 'tranching of MBS' (into senior, junior, etc.) refers to the ex-post allocation of losses. The equivalent of a 'senior tranche', which delivers consumption in both states, and would require holding both the Arrow securities.

[2][Razin, 2014, Chapter 7] provides further discussion of this 'complete markets' outcome.

[3]It is as if the full complement of the Arrow securities were available for purchase in desired

the 'down' tranches plus CDS contracts issued by the optimists and purchased by the pessimists. In the resulting Arrow/Debreu equilibrium at EA, the risky assets are held by a larger fraction of the population; and the price falls to EA to reflect the lower valuation of the marginal buyer.

It should be noted that with the introduction of CDS contracts, the optimists hold all the endowments of both risky assets and cash; and they sell insurance contracts to the pessimists. This allows the optimists to consume all the payoffs in the 'up' state while the pessimists consume all the payoffs in the 'down' state, just as in the Arrow-Debreu equilibrium. The asset price crash that their model predicts arises because optimists sell 'naked' insurance to the pessimists (i.e. CDS contracts collateralised by cash), raising the demand for cash, and lowering the price of risky assets.

One could perhaps interpret the 'bad news' shock in the investment banking model of Shin as the impact of introducing CDS contracts, which lowers the price of risky assets. For $R = 0.2$, for example, F&G calculated a fall of about a third as between p_T and p_C. In reality, however, the fall in the value on MBSs, net of insurance against failure, was far greater than one-third.

While F&G's ingenious analysis provides interesting insights, particularly on the role of heterogeneous beliefs, its application comes with several caveats. Most strikingly, there is no bankruptcy in their model. This reflects, it seems, to two strong technical assumptions: first, that all the agents are risk neutral, and second, that all 'naked' CDS contracts are fully collateralised (i.e. with liquid reserves available to cover all the losses). This would be surprising for unregulated OTC contracts; and the spectacular failure of AIG, for example, seems to contradict the assumption.[4] The sequence of events depends on financial innovation, but this is taken to be exogenous.

Unlike the treatment in Allen and Gale [2000], where a distinction is drawn between a new product and process industries (where diversity of opinion about risk is likely) and traditional industries (where it is not), and where a cost must be paid to establish one's type, there is only one risky investment in the F&G economy that is assumed; and all the agents know their type ex-ante. They assumed, moreover, that beliefs are not subject to market manipulation. This is, of course, quite contrary to the approach taken by Foster and Young [2010] on the gaming of performance fees; and of Akerlof and Shiller [2015] on collusion between bankers and rating agencies. We believe legal findings have an important role to play in helping to choose among

quantities with all the contracts fully collateralised.

[4] As the authors acknowledged in their discussion of the CDS market, [Fostel and Geanakoplos, 2012, p. 193].

such conflicting interpretations.

The approach taken in Curatola and Faia [2016] was to postulate heterogeneous tastes for risk-taking rather than the heterogeneous beliefs on outcomes, and to provide a complementary account of how these tastes evolve over time. What emerges as a result is a type of Minsky-cycle based on investor psychology rather than on the forecasting models of investment banks, as in Aymanns et al. [2016].

D.2 No-leverage equilibrium

D.2.1 Private decision in no leverage economy

Suppose, agent h is holding a portfolio containing w_e^h and y_e^h units of risk-free and risky assets, respectively. The agents optimise their portfolio by choosing their holdings of risk-free asset w^h and risky asset y^h and solve the following problem by:

$$\max_{w^h, y^h} \quad h\, U(w^h + y^h) \;+\; (1-h)\, U(w^h + Ry^h) \tag{D.1}$$

subject to:

$$\begin{aligned} w^h + py^h &\leq w_e^h + py_e^h \\ w^h &\geq 0 \\ y^h &\geq 0 \end{aligned}$$

The Lagrangian is:

$$\mathscr{L} = q^h U(w^h + y^h) + (1-q^h)U(w^h + Ry^h) + \lambda(w_e^h + py_e^h - w^h - py^h)$$

Using Khun-Tucker conditions, private decisions suggest that there are three group of agents, which are described as follows:

(1) Extreme optimists will be agents with $h > (\frac{p-R}{1-R})(\frac{1}{p})$ and their private decisions would be:

$$\begin{aligned} y^h &= \frac{w_e^h + py_e^h}{p} \\ w^h &= 0 \end{aligned}$$

(2) Cautious optimists will be agents with $\frac{p-R}{1-R} < h \leq (\frac{p-R}{1-R})(\frac{1}{p})$ and their private decisions would be:

$$\begin{aligned} y^h &= \frac{(h(1-R) - (p-R))}{(1-p)(p-R)}(w_e^h + py_e^h) \\ w^h &= \frac{(p-R) - hp(1-R)}{(1-p)(p-R)}(w_e^h + py_e^h) \end{aligned}$$

Note that the last term of the equation, $w_e^h + py_e^h$, represents equity.

