



Essays in Financial Economics

Ali Yavuz Polat

Department of Economics

University of Leicester

A thesis submitted for the degree of
Doctor of Philosophy at the University of Leicester.

September 2017

To my Father for the wisdom he provided throughout my life.
To my Mother and Siblings for their continuous and precious support.

Essays in Financial Economics

by

Ali Yavuz Polat

Abstract

This thesis examines the relationship between regulation and financial innovation, salience and financial fragility, and subprime mortgages and lending bubbles. In Chapter 1, we consider a model of two competing banks offering deposit contracts to households. The traditional bank can only invest in risk-free assets and offers risk-free deposits, while the other, risk-taking bank invests in risky assets and offers a risky contract with a higher return. Risk-averse households invest their wealth within the financial sector in order to trade in the subsequent period with a producer. The producer, representing the real economy, bears a capacity adjustment cost if he encounters demand fluctuations. Thus, by investing in the risky technology, the risk-taking bank creates an externality for the real sector through the contract it provides to households. Within this framework we investigate the role of regulation and financial innovation. The main result of this study is that, even in an extreme scenario where, by innovating, the risky bank can fully bypass regulation, regulation is still effective in reducing overall risk. This is due to the fact that regulation generates a “composition effect”, diverting funds away from the “innovative” bank towards the traditional one.

In Chapter 2, we show that salience theory can explain excess volatility of asset prices, and the resulting fire-sales in periods of financial turmoil. Here we classify risk into two types, idiosyncratic and systemic, and postulate that investors over-weigh the type of risk that is salient. Either component of risk (systemic or idiosyncratic) becomes salient when its realisation deviates sufficiently from a reference point. We show that a change in salience – from one component (idiosyncratic) to the other (systemic) – will generate excess volatility. Interestingly, higher risk aversion generally exacerbates excess volatility of prices.

In Chapter 3, we consider a model with two types of households; the poor with no initial endowment and the rich with positive endowment, and two types of assets; properties in a poor area and properties in a rich area. In the model, the poor agents require credit to buy an asset, whereas the rich can draw from their endowment. We show that credit-fuelled housing bubbles sometimes may improve welfare, making poorer individuals better-off. More precisely, there exist two types of equilibria in both markets: One is a bubble equilibrium, and the other is an equilibrium where asset prices are stable over time. While the poor always obtain a positive surplus in the bubble equilibrium, this is not necessarily true for the rich. Our results suggest there may be scope for market interventions aimed at sustaining the value of assets held by credit-constrained agents after the burst of a credit bubble.

Acknowledgements

In the name of Allah, the beneficent, the merciful “My Lord, increase me in knowledge.” Quran 20:114. All praises to Almighty Allah who gave me the courage and strength to complete this project and everything positive in my life is due onto my creator.

I would like to express my special appreciation and thanks to my supervisor Dr. Fabrizio Adriani for his continuous support during my Ph.D. study and for encouraging my research and for allowing me to grow as a researcher. I have been extremely fortunate to have a supervisor who gave me the freedom to explore on my own, and at the same time guided me to recover when my steps stumbled. I am also grateful to Dr. Andre Stenzel for his valuable comments and suggestions. I am truly thankful to my committee members, Professor Spiros Bougheas and Dr Aristotelis Boukouras, for their valuable comments and suggestions.

I must thank the administrative staff and academics of Department of Economics for their help and support. I am grateful to the Department for granting me the graduate teaching assistantship. Also during my study at the USA, the support of the Ministry of National Education, Republic of Turkey is gratefully acknowledged.

A special thanks to my family. Words cannot express how grateful I am to my mother and father for all of the sacrifices that they have made on my behalf. My siblings also deserve a special thanks, I highly value their support and prayers. I would like to express appreciation to my fiancée for her prayers, understanding and support. Last but not the least, I would also like to thank all of my friends who supported me in writing, and motivated me to strive towards my goal.

Declaration

Chapter 1 entitled “Financial Innovation and Regulation” has been presented at the following conference:

- IFABS 2016 Barcelona Conference, Barcelona, Spain
- Chapter 2 entitled “Salience and Financial Fragility” has been presented at the following conference:
 - 24th Annual Conference of the Multinational Finance Society, June 2017, Bucharest, Romania

Contents

Contents	v
List of Figures	vii
Introduction	1
1 Innovation or Loophole? Effect of Regulation under Financial Innovation	5
1.1 Introduction	6
1.2 The Model	10
1.3 Equilibrium Analysis	12
1.3.1 Producer	12
1.3.2 Households	14
1.3.3 Banks	14
1.3.4 Unregulated Equilibria	16
1.3.5 First Best	18
1.4 Innovation as Regulatory Arbitrage	21
1.4.1 Information Structure and First Best Implementation . . .	22
1.4.2 Innovation	22
1.5 Discussion of the Results and Robustness	27
1.5.1 Comparative Statics and Policy Implications	29
1.6 Conclusion	30
2 Salience and Financial Fragility	32
2.1 Introduction	33
2.1.1 Literature Review	35
2.1.2 Salience Theory	39
2.2 The Model	40
2.2.1 Salience	42
2.3 Results	44
2.3.1 Rational Investors	44
2.3.2 Salient Investors	45
2.4 Comparative Statics	50
2.5 Discussion and Policy Implications	55
2.6 Conclusions and Future Work	56

3	Subprime Mortgages and Credit Fuelled Bubbles	59
3.1	Introduction	60
3.1.1	Related Literature	63
3.2	The Model	66
3.3	Characterization of Equilibria	67
3.4	Extensions	85
3.4.1	Depreciation	85
3.4.2	Change in LTV Policy	87
3.5	Policy Implications and Conclusion	88
	Appendix	90

List of Figures

2.1	Price volatility w.r.t risk aversion γ	52
2.2	Price volatility with respect to systemic risk	53
2.3	Price volatility with respect to idiosyncratic risks	54
3.1	Case-Shiller Home Price Index (High vs Low Tier) for San Francisco, California	63
3.2	Case-Shiller Home Price Index (High vs Low Tier) for Miami, Florida	64
3.3	Poor asset market equilibrium for $r_2 \geq r$ where $p_t = B_t$ and $p_{t+1} = B_{t+1}$	72
3.4	Poor asset market equilibrium for $r_2 < r$	73
3.5	Constraints for the rich when $r_2 \geq r$	77
3.6	Constraints for the rich when $r_2 < r$	78
A.7	Outstanding Asset Backed Securities in US	91
A.8	Issuance of Securitization in billion \$	92
A.9	Case-Shiller Home Price Index (High vs Low Tier) for Los Angeles, California	93
A.10	Case-Shiller Home Price Index (High vs Low Tier) for New York, New York	94
A.11	Case-Shiller Home Price Index (High vs Low Tier) for Tampa, Florida	95
A.12	Case-Shiller National, 10-City and 20-City Home Price Indices for the USA	96

Introduction

In this thesis, I explore three different themes in financial economics. The first theme focuses on the interaction between regulation and financial innovation in the context of the discussion on the effectiveness of regulation. The theoretical analysis shows that regulation is still relevant and effective even when banks can bypass the regulation fully. This is due to the fact that risk-averse investors divert their wealth away from the risk-taking bank, towards the safer entity if their contract induces higher return volatility. This study contributes to the literature by providing a foundational base for the discussion on financial regulation. There are two competing banks in the model: The traditional bank, which can only invest in risk-free assets, and offers risk-free deposits; and the risk-taking bank, which offers a risky contract with a higher return, due the (risky) investment opportunities available. Risk averse households (HHs) optimally choose between these two contracts in the first period. In the second period, HHs exchange their realized total wealth with a producer. The producer, representing the real economy, bears a capacity adjustment cost if he encounters demand fluctuations. Thus by investing in the risky technology, the risk-taking bank causes an externality to the real sector through the contract it offers to households. In this way, risky contracts exacerbate demand fluctuations. The risk-taking bank faces a trade-off; by investing more in the risky asset, it can increase profits, but this implies fewer remaining reserves. As a result, more investment in the risky asset will imply a riskier contract for the households, who will then deposit less funds into the risk-taking bank, thus decreasing the bank's profits. The introduction of regulation, aimed at redressing the negative externality, forces the risk-taking bank to innovate. However, since innovation is costly, it also offers worse terms to investors. As a result, either the safe bank is able to attract more funds or the risk-taking bank reduces its exposure to the risky technology. In both cases,

overall risk taking is strictly lower than in the absence of regulation. We also analyse under what conditions regulation is more effective. This happens, when the innovation cost is high, or the total funds attracted by the risk-taking bank are low. These findings are consistent with the policy discussion regarding the downsizing of “too big to fail” financial institutions, primarily since the model suggests that as the funds attracted by the risk-taking bank grow larger, regulation becomes less effective. This effect is driven by the fact that the cost of devising ways around regulation is fixed and has accordingly a smaller impact on larger institutions.