(3) Pessimist will be agents with $h \leq (\frac{p-R}{1-R})$ and their private decisions would be:

$$y^h = 0$$
$$w^h = w_e^h + py_e^h$$

To describe the system of equations that characterises leverage equilibrium, there will be the marginal optimist h_m, who is indifferent between holding and not holding the safe asset, and the marginal pessimist h_s who is indifferent between holding and not holding the risky asset. In equilibrium, all the agents $h > h_b$ will sell all W, and hold only Y. All the agents $h_b \geq h > h_s$ will hold a combination of W and Y. All agents $h < h_s$ will sell all Y, and hold only durable consumption good W.

As a result, the demand function for risky assets given unanticipated news could be described by:

$$y^h(w_e^h, y_e^h) = \begin{cases} \frac{w_e^h + py_e^h}{p} & if \quad h \geq h_b \\ \\ \frac{(h(1-R)-(p-R))}{(1-p)(p-R)}(\omega_e^h + py_e^h) & if \quad h_b > h \geq h_s \\ \\ 0 & if \quad h < h_s \end{cases}$$

The holdings of risky assets could be described by:

$$w^h(w_e^h, y_e^h) = \begin{cases} 0 & if \quad h \geq h_b \\ \\ \frac{((p-R)-hp(1-R))}{(1-p)(p-R)}(\omega_e^h + py_e^h) & if \quad h_b > h \geq h_s \\ \\ w_e^h + py_e^h & if \quad h < h_s \end{cases}$$

where

$$h_s = \frac{p-R}{1-R}$$

and

$$h_b = \left(\frac{p-R}{1-R}\right)\frac{1}{p}$$

Due to strictly monotonicity and continuity for interior solution (cautious optimists) of yy in h and the connected set of heterogeneous agents $h \in [0, 1]$, there will be a unique marginal buyer h_m, who is indifferent between buying and selling

of risky assets; that is, $y = y_e$. The marginal optimist is described by $h_b = (\frac{p-R}{1-R})(\frac{1}{p})$ (he can be interpreted as being a cautious optimist who spends all his wealth for risky assets,$y^h = \frac{w_e+py_e}{1-R}\frac{1}{p}$). The marginal pessimist is described by $h_s = (\frac{p-R}{1-R})$ (he can be interpreted as being a cautious optimist who spends all the wealth for risk-free assets, $y^h = 0$).

D.2.2 Marginal optimist h_b and marginal pessimist h_s given improved asset quality

To prove that h_b is decreasing when there is unanticipated good news in which R shifts up to R', it gives us the following steps:

h_s is increasing with R

Step1

We will show that without allocational effect, h_s is increasing with R. From Eq(4.4), taking derivative of h_s with respect to R gives:

$$\frac{dh_s}{dR} = \frac{(1-R)\frac{dp}{dR} - (1-p)}{(1-R)^2} \tag{D.2}$$

where, $\frac{dp}{dR}$ could be obtained by taking derivative of Eq(4.5)

$$\frac{dp}{dR} = \frac{p(1-p^2)}{(1-R)(R+p(2p-R))} \tag{D.3}$$

It should be noted that no arbitrage condition $R < p < 1$, $\frac{dp}{dR}$ is strictly positive. Plugging Eq(D.3) into Eq(D.2), we shall see that $\frac{dh_s}{dR}$ is strictly positive and the inequalities below are satisfied.

$$\begin{aligned}
\frac{(1-R)\frac{dp}{dR} - (1-p)}{(1-R)^2} &> 0 \\
\frac{dp}{dR} &> \frac{(1-p)}{(1-R)} \\
\frac{p(1-p^2)}{(1-R)(R+p(2p-R))} &> \frac{(1-p)}{(1-R)} \\
(1+p)p &> (R+p(2p-R)) \\
p &< R
\end{aligned}$$

Step2

If we can prove that allocational effect will push up h_s further, then we can be sure that h_s is increasing with R. With good news, wealth is being transferred to optimists who were acquiring Y in the initial equilibrium. As the top group of optimists, who give high valuation to Y, become richer and are going to acquire more Y (h_m, who determine the asset price, increases), this will raise the price of Y. Since we know from Eq(4.4) that $\frac{\partial h_s}{\partial p} = \frac{1}{1-R} > 0$, we know that allocational effect also raises h_s. Thus, we can be ensured that h_s increases with R.

<u>h_b is decreasing with R</u>

Step1

From Eq(4.4), taking derivative of h_b with respect to R gives:

$$\frac{dh_b}{dR} = \frac{(1-R)R\frac{dp}{dR} - (1-p)p}{(1-R)^2p^2} \tag{D.4}$$

Plugging Eq(D.3) into Eq(D.4), we shall see that $\frac{dh_b}{dR}$ is strictly negative and the inequalities below are satisfied, given $0 < R < 1$ and no arbitrage condition $R < p < 1$.

$$\begin{aligned}
\frac{(1-R)R\frac{dp}{dR} - (1-p)p}{(1-R)^2p^2} &< 0 \\
\frac{dp}{dR} &< \frac{(1-p)p}{(1-R)R} \\
\frac{p(1-p^2)}{(1-R)(R+p(2p-R))} &< \frac{(1-p)p}{(1-R)R} \\
(1+p)R &< (R+p(2p-R)) \\
R &< p
\end{aligned}$$

Step2

However, the effect on h_b is uncertain as $\frac{\partial h_b}{\partial p} = \frac{R}{p^2(1-R)} > 0$, which implies that the allocational effect will lead h_b to increase. Thus, if the allocational effect dominates, like in the leverage economy case when $R' = 0.7$, h_b could be increasing and the economy is on the edge of a default. We shall see that later, and in this case the reversal to R could lead to a crash.

D.2.3 Numerical results

1) Initial equilibrium - Given $R = 0.2$, the market clearing and numerical results are described below.

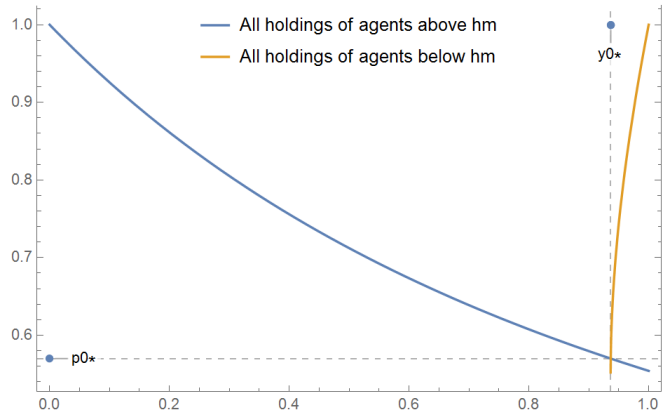


Figure D.2: No-leverage economy - an initial equilibrium's market clearing
Note: (1) the vertical axis represents price, and (2) the horizontal axis represent total amount of risky assets, where from the left axis to blue line is the amount of asset held by agents above h_m and from the right axis to yellow line the amount of asset held by agents above h_m

p0	0.57
h0	0.59

Type		h	y	w
Marginal optimist	hb	0.811	2.754	0.
Marginal buyer	hm	0.589	1.	1.
Marginal pessimist	hs	0.462	0.	1.57

Type	From	To	Size	y	w	Total Asset	Max. Loss
Extreme Optimists	0.8114	1	0.1886	0.5195	0.	0.2961	0.1922
Cautious Optimist	0.4625	0.8114	0.3489	0.4805	0.2739	0.5478	0.1778
Pessimist-Real	0	0.4625	0.	0.7261	0.7261	0.	
Total	0	1	1	1.	1.	1.57	0.37

Figure D.3: No-leverage economy - an initial equilibrium's numerical results

2) Equilibrium with unanticipated good news - Given $R' = 0.7$, the numerical results is described below.

p1	0.8728
h1	0.6238

Type		h	y	w
Marginal optimist	h'b	0.66	2.34	0.
Marginal buyer	h'm	0.624	1.273	0.844
Marginal pessimist	h's	0.576	0.	1.842

Type	From	To	Size	y	w	Total Asset	Max. Loss
Extreme Optimists	0.66	1	0.34	0.905	0.	0.7899	0.1564
Cautious Optimist	0.5761	0.66	0.0839	0.095	0.0801	0.163	0.0164
Pessimist-Real	0	0.5761	0.	0.9199	0.9199	0.	
Total	0	1	1	1.	1.	1.8728	0.1728

Figure D.4: No-leverage economy - a good news equilibrium's numerical results

3) Equilibrium with the reversal of good news - Given R'' is reversed to $= 0.2$, the numerical results is described below.

p2	0.5647
h2	0.6257

Type		h	y	w
Marginal optimist	h''b	0.807	2.754	0.
Marginal buyer	h''m	0.626	1.328	0.802
Marginal pessimist	h''s	0.456	0.	1.57

Type	From	To	Size	y	w	Total Asset	Max. Loss
Extreme Optimists	0.8073	1	0.1927	0.5308	0.	0.2997	0.1936
Cautious Optimist	0.4559	0.8073	0.3514	0.4692	0.2843	0.5492	0.1712
Pessimist-Real	0	0.4559	0.	0.7157	0.7157	0.	
Total	0	1	1	1.	1.	1.5647	0.3647

Figure D.5: No-leverage economy - a good news reversal equilibrium's numerical results