The second theme is concerned with financial instability and asset price volatility. The aim of this section is to understand how investors value assets, especially considering the build-up to and aftermath of the Great Recession. Specifically, we revisit the question of why complex and opaque securities are overpriced ex-ante, and why later, during a crisis period, they are sold at fire-sales prices. The central idea here is that investors perceive different types of risk, idiosyncratic and systemic, disproportionately, focusing more on the type of risk that is salient. Either type of risk (systemic or idiosyncratic) is salient when its realization deviates sufficiently from a reference point. We show that assets are mispriced due to the salient thinking of investors, which creates excess price volatility. We also conduct comparative statics with respect to the model parameters. The most interesting result is that price volatility and mispricing increases with risk aversion. When risk aversion increases due to mean-variance utility, investors’ expected utility becomes more sensitive to a change in perceived variance. Initially, when idiosyncratic risk is salient, agents are willing to pay a premium for assets that carry less idiosyncratic risk. The more risk-averse agents are, the more prevalent the mispricing of these assets. However, when a shock hits the market making systemic risk more salient, these investors will try to rebalance their portfolios in favour of assets that carry less systematic risk. The more risk-averse agents are, the more aggressively they will rebalance their portfolios, thus creating excess volatility. In contrast, in the extreme case of risk neutrality, investors do not care about which type of risk they are facing - systemic or idiosyncratic - and focus solely on the expected returns. Thus, their demand is unaffected by which type of risk is salient. The policy implications of this paper involve both ex-ante and ex-post policies. In terms of ex-post interventions, the proposed model and

results suggest that there is scope for interventions in financial markets such as Troubled Asset Relief Programmes (TARP). However, it should be noted that we do not conduct a welfare analysis, which means that the effectiveness of these interventions should be discussed and analysed in relation to a specific context.

The final theme that is discussed is asset price bubbles, taking into consideration the empirical evidence relating to subprime lending bubbles during the financial crisis. This paper aims to show that credit-fuelled bubbles - where the price of an asset is above the fundamental - can sometimes make poor agents (subprime borrowers) better-off and improve welfare. This argument relies on the idea that asset price bubbles can improve the intergenerational allocation of resources in the presence of financial frictions, such as borrowing constraint. We consider a model with two types of households; the poor (with no initial endowment), and the rich (with some endowment); and two types of assets, properties in a poor area (poor asset), and properties in a rich area (rich asset). In the model, poor agents need credit to purchase an asset, whereas the rich can draw from their endowment. In other words, the poor need a 100% loan-to-value (LTV) mortgage (can be thought of as subprime loans), whereas the rich can make a downpayment. Thus, the equilibrium price for the poor asset is determined by the availability of credit. We show that there exist two types of equilibria for the properties in the poor region. One is a bubble equilibrium, which exists if credit growth is sufficiently large, i.e. a credit-fuelled bubble. The other is an equilibrium where the asset price is stable over time, but the asset is significantly underpriced. In the bubble scenario, prices grow fast enough to enable the poor household who purchased the asset to not only pay back their loan by selling the asset when old, but also to receive a positive surplus. In the no-bubble scenario credit growth is small, implying that price growth is not sufficient to enable the agents to pay back their debt by selling the asset in the future. Thus, in the no-bubble case, the equilibrium price will be zero, i.e. the asset is significantly underpriced. The implication of this result is that the bubble scenario makes the poor better-off and improves welfare. For the rich asset market, there are also two types of equilibria; a bubble equilibrium, and a no-bubble equilibrium with a price at the fundamental value. In the rich asset market, the price is bounded below by the fundamental value, since the rich can always purchase the asset by drawing funds from their endowment. This in contrast, is not true for the poor

asset, which is always underpriced in the absence of bubbles. In terms of policy, our results suggest that there may be scope for market interventions aimed at sustaining the value of the assets held by credit-constrained agents after the burst of a credit bubble.

Chapter 1

Innovation or Loophole? Effect of Regulation under Financial Innovation

Chapter Abstract

We build a tractable model of two competing banks offering deposit contracts to households. One bank, the traditional one, can only invest in risk-free assets and offers risk-free deposits, while the other invests in risky assets and offers a risky contract with a higher return. Risk-averse households solve a canonical portfolio optimization problem and invest their wealth within the financial sector to trade in the subsequent period with a producer. The producer, representing the real economy, bears a capacity adjustment cost if he faces demand fluctuations. Thus, by investing in the risky technology, the risk taking bank causes an externality on the real sector through the contract it offers to households. Risky contracts exacerbate demand fluctuations. Within this framework we investigate the role of regulation and financial innovation. We take a narrow view of financial innovation interpreted as a means to avoid regulation on risk exposure. Our main result is that even in the extreme scenario where, by innovating, the risky bank can fully bypass regulation, regulation is still effective in reducing overall risk. This is due to the fact that regulation generates a “composition effect” diverting funds away from the “innovative” bank towards the traditional one.

1.1 Introduction

The last financial crisis showed that financial innovation can be used to avoid regulation (Kane, 1981, 2012; Wall, 2014). It is often claimed that financial institutions regularly seek innovative ways through which to avoid regulation (The New York Times, 2014)¹, so that regulation is simply a costly activity with negligible gain. Surprisingly, the ineffectiveness of financial regulation is also asserted (in)directly by former regulators. In particular, Greenspan (2011) argues that increased equity capital requirements merely act as an excess buffer despite their significant cost, and appear to have little use. The implication of Greenspan’s claim is that regulation is inefficient and should be kept at a minimal level. Similarly, Howard Davies, the former Chairman of the UK Financial Services Authority and former Deputy Governor of the Bank of England, has criticised supporters of tougher regulation as being in “dangerous territory” and that the arguments given towards building up capital ratios to make banks survive without government aid is “reckless prudence”.² A common theme among critics of regulation is that banks usually find ways to bypass regulatory provision through financial innovations. In this paper, we analyse the interaction between regulation and financial innovation in the context of regulation effectiveness by focusing on the question of whether regulation can still be effective when risk-taking banks bypass such regulation through innovation.

In this paper, we consider financial innovations whose sole purpose is regulatory avoidance.³ We find that, even when banks successfully find ways to avoid mandatory restrictions on risk taking, regulation still has an important effect. This is because through a “composition effect” regulation diverts funds away from

¹The Editorial Board, (2014, July 2). Another Failure to Regulate Derivatives. *The New York Times*, Retrieved from <http://www.nytimes.com/2014/07/03/opinion/another-failure-to-regulate-derivatives>

²As a response to the comment of Greenspan, A. (2011) ‘Regulators must risk more, and intervene less,’ Financial Times, July 26, 2011.

³Acharya et al. (2013) argue that securitized assets were used for regulatory arbitrage rather than as a tool for risk transfer. Stein (2010) also mentions that one of the drivers of securitization was regulatory arbitrage.

risky banks. Innovation allows risk taking banks to offer opaque new assets and to avoid regulation by retaining private information on risk exposure. The central idea of the paper is that even though risky banks can fully avoid regulation, still a lower amount of risky asset is invested compared to the unregulated case due to the composition effect. This result implies that regulation is not neutral.