D.3 Leverage equilibrium

D.3.1 Private decision in leverage economy

Now, suppose, agent h is currently holding (endowed with) portfolio containing w_e^h and y_e^h units of risk-free and risky assets, respectively. The current borrowing of the agent is b_e^h. Negative value of b^h implies borrowing while positive value implies lending. The agent optimises his portfolio by choosing holdings of risk-free asset w, risky asset y, and borrowing amount b. The problem thus becomes:

$$\max_{w^h, y^h, b^h} \quad h\, U(w^h - b^h + y^h) \;+\; (1-h)\, U(w^h - b^h + Ry^h) \tag{D.5}$$

subject to:

$$\begin{aligned} w^h - b^h + py^h &\leq w_e^h - b_e^h + py_e^h \\ w^h &\geq 0 \\ y^h &\geq 0 \\ b^h &\leq yR \end{aligned}$$

Note that the right-hand side of the budget constraint represents agent h' wealth or net worth. We can simplify the problem by defining net borrowing, $\omega^h = w^h - b^h$. For simplicity, we assume that an agent will not borrow until a safe asset is used up; that is, an agent will borrow ($b^h < 0$) iff $\omega^h < 0$ (remember, the interest rate is equal to zero). This group of agents who borrow in equilibrium is optimistic about the economy, and would like to shift their consumption from D to U, by borrowing to have additional investment in risky assets. If state D is realized, they will need to return their promises to a lender by using risky asset payoff. The problem can then be written as:

$$\max_{\omega^h, y^h} \quad h\, U(\omega^h + y^h) \;+\; (1-h)\, U(\omega^h + Ry^h) \tag{D.6}$$

subject to:

$$\omega^h + py^h \leq \omega_e^h + py_e^h$$
$$\omega^h \geq -Ry^h$$
$$y^h \geq 0$$

The Lagrangian is:

$$\mathscr{L} = q^h U(\omega^h + y^h) + (1 - q^h)U(\omega^h + Ry^h) + \lambda(\omega_e^h + py_e^h - \omega^h - py^h)$$

Using Khun-Tucker conditions, the solution to this problem divides the agents into three groups, which are described as follows:

(1) Extreme optimists, who are using leverage to gain additional risky assets and have $\omega^h \leq 0$, will be agents with $h \geq \left(\frac{p-R}{1-R}\right)\frac{1}{p}$ and their private decisions would be:

$$y^h = \frac{h(1-R) - (p-R))}{(1-p)(p-R)}(\omega_e^h + py_e^h)$$
$$\omega^h = \frac{(p-R) - hp(1-R)}{(1-p)(p-R)}(\omega_e^h + py_e^h)$$

(2) Cautious optimists, who are not using leverage and holding a combination of risky and risk free asset with $\omega^h > 0$, will be agents with $\frac{p-R}{1-R} \leq h < \left(\frac{p-R}{1-R}\right)\frac{1}{p}$, and their private decisions would be:

$$y^h = \frac{h(1-R) - (p-R))}{(1-p)(p-R)}(\omega_e^h + py_e^h)$$
$$\omega^h = \frac{(p-R) - hp(1-R)}{(1-p)(p-R)}(\omega_e^h + py_e^h)$$

(3) Pessimist, who are holding only risk free asset in their portfolio, will be agents with $h < \left(\frac{p-R}{1-R}\right)$, and their private decisions would be:

$$y^h = 0$$
$$\omega^h = \omega_e^h + py_e^h$$

To describe the system of equations that characterises leverage equilibrium, there will be the marginal pessimist h_s, who is indifferent in holding Y, and the least-optimistic agent who is indifferent between W and Y. In equilibrium, all the

agents $h > h_s$ will hold a combination of W and Y. All the agents $h < h_s$ will sell all their holdings of Y, and hold only durable consumption good W.

In addition, since leverage increases the span of the consumption set for agents (the optimists in particular), in equilibrium there will be the marginal optimist h_b, who is indifferent between the leverage buying and cash buying of Y. All agents $h > h_b$ will sell a promise to finance their buying of Y. All agents $h_b > h > h_s$ will hold a combination W and Y.

As a result, the demand function for risky assets in response to unanticipated news could be described by:

$$y^h(w_e^h, y_e^h) = \begin{cases} \frac{(h(1-R)-(p-R))}{(1-p)(p-R)}(\omega_e^h + py_e^h) & if \quad h \geq h_s \\ \\ 0 & if \quad h < h_s \end{cases}$$

The holdings of risky assets could be described by:

$$\omega^h(w_e^h, y_e^h) = \begin{cases} \frac{((p-R)-hp(1-R))}{(1-p)(p-R)}(\omega_e^h + py_e^h) & if \quad h \geq h_s \\ \omega_e^h + py_e^h & if \quad h < h_s \end{cases}$$

where,

$$h_s = \frac{p-R}{1-R}$$

and,

$$h_b = \left(\frac{p-R}{1-R}\right)\frac{1}{p}$$

D.3.2 Existence of a default and a default threshold

As the borrowing constraint of agents is never binding except for the most optimistic agent with $h = 1$, for $h < 1$ we have:

$$0 < \omega' + R'y'$$

Given the reversal of good news, a default condition for agent h is described by negative equity that is,

$$\omega' + p''y' < 0$$

If $p'' \geq R'$, there will be no default in equilibrium because,

$$0 \quad < \omega' + R'y' \leq \quad \omega' + p''y'$$

If $p'' < R'$, there will always be a default in equilibrium. Since $\omega' + R'y'$ is decreasing in h and $\omega' + R'y' = 0$ for the top optimist $h = 1$, there will be agents in which,

$$\omega' + p''y' \quad < 0 < \omega' + R'y'$$

And for $h = 1$, whose borrowing constraint is binding $\omega' = -R'y'$,

$$\omega' + p''y' = -R'y' + p''y' = (p'' - R')y' < 0$$

As agent h's equity $\omega' + p''y'$ is decreasing in h, there will be a unique default threshold h_d that satisfies, $\omega' + p''y' = 0$. Thus, the default threshold, thus, could be described by:

$$h_d = Min\left(\frac{(1-p'')(p'-R')}{(p'-p'')(1-R')}\ ,\ 1\right)$$

D.3.3 Numerical results

1) Initial allocation - Given $R = 0.2$, the market clearing and numerical results are described below.

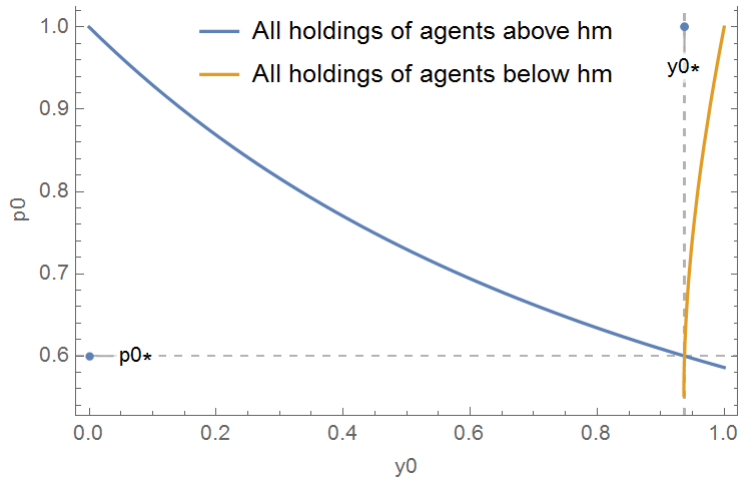


Figure D.6: Leverage economy - an initial equilibrium's market clearing

p0	0.6
h0	0.625

Type		h	y	w
Binded-constraint optimist	1	1.	4.	-0.8
Marginal optimist	h0b	0.833	2.667	0.
Marginal Buyer	hm	0.62	1.	1.
Marginal pessimist	h0s	0.5	0.	1.6

Type	From	To	Size	y	w	Total Equity	Max. Loss	Leverage
Extreme Optimists	0.8333	1	0.1667	0.5556	-0.0667	0.2667	0.2222	1.25
Cautious Optimists	0.5	0.8333	0.3333	0.4444	0.2667	0.5333	0.1778	
Pessimists	0	0.5	0.5	0.	0.8	0.8	0.	
Total	0	1	1	1.	1.	1.6	0.4	

Figure D.7: Leverage economy - an initial equilibrium's numerical results

2) 'Good news' equilibrium - Given $R' = 0.7$, the market clearing and numerical results are described below.

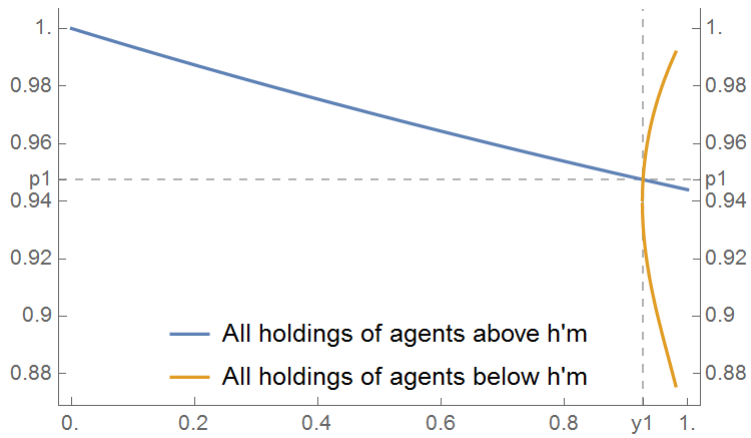


Figure D.8: Leverage economy - a 'good news' equilibrium's market clearing

p1	0.9475
h1	0.8741

Type		h	y	w
Binded-constraint optimist	1	1.	12.081	-8.457
Marginal Buyer	h'm	0.874	2.993	-0.196
Marginal optimist	h1'b	0.871	2.776	0.
Marginal pessimist	h1's	0.825	0.	2.503