More precisely, we consider a model of two competing banks: *BankA* can only invest in risk-free assets and offers risk-free deposits, whereas *BankB* offers a risky contract with a higher return, having access to a risky technology.⁴ Risk averse households (HHs) optimally choose between these two contracts in the first period. In the second period, HHs exchange their realized total wealth with a producer. The producer, representing the real economy, bears a capacity adjustment cost if he faces demand fluctuations. Thus, by investing in the risky technology, the risk-taking bank causes an externality to the real sector through the contract it offers to households. In this way, risky contracts exacerbate demand fluctuations. The risk-taking bank (*BankB*) faces a trade-off; by investing more in the risky asset, it can increase profits, but this implies fewer remaining reserves. As a result, more investment in the risky asset will imply a riskier contract for the households, who will then deposit less funds into *BankB* (thus decreasing the bank's profits). The introduction of regulation, aimed at redressing the negative externality, forces *BankB* to innovate. However, since innovation is costly, *BankB* also offers worse terms to investors. As a result, either the safe *BankA* is able to attract more funds or the risk-taking bank reduces its exposure to the risky technology. In both cases, overall risk taking is strictly lower than in the absence of regulation. We also analyse under what conditions regulation is more effective. This happens, when the innovation cost is high, or the total funds attracted by the risk-taking bank are low. These findings are consistent with the policy discussion regarding the downsizing of “too big to fail” financial institutions, primarily since the model suggests that as the funds attracted by

⁴In this sense *BankB* is a monopoly on the risky project. We will comment on the effect of this monopoly later when the regulation case is presented.

the risk-taking bank grow larger, regulation becomes less effective. This effect is driven by the fact that the cost of devising ways around regulation is fixed and has accordingly a smaller impact on larger institutions.

The literature on financial innovation has mostly focused on the effect of innovation as a contributor to financial crises. [Calomiris \(2009\)](#) discusses the motivation for innovation in regards to securitization activities, and claims that policy actions were the main cause of the crisis rather than inaction. [Gennaioli et al. \(2012, 2013\)](#) analyse a model where, due to the high demand for riskless bonds from investors, banks engage in securitization. Under rational expectations, securitization does not cause any fragility since agents understand all possible states (risks) perfectly. However, they show that, in a model where agents neglect the worst state of nature, such neglected risk causes excessive security issuance and thus fragility. Assuming a competitive financial market, [Thakor \(2012\)](#) analyses incentives for innovation, noting that banks have the choice to either make standard loans (where there is an agreement in the market on default probabilities) or to innovate and issue new loans on which there is no agreement as to default. In equilibrium, due to competition standard loans make zero profit whereas new loans make a positive profit since there is no agreement on default probabilities. The problem is that this disagreement about new loans may cause investors to withdraw funding and thereby cause a financial crisis. This trade-off between making a positive profit and the possibility of fragility determines the level of innovation. Different from earlier literature, the model presented in this paper explicitly links innovation to regulatory avoidance and analyses the effect of regulation aimed at reducing risk.

The claims regarding the effectiveness of regulation expounded upon in public debates tend to exaggerate the cost of regulation, and confuse private and social costs.⁵ More importantly, the extreme social and economic costs of low probability catastrophic events (such as crises) are overlooked. Thus, it is crucial to

⁵For a detailed and a comprehensive discussion, see [Admati et al. \(2011\)](#). [Kisin and Manela \(2016\)](#) also find evidence that higher capital requirements will simply imply a modest cost.

address the externalities caused by the financial industry. This paper analyses these externalities by including a real sector which bears a cost due to the risk-taking behaviour of the banks. We contribute to the literature by addressing a very crucial discussion as to the effectiveness of regulation by offering a possible mechanism, namely the composition effect, through which regulation is effective in the presence of externalities.

There are also discussions as to the effectiveness and use of some policies - such as a deposit insurance system. [Allen and Gale \(2003\)](#) argue that a deposit insurance system itself is a bad policy and does not justify another bad policy of capital requirements. This claim should be judged depending on the possible alternative scenarios. If it is claimed that we should impose no regulation and let the market work, then concerns as to the soundness of the financial system should be addressed clearly. [Allen and Gale \(1998, 2000\)](#) argue that in a “laissez-faire system” under “standard” conditions, financial crises may indeed be socially optimal. However, unlike their results derived from the model of [Diamond and Dybvig \(1983\)](#), the bank run is not usually due to equilibrium behaviour, but rather due to banks being highly leveraged so that the accumulation of risks causes a crisis ([Goldstein and Razin, 2013](#)). Due to the nature of the financial sector, according to basic economic intuition, the sector cannot be left as a “laissez-faire system” since the consequence of its actions directly affect the real economy (by causing significant externalities), even assuming away any government guarantee.⁶ [Plantin \(2014\)](#) assuming risk neutral agents, shows that it may be socially optimal to loosen capital requirements for the regulated banking sector in order to dry up the liquidity in the shadow banking sector. This will prevent the regulated banking sector to offload more risk in the shadow sector to bypass capital regulations.

In another branch of the literature, it is claimed that the current regulatory

⁶The nature of the financial system is different in the sense that, in any other sector, the public does not directly care about the soundness of a single firm. However, due to the contagion effects and direct externalities on the real economy, citizens are adversely affected by the failure of a financial institution as this can trigger a financial crisis.

framework is problematic since it is imperfect. Barth et al. (2004) conclude that their objection to the current regulatory framework does not mean regulation is useless, but rather that it needs to be done properly by forcing accurate information disclosure and by avoiding moral hazard incentives of the current deposit insurance framework. This argument is helpful due to it invoking a search for optimal and feasible regulation. However, before discussing such optimal regulation, we should construct a base for possible policy channels in order to agree upon the relevancy of the regulation.

The paper proceeds as follows. In Section 1.2, we present the model. In Section 1.3, we analyse the solution. Here, we present unregulated equilibria where banks are free to choose their investment level and relevant contracts. We then consider the first best solution where a regulator, as a social planner, decides the socially optimal investment levels and contracts. Section 1.4 examines what happens when banks can bypass regulation using costly innovation technology. Section 1.5 discusses the results by comparing innovation equilibria with unregulated equilibria and further presents comparative statics and a number of policy implications. Section 1.6 concludes and offers a possible future extension .

1.2 The Model

There are two time periods $t = 1, 2$. The agents in the economy are; households (HH), a producer and two banks. There is a consumption good valued by all the agents and a second consumption good only valued by HHs.

Banks: There are two types of risk-neutral banks A, B , without any equity. They can only invest by borrowing from HHs. Assume also that banks cannot default strategically, so they have to pay back the promised contractual return to the HHs. Banks do not value the final product of the producer and the profits of banks are redistributed to HHs.

- *Bank A* only deals with traditional banking activity, thereby offering a safe contract (excess return is normalized to 0 for the safe technology.)

- *BankB* can invest both in the safe technology (like *BankA*) and in a sophisticated risky technology. The bank chooses λ which denotes the fraction invested in the risky asset. Thus, $1 - \lambda$ denotes the fraction of capital invested in safe assets (“reserves”). There are two states of nature; a good state that occurs with probability $q \in (0, 1)$ and a bad state that occurs with probability $1 - q$. The risky asset pays r_B in the good state and 0 in the bad state. $E[r_B] = qr_B$ is assumed to be greater than 1 (i.e., the risky technology has an expected return greater than the safe asset, i.e. it has an expected excess return greater than 0). Finally, *BankB* offers a contract (r_H, r_L) to HHs paying r_H in the good state and r_L in the bad state.

Households(HHs): There is a continuum of mass 1 of households, each with time 1 wealth, $W > 0$. Households are risk-averse and have mean-variance utility $EU[W] = E[W] - 1/2Var(W)$. Households consume in the last period after trading with the producer and value both consumption goods equally.