Type	From	To	Size	y	w	Total Equity	Max. Loss	Leverage
Extreme Optimists	0.8707	1	0.1293	0.9376	-0.5249	0.3634	0.232	2.44437
Cautious Optimists	0.825	0.8707	0.0457	0.0624	0.0582	0.1174	0.0155	
Pessimists	0	0.825	0.825	0.	1.4667	1.4667	0.	
Total	0	1	1	1.	1.	1.9475	0.2475	

Figure D.9: Leverage economy - a 'good news' equilibrium's numerical results

3) 'good news' reversal equilibrium' - Given R'' is reversed to $= 0.2$, the market clearing and numerical results are described below.

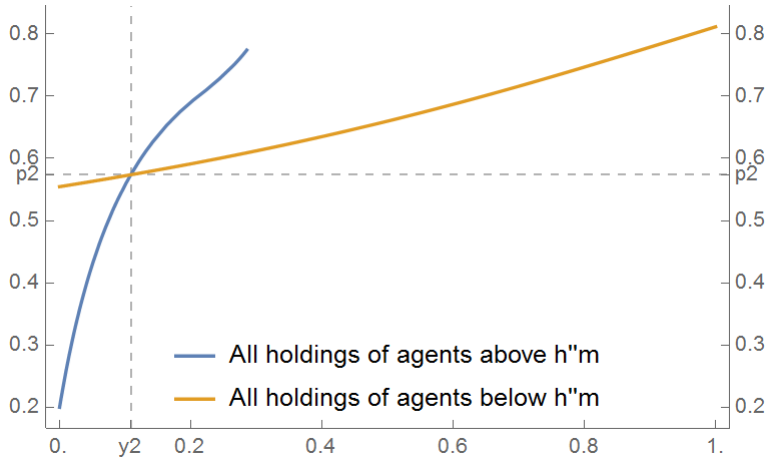


Figure D.10: Leverage economy - a 'good news' reversal equilibrium's market clearing

p2	0.5737
h2	0.8752

Type		h	y	w
Default Threshold	h''d	0.9409	0.	0.
Marginal Buyer	h''m	0.8752	3.0671	-0.2629
Marginal optimist	h''b	0.8143	4.3115	0.
Marginal pessimist	h''s	0.4672	0.	1.6

Type	From	To	Size	y	w	Total Equity	Max. Loss	Leverage
Default Optimist	0.9409	1	0.0591	0.	-0.0443	-0.0443	0.	
Extreme Optimists	0.8143	0.9409	0.1266	0.349	-0.0223	0.1779	0.1304	1.12513
Cautious Optimists	0.4672	0.8143	0.3471	0.651	0.3191	0.6926	0.2433	
Pessimists	0	0.4672	0.4672	0.	0.7474	0.7474	0.	
Total	0	1	1	1.	1.	1.5737	0.3737	

Figure D.11: Leverage economy - a 'good news' reversal equilibrium's numerical results

Bibliography

Acemoglu, D. and Robinson, J. A. (2013). *Why nations fail: The origins of power, prosperity, and poverty.* Broadway Business.

Adrian, T. and Shin, H. S. (2010). Liquidity and leverage. *Journal of Financial Intermediation*, 19(3):418–437.

Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, 84(3):488.

Akerlof, G. A. and Shiller, R. J. (2015). *Phishing for phools: The economics of manipulation and deception.* Princeton University Press, Princeton, NJ.

Allen, F. and Carletti, E. (2012). Systemic risk and macroprudential regulation. In *The Global Macro Economy and Finance*, pages 191–210. Springer.

Allen, F. and Gale, D. (2000). *Comparing financial systems.* MIT press.

Allen, F. and Gale, D. (2004). Financial Fragility, Liquidity, and Asset Prices. *Journal of the European Economic Association*, 2(6):1015–1048.

Allen, F. and Gale, D. (2009). *Understanding financial crises.* Oxford University Press.

Arnold, V. I. (1984). *Catastrophe Theory.* Springer Berlin Heidelberg.

Aymanns, C., Caccioli, F., Farmer, J. D., and Tan, V. W. C. (2016). Taming the Basel leverage cycle. *Journal of Financial Stability*, 27:263–277.

Behn, M., Haselmann, R., and Wachtel, P. (2016). Procyclical capital regulation and lending. *The Journal of Finance*, 71(2):919–956.

Bhattacharya, S., Goodhart, C. A., Tsomocos, D. P., and Vardoulakis, A. P. (2015). A reconsideration of minsky's financial instability hypothesis. *Journal of Money, Credit and Banking*, 47(5):931–973.