The Producer represents the real economy in the model. Bearing the capacity adjustment cost, as described below, the producer captures the negative externality caused by the financial sector upon the real sector. This point is quite crucial in the model since it brings the reason for regulating the banking sector.⁷ As in Plantin (2014), the producer produces a good in the second period requiring $\alpha < 1$ units of the consumption good as input. Since the producer is a monopoly, he sets a maximal output price of one per unit⁸ so there are potential gains from trade. The initial endowment of HHs, W , must be stored until $t = 2$ so that they can trade with the producer. The banks, as intermediaries in the model, provide such a service. The producer has to determine a capacity level N_1 ex-ante, before observing HHs demand. He can adjust the capacity ex-post to N_2 at date 2, but

⁷In reality, financial downturns affect the real sector directly or indirectly, so one cannot claim that the financial sector should be left alone since there exist significant externalities affecting the real sector.

⁸Since HHs value one unit of consumption good exactly the same as a unit of the production good.

there is an adjustment cost;

$$\frac{k}{2}(N_2 - N_1)^2.$$

We assume that the adjustment cost k is not too large, i.e., $k < \frac{1-\alpha}{Wr_B}$. This ensures that the producer always adjust the capacity in full.

The risky contract (r_H, r_L) ⁹ offered by *BankB* implies a stochastic demand for the producer since HHs have to store their wealth within the financial system. Thus, risk-taking by *BankB* has a real negative externality on the producer through the demand of HHs. The technology works as follows. HHs invest their wealth within the financial sector and then the proceeds from the banks are used to buy the consumption good from the producer. In that sense HHs do not consume their endowments since investing within the financial sector makes them better off.

The timing of the model: At $t = 1$, *BankB* offers (r_H, r_L) . Given (r_H, r_L) , HHs optimize their investment W_B . At $t = 2$, production takes place and HHs trade with the producer.

Considering that λ refers to the risky investment level chosen by *BankB*, $\lambda W_B(r_H, \lambda)$ is denoted as the **total risky investment** in equilibrium.

1.3 Equilibrium Analysis

1.3.1 Producer

Solving backwards from $t = 2$, the producer charges a maximal output price of one per unit since he is a monopolist. This is because HHs value both the consumption good and the production good exactly the same. Ex-post at $t = 2$ the producer finds it optimal to adjust the capacity equal to the realized wealth of HHs. To see this, we revisit Lemma 1 and 2 of [Plantin \(2014\)](#). Suppose that the realized wealth of HHs at $t = 2$ is $W_2 \geq N_1$. Then after observing W_2 , the

⁹Debt is also considered in the set of contracts since we do not assume any specific contract ex-ante, but debt turns out to not be optimal for Bank B .

producer chooses $N_2 \in [N_1, W_2]$ in order to maximize profits;

$$(1 - \alpha)N_2 - \frac{k}{2}(N_2 - N_1)^2.$$

Assuming that k is small enough, i.e., $k < \frac{1-\alpha}{Wr_B}$ ¹⁰ and the producer can charge the maximal price of one, profit is maximized at $N_2 = W_2$. The intuition for this is simple. The producer can extract all surplus from HHs so that he produces the maximum possible output compatible with ex-post wealth W_2 .

Now, suppose that realized wealth of HHs is smaller than the initial capacity $W_2 < N_1$, so that the producer has to choose N_2 . Since $N_2 < W_2$ is never optimal, it should be such that $N_2 \in [W_2, N_1]$. As a monopoly, the producer can at most obtain W_2 from HHs. Then, profit can be written as;

$$W_2 - \alpha N_2 - \frac{k}{2}(N_2 - N_1)^2.$$

Note that, whenever k is small enough, $N_2 = W_2$.

The intuition of this result is as follows. Only HHs value the production good, so that if the producer chooses a production level at $t = 2$ bigger than the realized wealth, the excess production cannot be sold. Again it is not profitable to set a level smaller than the realized wealth of HHs $N_2 < W_2$, since the producer loses the opportunity to sell more and to extract more from HHs.

Ex-ante (i.e. at $t = 1$) knowing that the second period wealth \tilde{W} ¹¹ of HHs is stochastic, the producer sets the initial capacity equal to the expected wealth of

¹⁰In order to have the maximum at $N_2 = W_2$ it needs to be such that $k \leq \frac{1-\alpha}{N_2-N_1}$. Since N_2, N_1 are endogenous restrict attention to $k \leq \min\{\frac{1-\alpha}{N_2-N_1}\}$. The maximum wealth that can be realized for households is $W_A + W_B r_B$ (if the good state is realized $r_H \leq r_B$), where as the minimum is $W_A + W_B 0$ (if the bad state is realized $r_L \geq 0$). Thus, the maximum wealth difference that can occur is $W_B r_B$. Also it is clear that it cannot be optimal for the producer to adjust capacity bigger than the maximum wealth difference so that $N_2 - N_1 \leq W_B r_B < W r_B$. Thus, $k \leq \min\{\frac{1-\alpha}{N_2-N_1}\} < \frac{1-\alpha}{W r_B}$. Note that smaller k also implies that the regulator cares less about the externality (since the cost on the producer will be lower). As will be discussed when regulator result is presented, either r_B or k should be greater than a specific value so that the regulator solution is binding.

¹¹To simplify the notation, the time subscripts will be omitted for the rest of the paper.

HHs, $E[\tilde{W}]$. Thus, the total expected profit of the producer is;

$$(1 - \alpha)(W + W_B(E[r] - 1)) - \frac{k}{2}Var(W_B)$$

where the first part $(1 - \alpha)E[\tilde{W}] = (1 - \alpha)E[W_A + W_B r] = (1 - \alpha)(W + W_B(E[r] - 1))$ is the profit of production and the second part $\frac{k}{2}Var(\tilde{W}) = \frac{k}{2}Var(W_B)$ is the negative externality generated by the economy's risk exposure.

1.3.2 Households

Assuming $u(\cdot)$ satisfies the usual conditions $u' > 0, u'' < 0$; the household solves the following problem;

$$\max_{W_A, W_B \in [0, W]} E[U(W_A + W_B r)] \equiv \max_{W_B \in [0, W]} E[U(W + W_B(r - 1))]$$

where $W_A + W_B = W$ and r is the realization of HH return, which is either r_H or r_L , and $E[r] = qr_H + (1 - q)r_L$. Note that this is a canonical portfolio optimization problem with one safe and one risky asset. FOC is;

$$E[U'(W + W_B(r - 1))(r - 1)] = 0.$$

Assuming mean-var utility $EU[\tilde{W}] = E[\tilde{W}] - 1/2Var(\tilde{W})$, the FOC then implies;

$$W_B(r_H, r_L) = \frac{E[r] - 1}{q(1 - q)(r_H - r_L)^2} = \frac{E[r] - 1}{Var(r)}.$$

1.3.3 Banks

BankA just provides a safe storage technology to HHs so that they can store their wealth until $t = 2$. Consequently, we can just focus on *BankB*'s problem.

Given the optimal solution of the HH as $W_B(r_H, r_L) = W_B$, *BankB* solves the following by choosing the fraction of wealth invested in the risky project λ and

the contract (r_H, r_L) as the payment in good and bad states.

$$\max_{\lambda \in [0,1], r_H, r_L \in R^+} \Pi_B \equiv (1 - \lambda)W_B + \lambda W_B E[r_B] - W_B E[r]$$

s.t.

$$r_L W_B \leq (1 - \lambda)W_B \tag{1.1}$$

$$r_H W_B \leq (1 - \lambda)W_B + \lambda W_B r_B \tag{1.2}$$

where λ is the fraction of HHs wealth invested in the risky project. The left-hand side of (1.1) is the total amount that should be paid to HHs if the bad state is realized, as specified in the contract. The right-hand side is the funds available to the bank if the bad state is realized. (1.2) is similar: The LHS is the amount offered in the contract for the good state and RHS is what the bank has in the good state. Thus, constraints (1.1) and (1.2) are liability constraints stating that in both states the bank is solvent.

Lemma 1.1 *Constraint (1.1) is binding $r_L = (1 - \lambda)$*

Proof. Suppose (1.1) holds with strict inequality, then for fixed r_L and r_H , $W_B(r_H, r_L)$ is fixed. This implies that the profit function Π_B is monotonically increasing in λ (since it is assumed that $E[r_B] = qr_B > 1$). However, if there is a slack in (1.1), profit can be increased if λ is increased by $\epsilon > 0$, still satisfying (1.1). (1.1) cannot then hold with strict inequality.¹² Note that RHS of (1.2) is increasing in λ so that this constraint is not violated when λ is increased by $\epsilon > 0$. ■

Lemma 1.2 *Constraint (1.2) holds with strict inequality, $r_H < (1 - \lambda) + \lambda r_B$*

Proof. Following from Lemma 1.1, since $r_L = 1 - \lambda$, the bank earns zero profit in the bad state. If (1.2) were to be satisfied with equality, then the bank would make zero profit also in the good state and so total expected profit would be zero.

¹²This observation will be very helpful for the characterization of equilibrium and the interpretation of results.

However note that $\{\lambda = 0, r_H = 1, r_L = 1\}$ cannot be an equilibrium since there exist a profitable deviation $\{\lambda = \epsilon, r_H \neq 1, r_L = 1 - \epsilon\}$, $\epsilon > 0$ for the bank which gives positive profits. Thus $\lambda > 0$. Both (1.2) and (1.1) holding with equality implies $E[r] = qr_H + (1 - q)r_L = 1 - \lambda + \lambda qr_B > 1$ since $qr_B > 1$. HHs will then choose $W_B > 0$ since $E[r] > 1$. However, the bank can make positive profit by offering $r^H - \epsilon$ for some $\epsilon > 0$ since HHs will still choose a positive investment, $W_B > 0$ so long as $E[r] - q\epsilon > 1$. This argument shows that (1.2) holds with strict inequality in equilibrium. ■

1.3.4 Unregulated Equilibria

Substituting $r_L = (1 - \lambda)$ in Π_B , we can rewrite *BankB's* problem as follows;

$$\max_{\lambda \in [0,1], r_H \in [1, r_B]} (1 - \lambda)W_B + \lambda W_B E[r_B] - W_B E[r] \equiv q W_B [1 + \lambda(r_B - 1) - r_H]$$

$$r^H < 1 + \lambda(r_B - 1).$$

Proposition 1.3 *There exists a continuum of unregulated equilibria, which consist of optimal contracts (r_H^*, r_L^*) and risky investment levels λ^* , such that*

$$(1 - r_L^*) = \lambda^* = (r_H^* - 1) \frac{(qr_B + 1)}{r_B(2 - q) - 1} \quad (1.3)$$

for $\lambda^* = 1 - r_L^* \in (\frac{\phi}{W}, 1]$ and $r_H^* \in (1, r_B]$.

In any equilibrium,

(i) *The amount invested in the risky asset is the same and given by*

$$\lambda^* W_B(r_H^*, \lambda^*) = \frac{q^2 r_B^2 - 1}{4(1 - q)qr_B^2} \equiv \phi. \quad (1.4)$$

(ii) The decentralized profits of Bank B, Π_B^{Dec} are the same and given by;

$$\Pi_B^{Dec} = \frac{(qr_B - 1)^2}{4(1 - q)r_B}.$$

Proof. Solving for the FOCs for λ and r_H , we obtain;

$$\lambda^* = \frac{r_H - 1}{r_B - 1} - \frac{W_B(\lambda, r_H)}{\frac{\partial W_B}{\partial \lambda}} \quad \text{and} \quad r_H^* = 1 + \lambda(r_B - 1) - \frac{W_B(\lambda, r_H)}{\frac{\partial W_B}{\partial r}}.$$

From these FOCs, we obtain $\frac{\partial W_B}{\partial r_H}(r_B - 1) = -\frac{\partial W_B}{\partial \lambda}$

Then, using the HH solution, $W_B(r_H, r_L)$, imposing mean-variance utility as shown above and substituting $r_L = (1 - \lambda)$, we obtain;

$$W_B(\lambda, r_H) = \frac{q(r_H - 1) - (1 - q)\lambda}{q(1 - q)(r_H + \lambda - 1)^2}.$$

Using this and substituting its partial derivatives (as above) results in equation (1.3).¹³

Using (1.3), substituting $\lambda^* = (r_H^* - 1)\frac{(qr_B + 1)}{r_B(2 - q) - 1}$ in $W_B(\lambda^*, r_H^*)$ as given above we get

$$W_B(\lambda^*, r_H^*) = \frac{q(r_H^* - 1) - (1 - q)(r_H^* - 1)\frac{(qr_B + 1)}{r_B(2 - q) - 1}}{q(1 - q)\left(r_H^* + (r_H^* - 1)\frac{(qr_B + 1)}{r_B(2 - q) - 1} - 1\right)^2}$$

Total risky investment is calculated by $\lambda^*W_B(\lambda^*, r_H^*)$ as given above. Also substituting λ^* in the profit function given above, the profit of Bank B is given by;

$$\frac{q(r_H^* - 1) - (1 - q)(r_H^* - 1)\frac{(qr_B + 1)}{r_B(2 - q) - 1}}{q(1 - q)\left(r_H^* + (r_H^* - 1)\frac{(qr_B + 1)}{r_B(2 - q) - 1} - 1\right)^2} \left[1 + (r_H^* - 1)\frac{(qr_B + 1)}{r_B(2 - q) - 1}(r_B - 1) - r_H^*\right] = \frac{(qr_B - 1)^2}{4(1 - q)r_B}$$

Note that we need to check two constraints namely $W_B \leq W$ and $r^H < 1 + \lambda(r_B - 1)$. As it is proved in the proof of Lemma 1.2 $\lambda = 0$ cannot be an equilibrium which means $r^H < 1 + \lambda(r_B - 1)$ holds since $r_B > 1$. Also note that HHs cannot invest in Bank B more than their wealth $W_B \leq W$. Thus using (1.4) we restrict $\lambda > \frac{\phi}{W}$. ■

¹³In using (1.3) it is assured that the above liability constraint $r^H < 1 + \lambda(r_B - 1)$ is satisfied.

The intuition behind the proposition is as follows: In equilibrium, *BankB* is free to choose different optimal contracts¹⁴ (r_H, r_L) such that if a higher return r_H is offered in the good state, λ should also be higher. Possessing a higher λ also implies the return in the bad state r_L should be lower (from (1.1)). But this means more risk is transferred to HHs since $r_H - r_L$ will be higher. Thus, HHs will invest less W_B in *BankB*. Finally, the increased expected profit, as is due to higher investment in the risky project (higher λ) is exactly offset by the decreased demand W_B . The reason for this is that, from the viewpoint of the HHs, direct investment in *BankA* or *BankB* accumulating reserves on their behalf are perfect substitutes. Thus, if *B* offers a contract with a higher variance, HHs will undo the higher risk transferred to them in their contract by investing more in the safe bank W_A and less in the risky bank W_B . This implies that the total amount of risky investment, in any equilibrium among the continuum, is the same (as shown in (1.4)). Another way to understand the proof of Proposition 1.3 is by observing that the expression $r_H W_B$ is constant. In other words, HHs do not care the level of λ since the bank offers a constant return.

1.3.5 First Best

We next characterize a regulator who can achieve the first best since we assume that there is no information asymmetry or enforcement problem. The regulator chooses the socially optimal pair of (λ^R, r_H^R) considering the externality on the producer. Since only HHs value the production good, the total demand for the producer is whatever wealth consumers (HHs) possess at the end of period 2. Thus, as long as HHs choose $W_B > 0$, the producer has to bear an adjustment cost due to the stochasticity of the HHs' wealth.¹⁵

In this section, we simply assume that the regulator can observe the risky investment level λ and r_H so that he can perfectly enforce the regulation.