Blinder, A. S. (2013). *After the music stopped: The financial crisis, the response, and the work ahead.* Penguin Books, New York, NY.

Borio, C., Furfine, C., Lowe, P., et al. (2001). Procyclicality of the financial system and financial stability: issues and policy options. *BIS papers*, 1(March):1–57.

Brunnermeier, M. K., Eisenbach, T. M., and Sannikov, Y. (2012). Macroeconomics with financial frictions: A survey. Working Paper 18102, National Bureau of Economic Research.

Calomiris, C. W. (2009). The subprime turmoil: What's old, what's new, and what's next. *The Journal of Structured Finance*, 15(1):6–52.

Calomiris, C. W. and Haber, S. H. (2014). *Fragile by design: The political origins of banking crises and scarce credit.* Princeton University Press, Princeton, NJ.

Carlin, W. and Soskice, D. W. (2014). *Macroeconomics: Institutions, instability, and the financial system.* Oxford University Press, Oxford.

Catarineu-Rabell, E., Jackson, P., and Tsomocos, D. P. (2005). Procyclicality and the new basel accord-banks' choice of loan rating system. *Economic Theory*, 26(3):537–557.

Cifuentes, R., Ferrucci, G., and Shin, H. S. (2005). Liquidity risk and contagion. *Journal of the European Economic Association*, 3(2-3):556–566.

Coimbra, N. and Rey, H. (2017). Financial cycles with heterogeneous intermediaries. Working Paper 23245, National Bureau of Economic Research.

Curatola, G. and Faia, E. (2016). Divergent Risk-Attitudes and Endogenous Collateral Constraints. Discussion Papers 11678, Centre for Economic Policy Research.

Danielsson, J., Embrechts, P., Goodhart, C., Keating, C., Muennich, F., Renault, O., Shin, H. S., et al. (2001). An academic response to Basel II. *Special Paper - LSE Financial Markets Group.*

Dávila, E. and Korinek, A. (2017). Pecuniary externalities in economies with financial frictions. *The Review of Economic Studies*, 85(1):352–395.

Diamond, D. W. and Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of political economy*, 91(3):401–419.

Diamond, D. W. and Rajan, R. G. (2011). Fear of fire sales, illiquidity seeking, and credit freezes. *The Quarterly Journal of Economics*, 126(2):557–591.

Foote, C. L., Gerardi, K. S., and Willen, P. S. (2012). Why did so many people make so many ex post bad decisions? the causes of the foreclosure crisis. Working Paper 18082, National Bureau of Economic Research.

Fostel, A. and Geanakoplos, J. (2012). Tranching, CDS, and Asset Prices: How Financial Innovation Can Cause Bubbles and Crashes. *American Economic Journal: Macroeconomics*, 4(1):190–225.

Fostel, A. and Geanakoplos, J. (2015). Leverage and Default in Binomial Economies: A Complete Characterization. *Econometrica*, 83(6):2191–2229.

Foster, D. P. and Young, H. P. (2010). Gaming Performance Fees By Portfolio Managers. *The Quarterly Journal of Economics*, 125(4):1435–1458.

Geanakoplos, J. (2003). Liquidity, default, and crashes endogenous contracts in general. In *Advances in economics and econometrics: theory and applications: eighth World Congress*, volume 2, pages 170–205.

Geanakoplos, J. (2010). The leverage cycle. *NBER macroeconomics annual*, 24(1):1–66.

Gennaioli, N., Shleifer, A., and Vishny, R. (2012). Neglected risks, financial innovation, and financial fragility. *Journal of Financial Economics*, 104(3):452–468.

Goodhart, C. A. (2011). *The Basel Committee on Banking Supervision: a history of the early years 1974–1997*. Cambridge University Press.

Goodhart, C. A., Sunirand, P., and Tsomocos, D. P. (2006). A model to analyse financial fragility. *Economic Theory*, 27(1):107–142.

Goodhart, C. A. and Tsomocos, D. P. (2011). The role of default in macroeconomics. Discussion paper no. 2011-e-23, Institute for Monetary and Economic Studies, Bank of Japan.

Goodhart, C. A., Tsomocos, D. P., and Vardoulakis, A. P. (2010). Modelling a housing and mortgage crisis. Financial markets group discussion papers no. 649, London School of Economics.

Gorton, G. B. (2010). *Slapped by the invisible hand: The panic of 2007*. Oxford University Press, Oxford.

Hellmann, T. F., Murdock, K. C., and Stiglitz, J. E. (2000). Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough? *American Economic Review*, 90(1):147–165.

Holmstrom, B. (2008). Commentary: The panic of 2007. In *Proceedings-Economic Policy Symposium-Jackson Hole*, pages 263–273. Federal Reserve Bank of Kansas City.

Holmstrom, B. and Tirole, J. (1997). Financial Intermediation, Loanable Funds, and The Real Sector. *The Quarterly Journal of Economics*, 112(3):663–691.