¹⁴Since this is a continuum of equilibria, bank can choose any contract from the continuum set of contracts.

¹⁵Since $qr_B > 1$, from the well-known result of canonical portfolio optimization with a risky and a risk-free asset, the risk-averse HH will always choose a level of $W_B > 0$.

The regulator solves the following problem;

$$\max_{\lambda \in [0,1], r_H \in R^+} E[\Pi_B + U_{HH}] + (1 - \alpha)(W + W_B(E[r] - 1)) - \frac{k}{2} \text{Var}(W_B) \quad (1.5)$$

where $(1 - \alpha)(W + W_B(E[r] - 1))$ is the producer's profit and the last term is the externality.¹⁶ Recall that α is the cost of producing each unit so $(1 - \alpha)$ is the per unit profit of the producer.

Proposition 1.4 *There exists a continuum of first best (regulator's) equilibria which consist of contracts (r_H^R, r_L^R) and risky investment levels λ^R , such that;*

$$\lambda^R = (r_H^R - 1) \frac{2 - q(r_B - k) - \alpha(1 - q)}{qr_B + (k + \alpha)(1 - q) - 1} \quad (1.6)$$

for $\lambda^R = 1 - r_L^R \in (\frac{\phi}{W}, 1]$ and $r_H^R \in (1, r_B]$.

In any equilibrium:

(i) *The amount invested in the risky asset is the same and given by*

$$\lambda^R W_B(r_H^R, \lambda^R) = \frac{(qr_B + \alpha(1 - q) - kq - 2)(q(1 + r_B) + \alpha(1 - q) - 2)}{(1 + k)^2(-1 + q)q} \equiv \kappa.$$

The total risky investment under regulation is less than that of unregulated case, namely $\kappa < \phi$ if

$$(a) \quad r_B > \frac{q \left(\sqrt{\frac{q^2(2\alpha+k-1)^2 - 4q(\alpha(2\alpha+k-5) - 4k) + 4(\alpha-2)^2}{q^2}} + k - 1 \right) + 2\alpha(q-1) + 4}{4q}, \quad k > 0 \quad \text{or}$$

$$(b) \quad k > \frac{1 + r_B(2\alpha(1 - q) - 4) + q(2r_B + 1))}{qr_B - 1}.$$

(ii) *The profit of Bank B is the same and given by;*

$$\Pi_B^{Reg} = \frac{(\alpha(1 - q) + q(r_B + 1) + 2)(r_B(\alpha(1 - q) + qr_B - 2) + k(qr_B - 1) - 1)}{(k + 1)^2(1 - q)}.$$

Proof. Similar to Proposition 1.3, the equilibria is obtained by simultaneously

¹⁶ Π_A can be ignored since his choice has no effect.

solving the FOCs of the above problem. Using

$$E[\Pi_B] \equiv (1 - \lambda)W_B + \lambda W_B E[r_B] - W_B E[r] = W_B(1 - \lambda + \lambda E[r_B] - E[r])$$

and

$$E[U_{HH}] = E[W] - 1/2 \text{Var}(W) = W + W_B(E[r] - 1) - \frac{1}{2} \text{Var}(W_B),$$

(1.5) becomes

$$W_B(1 - \lambda + \lambda E[r_B] - E[r]) + (2 - \alpha)(W + W_B(E[r] - 1)) - \frac{k + 1}{2} \text{Var}(W_B).$$

Using this maximization argument and optimal portfolio solution of HH as

$$W_B(\lambda, r_H) = \frac{q(r_H - 1) - (1 - q)\lambda}{q(1 - q)(r_H + \lambda - 1)^2},$$

the FOCs w.r.t λ and r_H gives us (1.6) Then, total risky investment κ , and total profit Π_B^{Reg} is calculated using (1.6). ■

If the return of the risky project r_B or the externality parameter k is big enough as stated above, the first best will result in lower (λ^R, r_H^R) compared to the unregulated (λ^*, r_H^*) . For the parametric restriction given in part (i) (r_B or k being large enough), Total Risky Investment (TRI) will also be lower. The intuition of the need for this parametric restriction is as follows. In the model, *BankB* has monopoly power on the risky project. Abstracting from the externality, there will be a monopoly distortion; i.e. if the externality is zero ($k = 0$), *BankB* will choose too low (λ^*, r_H^*) and will invest a suboptimal fraction of funds in the risky project. Thus, we need either a larger externality or a larger variance of the risky project (since in the bad state, the project returns zero and having a larger r_B means a bigger variance) or both to have a binding regulation. Also note that for the parametric relation in (i)(b) where $k > \underline{k} = \frac{1 + r_B(2\alpha(1 - q) - 4) + q(2r_B + 1)}{qr_B - 1}$, the cut-off \underline{k} increases with q or

r_B .¹⁷ This is intuitive considering that qr_B is the expected return of the risky project. When the expected return is higher, the regulator cares less about the externality since the socially efficient level of risky investment level should be higher. Thus, considering the trade-off that the regulator faces, in order to outweigh the monopoly distortion mentioned above, the externality should be stronger for higher levels of expected return $E[r_B] = qr_B$.

1.4 Innovation as Regulatory Arbitrage

Notably, with the help of improvements in IT technology over the last 20 years, financial markets have experienced fast innovation. Some innovations have improved access to information and have decreased transaction times. This study instead treats innovation as a tool through which to avoid regulation. This assumption is motivated by what we observed prior to the last financial crisis. For example, financial institutions developed products comprised of different already complex underlying assets, essentially adding another layer of complexity. This over-complex structure helped financial institutions to hide the real risks they were taking. Offloading the risk to off-balance sheet entities was another widely used strategy designed to avoid capital regulations and controls.¹⁸ Acharya et al. (2013) argue that securitized assets were used for regulatory arbitrage rather than the commonly cited motivation of securitization acting as a tool for risk transfer. Stein (2010) also explains how securitization was used to facilitate regulatory arbitrage:

“It has become apparent in recent years that another important driver

¹⁷Note that $k > \underline{k}$ is compatible with our earlier assumption on k .

¹⁸In a special report by Moody’s Investors Service, offloading risk is explicitly described. Here, it is asserted that: “ABCP programs also offer advantages to their bank sponsors. The programs are typically structured and accounted for by the banks as an off-balance sheet activity. If the bank were to provide a direct corporate loan, even one secured by the same assets, it would appear on the bank’s balance sheet as an asset and the bank would be obligated to maintain regulatory capital for it. An ABCP program permits the Sponsor (i.e., the commercial bank) to offer receivable financing services to its customers without using the Sponsor’s balance sheet or holding incremental regulatory capital.”

of securitization activity is regulatory arbitrage, a purposeful attempt by banks to avoid the rules that dictate how much capital they are required to hold. The most obvious alternative explanation is that banks exploited a regulatory loophole: If they held the loans directly on their balance sheets, they faced a regulatory capital requirement on these loans; but if they securitized the loans and parked them in an off-balance-sheet vehicle (albeit one with essentially full recourse to the banks in the event of trouble), the regulatory capital requirement was much reduced.”

In the model, we assume *BankB* can pay a fixed cost to avoid regulation by making λ private information. Crucially, we show that even though the cost of innovation is fixed and enters the profit function as a sunk cost, it still has a real effect arising from the demand side. If innovation cost were assumed to have a variable part, the results of this paper would be strengthened, thus implying a further decrease in the optimal share of risky investment λ^* chosen by the bank.

1.4.1 Information Structure and First Best Implementation

If the model is extended such that the regulator observes λ and r_H only with some probability, the first best can still be implemented by threatening a large enough¹⁹ fine whenever *BankB* is caught having a λ or r_H different from the regulated level. Thus, for the sake of simplicity, we will assume that in the absence of financial innovations, the first best can be implemented directly.

1.4.2 Innovation

By innovating and offering a new opaque product the bank can keep the risky investment level λ as private information. Thus, the regulator cannot impose any specific risky investment level. This is seemingly similar to what happened prior

¹⁹Namely extracting all the profits as a fine.

to the financial crisis. For instance, by issuing highly opaque derivative securities (e.g., ABSs, CDOs, CDO²), the banks gained an information advantage over the regulators.

The innovation works as follows: *BankB* can transfer the risky project to a Structured Investment Vehicle (SIV) so that securities are not issued within the banking entity but rather through the SIV. Thus, the regulator cannot observe the risky investment (λ) through *BankB*'s balance sheet, with this implying that no regulation can be enforced (i.e., λ^R and r_H^R cannot be enforced). However, this innovative way of offering new securities comes with a cost.

From the bank's point of view, this means that innovation has a trade-off.²⁰ It will help the bank to avoid the regulation but with a cost of c . Thus, depending on the level of the cost c , the bank will either go for innovation and avoid the regulation or just comply with the regulation without innovation.²¹

To simplify the analysis and to make the point more clear, we assume that by paying the fixed cost c , *BankB* can fully avoid regulation (i.e., even if the bank chooses a different λ level than the enforced λ^R , there is still no risk of being punished since the probability of being caught becomes zero).²² Moreover assume that HHs know the fraction of assets invested in the risky assets and make informed decisions.

Then, the new problem for the *BankB* is as follows;

$$\max_{\lambda \in [0,1], r^H \in R^+, I \in \{0,1\}} \Pi(\lambda) = (1-p(I))((1-\lambda)W_B + \lambda W_B E[r_B] - W_B E[r]) + (p(I)\Pi_B^R) - cI$$

where $I \in \{0,1\}$ is the choice variable for innovation, $p(I)$ is the probability

²⁰Kisin and Manela (2016): "While the loophole benefited banks by relaxing their regulatory constraints, using it was costly, as banks had to pay an incremental cost for using ABCP conduits. Therefore, for constrained banks that use the loophole, the ratio of the marginal cost of using the loophole to its marginal capital relief reveals the shadow cost of the regulatory capital constraint."

²¹Namely, when the cost is bigger than the extra return captured by avoiding regulation. See Lemma 1.5.

²²This simplification helps us to make the argument (as discussed in Section 1.5) more clear under a simple framework. The intuition is as follows: Although we assume the bank can fully bypass regulation, such regulation still has an effect. Therefore, if we relax this assumption, our results will hold even stronger.

that the regulator catches (and punishes) the bank for the given I , and Π_B^R is the profit under regulation. For simplicity, we assume that innovation ensures perfect regulatory avoidance. *BankB* first decides whether to innovate or not, at the beginning of $t = 1$. Regulation is fully avoided if the bank chooses to innovate, thereby implying that $p(I = 1) = 0$. If it decides not to innovate $p(I = 0) = 1$, then whenever the bank chooses a different $\lambda \neq \lambda^R$, all profits will be fined so that the bank has to comply with the regulation (i.e., $\Pi(\lambda) = \Pi_B^R$). Thus, if the bank innovates $I = 1$, then the probability part disappears and the cost is paid. If $I = 0$, the bank has to consider the punishment.

Lemma 1.5 *There exists a threshold level \bar{c} for the rating cost, such that BankB will innovate and avoid regulation if $c < \bar{c}$ and will not innovate and comply otherwise. Formally;*

- (i) *if $c < \bar{c}$ the bank will choose to innovate and avoid regulation since $\Pi_B^{innovation} > \Pi_B^{Reg}$.*
- (ii) *if $c \geq \bar{c}$ the bank will not innovate and just comply with the regulation since $\Pi_B^{innovation} \leq \Pi_B^{Reg}$.*

Proof. Considering the maximization problem given above and in applying the same logic used in the proof of Proposition 1.6, innovation profit can be found from the comparison with the deregulated profit. For any given $r_H^i = r_H^*$ using (1.3) and (1.9) $r_L^i = r_L^*$ also holds, thus implying that $W_B(r_H^*, r_L^*) = W_B(r_H^i, r_L^i)$ since the optimal contracts are the same. Moreover $r_H^i = r_H^*$ implies $\left(\lambda^i + \frac{c}{W_B(r_H, \lambda)}\right) = \lambda^*$. Thus, we can rewrite the equilibrium profit,

$$\Pi_B^{innovation} = (1 - \lambda^i)W_B(r_H^i, r_L^i) + \lambda^i W_B(r_H^i, r_L^i)E[r_B] - W_B(r_H^i, r_L^i)E[r] - c$$

where $E[r] = qr_H^i + (1 - q)r_L^i = qr_H^* + (1 - q)r_L^*$. Substituting $\lambda^i = \lambda^* - \frac{c}{W_B(r_H, \lambda)}$, we can write innovation equilibrium profit in terms of deregulated equilibrium

profit. Formally,

$$\begin{aligned}
\Pi_B^{innovation} &= W_B(r_H^*, r_L^*)(1 + \lambda^i(E[r_B] - 1) - E[r]) - c = \\
&W_B(r_H^*, r_L^*)(1 + (\lambda^* - \frac{c}{W_B(r_H, \lambda)})(E[r_B] - 1) - E[r]) - c = \\
&W_B(r_H^*, r_L^*)(1 + \lambda^*(E[r_B] - 1)) - W_B(r_H^*, r_L^*)\frac{c}{W_B(r_H^*, r_L^*)}(E[r_B] - 1) - c = \\
&\Pi_B^{Dec} - cE[r_B] = \frac{(qr_B - 1)^2}{4(1 - q)r_B} - cE[r_B].
\end{aligned}$$

Under innovation, the bank's equilibrium profit is reduced by $-cE[r_B] = -cqr_B$ compared to the unregulated case. Thus,

$$\Pi_B^{innovation} = \Pi_B^{Dec} - cqr_B = \frac{(qr_B - 1)^2}{4(1 - q)r_B} - cqr_B.$$

Using this final expression and comparing it with Π_B^{Reg} , we found the threshold

$$\bar{c} = \frac{(1 + k - kqr_B + r_B(-4 + q + 2qr_B + 2\alpha - 2q\alpha))^2}{4(1 + k)^2(1 - q)qr_B^2}.$$

■

If the innovation cost c is large enough, the bank will not prefer to pay the cost and no innovation will occur. We are then back to the full regulation enforcement scenario. To focus on the interesting scenario, we will analyse the case where innovation is profitable. Hence, under innovation ($I = 1$) *BankB*'s problem can be rewritten as;

$$\max_{\lambda \in [0, 1], r^H \in R^+} \Pi(\lambda) = (1 - \lambda)W_B + \lambda W_B E[r_B] - W_B E[r] - c$$

s.t.

$$r^L = (1 - \lambda) - \frac{c}{W_B} \tag{1.7}$$

$$r^H < 1 + \lambda(r_B - 1) - \frac{c}{W_B} \tag{1.8}$$

where (1.7) and (1.8) are the liability constraints similar to the unregulated case.

The difference is that the innovation cost tightens both constraints. Another important point is that as W_B becomes smaller (i.e., HHs invest less in *BankB*) the constraints tighten further. Thus, by observing the wealth invested in the financial system, the regulator may have an idea, in advance, as to the effect of their policy making.²³

Proposition 1.6 *Innovation equilibria consist of a set of contracts (r_H^i, r_L^i) and risky investment level λ^i such that;*

$$(1 - r_L^i) = \left(\lambda^i + \frac{c}{W_B(r_H, \lambda)} \right) = (r_H^i - 1) \frac{(qr_B + 1)}{r_B(2 - q) - 1} \quad (1.9)$$

for $\lambda^i \in (\frac{\phi}{W}, 1]$ and $r_H^i \in (1, r_B]$.