Kashyap, A. K., Tsomocos, D. P., and Vardoulakis, A. P. (2014a). How does macroprudential regulation change bank credit supply? Working Paper 20165, National Bureau of Economic Research.

Kashyap, A. K., Tsomocos, D. P., and Vardoulakis, A. P. (2014b). Principles for macroprudential regulation. *Financial Stability Review*, (18):173–182.

Kashyap, A. K., Tsomocos, D. P., and Vardoulakis, A. P. (2017). Optimal bank regulation in the presence of credit and run risk. Finance and economics discussion series 2017-097, Federal Reserve Board, Washington, D.C.

Kay, J. (2017). Justice has settled for second best in the Rolls Royce scandal. https://www.ft.com/content/2c2ffb2e-df08-11e6-86ac-f253db7791c6 [Online; posted 20-January-2017].

Kiyotaki, N. and Moore, J. (1997). Credit cycles. *Journal of political economy*, 105(2):211–248.

Lo, A. W. (2017). *Adaptive Markets: Financial Evolution at the Speed of Thought*. Princeton University Press.

Mian, A. and Sufi, A. (2015). *House of debt: How they (and you) caused the Great Recession, and how we can prevent it from happening again*. University of Chicago Press.

Miles, D. (2015). Housing, Leverage, and Stability in the Wider Economy. *Journal of Money, Credit and Banking*, 47(S1):19–36.

Miller, M. and Zhang, L. (2015). The Hedgehog and the Fox: From DSGE to Macro-Pru. *The Manchester School*, 83:31–55.

Miller, M., Zhang, L., and Rastapana, S. (2016). A comedy of errors: misguided policy, mis-sold mortgages and more. Discussion Paper 11533, Centre for Economic Policy Research.

Milne, A. (2009). *The fall of the house of credit: what went wrong in banking and what can be done to repair the damage?* Cambridge University Press.

Pederzoli, C., Torricelli, C., and Tsomocos, D. P. (2010). Rating systems, procyclicality and basel ii: an evaluation in a general equilibrium framework. *Annals of Finance*, 6(1):33.

Rajan, R. G. (2006). Has finance made the world riskier? *European Financial Management*, 12(4):499–533.

Rajan, R. G. (2011). *Fault Lines.* Princeton University Press, Princeton, NJ.

Rakoff, J. S. (2014). The financial crisis: why have no high-level executives been prosecuted? http://www.nybooks.com/articles/2014/01/09/financial-crisis-why-no-executive-prosecutions/ [Online; posted 9-January-2014].

Razin, A. (2014). *Understanding global crises: an emerging paradigm.* MIT Press.

Reinhart, C. M. and Rogoff, K. S. (2009). *This time is different: Eight centuries of financial folly.* Princeton University Press.

Repullo, R. and Suarez, J. (2013). The procyclical effects of bank capital regulation. *The Review of financial studies*, 26(2):452–490.

Shin, H. S. (2010). *Risk and liquidity.* Oxford University Press.

Shubik, M. (1973). Commodity money, oligopoly, credit and bankruptcy in a general equilibrium model. *Economic Inquiry*, 11(1):24–38.

Shubik, M. and Wilson, C. (1977). The optimal bankruptcy rule in a trading economy using fiat money. *Journal of Economics*, 37(3-4):337–354.

Simsek, A. (2013). Belief disagreements and collateral constraints. *Econometrica*, 81(1):1–53.

Sorkin, A. (2009). *Too big to fail: The inside story of how Wall Street and Washington fought to save the financial system.* Penguin Group, New York, NY.

Stockman, D. (2013). *The great deformation: The corruption of capitalism in America.* Hachette UK.

Tella, S. D. (2017). Optimal regulation of financial intermediaries. Working Paper 23586, National Bureau of Economic Research.

Tsomocos, D. P. (2003a). Equilibrium analysis, banking and financial instability. *Journal of mathematical economics*, 39(5-6):619–655.

Tsomocos, D. P. (2003b). Equilibrium analysis, banking, contagion and financial fragility. *Bank of England. Quarterly Bulletin*, 43(1):53.

Yan, J. (2017). Belief heterogeneity, collateral and endogenous leverage. Job market paper, Said Business School, University of Oxford, UK.

Yellen, J. (2017). Financial stability a decade after the onset of the crisis. In *Presentation at "Fostering a Dynamic Global Recovery," A Symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming (Aug. 25)*.

Zame, W. R. (1993). Efficiency and the role of default when security markets are incomplete. *The American Economic Review*, pages 1142–1164.

Zeeman, E. C. (1974). On the unstable behaviour of stock exchanges. *Journal of Mathematical Economics*, 1(1):39–49.

Zhang, Y. (2017). Procyclical-leverage, asymmetric information and financial market instabilities. Unpublished master's thesis, University of Warwick, Coventry, UK.