Innovation equilibria result in a lower level of risky investment compared to that of the unregulated case: In any equilibrium, for any given $r_H^i = r_H^$, $\lambda^i < \lambda^*$ always holds.*

Proof. The FOCs stated in the proof of Proposition 1 is a general solution in terms of W_B . So the only difference in the innovation case is in terms of optimal $W_B(\lambda, r_H)$ since the liability constraints have changed. Substituting (1.7) in W_B (HH's optimal investment) found in Section 1.3.2, we obtain

$$W_B(\lambda, r_H) = \frac{qr_H - (1 - q)(\lambda + c/W_B(\lambda, r_H))}{q(1 - q)(r_H + \lambda + c/W_B(\lambda, r_H) - 1)^2}.$$

Then using this result and substituting in FOCs we get equation (1.9).²⁴

The last part of the proposition is from just comparing equations (1.3) and (1.9). For any given $r_H^i = r_H^*$, $\left(\lambda^i + \frac{c}{W_B(r_H, \lambda)} \right) = \lambda^*$ implying that $\lambda^i < \lambda^*$ since $\frac{c}{W_B(r_H, \lambda)} > 0$ always holds. ■

These results²⁵ imply that even under the extreme scenario of full avoidance

²³This observation may have some interesting policy implications.

²⁴Note, the solution of W_B is intentionally used as such without explicitly solving it for the parameters. The reason is to make the result similar to that of the unregulated case to intuitively compare them. Also, the explicit solution is quite messy and unintuitive.

²⁵Note that equation (1.9) is intentionally left as above instead of writing the explicit solution in terms of λ . The reason is to directly compare with equation (1.3)

of regulation, still regulation is effective and not irrelevant. This indicates that if the full avoidance assumption is relaxed, the results hold even stronger.

Note that, the bank offers the same contract (r_H, r_L) under innovation compared to the unregulated contract,²⁶ whereas the risky investment level has to be lower for a given contract. The intuition is that innovation requires resources. If the bank wishes to offer the same payment in the good state of nature, while keeping the same risky investment as in the no regulation case, it needs to reduce its payment in the bad state. This however implies that HHs face higher risk and lower expected return. In order to boost the payment in the bad state, *BankB* thus needs to reduce its risky investment.

1.5 Discussion of the Results and Robustness

Since we assume that the bank can bypass regulation fully, we will directly compare the innovation result with the unregulated result.

Definition 1.1 *The composition effect refers to the absolute difference between total risky investment in the unregulated equilibrium and in the innovation equilibrium.*

Proposition 1.7 *Compared with the unregulated case, total risky investment level is reduced due to the composition effect. Formally; $\lambda^i W_B(r_H^i, \lambda^i) < \lambda^* W_B(r_H^*, \lambda^*)$ holds for all λ^i, r_H^i satisfying (1.9) and λ^*, r_H^* satisfying (1.3). Given the lower externality on the producer, regulation is effective even under the full avoidance assumption.*

Proof.

In focusing on equations (1.3) and (1.9) the only extra term in the innovation case, (1.9), is the extra term $\frac{c}{W_B(r_H, \lambda)} = \frac{c}{W_B(r_H, r_L)}$, which is always positive

²⁶See equations (1.3) and (1.9)

since both $c > 0$ and $W_B(r_H, \lambda) > 0$.²⁷ We will proceed as follows to show that regulation is effective: Pick a pair of (r_H, λ) such that $r_H^* = r_H^i$. I.e. pick an equilibrium pair where the payment in the high state is the same in both cases. This also implies $r_L^* = r_L^i$ from the equations (1.3) and (1.9). From the HH's point of view, in having the same contract in both the unregulated and the innovation case, $(r_H^*, r_L^*) = (r_H^i, r_L^i)$, with this implying $W_B(r_H^*, r_L^*) = W_B(r_H^i, r_L^i)$. However, from (1.9) and (1.3), $r_H^* = r_H^i$ means $\lambda^i < \lambda^*$ due to the extra term, $\frac{c}{W_B(r_H, \lambda)} > 0$. Thus;

$$\lambda^i W_B(r_H^i, \lambda^i) < \lambda^* W_B(r_H^*, \lambda^*). \quad (1.10)$$

This implies that the total risky investment is reduced in the innovation case. Since this is true for any arbitrary $r_H^* = r_H^i$ level and since total risky investment level is identical in any equilibrium, this implies that (1.10) always holds. In other words, regulation is not redundant and still has an effect even under the assumption that the bank avoids regulation fully. ■

Using (1.10), the interpretation of the effect of regulation is as follows: After innovating, if the bank invests the same level in the risky project compared with the unregulated case, $\lambda^i = \lambda^*$; from (1.9) and (1.3) the low state return has to be lower $r_L^i < r_L^*$ and the high state return has to be higher $r_H^i > r_H^*$, with this implying a higher variance (i.e. higher risk transferred to HHs). But the bank cannot offer the same $E[r]$ as the unregulated case since the constraints are tighter post-innovation. Thus, HHs will choose a lower level $W_B(r_H, r_L)$. We call this *the composition effect* of the regulation, where HHs funds are diverted away from *BankB* to *BankA*. Since this result is under the assumption of full avoidance, the results hold even stronger when we relax it and assume a partial avoidance. The reason for this is that, under partial avoidance, the bank will consider the probability of getting caught which will decrease the incentives to

²⁷From the HH's canonical portfolio optimization problem, we know that the HH always invest some amount in the risky asset since *B* always offers $E[r] > 1$ in equilibrium.

innovate implying a higher threshold \bar{c} .

Note that since there exists a continuum of equilibria, in an innovation equilibrium it is possible that *BankB* have a higher risky investment than the unregulated equilibria. Even though the risky investment is higher under innovation, still composition effect always materializes since the risky investment level is always lower in any innovation equilibrium. This implies that regardless of the risky investment level innovation equilibria lead to a decreased level of externality.

Several assumptions of the model may need to be discussed. In the model, *BankB* is a monopoly in the risky investment technology. If there were many banks in a perfect competition environment, this would have only changed the final wealth of the consumer and the rest of the analysis would have stayed the same. The assumption that the producer is a monopoly in the model is not restrictive. Even if we assume perfect competition, it will not change our results qualitatively. Under perfect competition, the producer will sell the production good at a price below the maximum of 1 ($p^* < 1$) which can be derived from the zero profit condition. In turn, this implies that for a given date-2 wealth of HHs, the total production will be $\frac{W}{p^*}$, as will change only the total utility of the HHs. All of the remaining analysis will follow.

1.5.1 Comparative Statics and Policy Implications

Observing an innovative product offered by *BankB*, both HHs and the regulator can infer that the cost of innovation is not high. In that case, the effect of regulation may not be too significant if W_B is large.²⁸ This also has an interesting implication as to the market size. It suggests that the larger is the innovative asset market, the higher the incentives to avoid regulation. This is aligned with the recent policy discussions as to downsizing large financial institutions. This further explains how the massive issuance of innovative derivative products (such

²⁸Note, the difference between the constraints in the innovation case pertains to the extra $-c/W_B$ term in r.h.s. Thus, if c is small and W_B is big, the innovation case will not differ significantly from the unregulated case.

as ABSs) was used as a tool to avoid regulation, resulting in excessive risk-taking and externality. Although our model is static, if we think in a dynamic context, as the market size becomes bigger, we will observe more opaque and innovative securities where underlying risks cannot be detected by either the investors or regulatory authorities.

Lemma 1.8 *The ratio of $\frac{c}{W_B}$ will determine the effect of regulation on the total risky investment level. The smaller the ratio, the closer the total risky investment level under the innovation case to the unregulated case, implying a smaller effect of regulation. The smaller the cost c , the bigger the incentives to innovate. Also, the higher the funds (W_B) that can be attracted from HHs, the bigger the incentives.*

Proof. The proof directly follows from the proof of Proposition 1.7. ■

Lemma 1.8 is helpful for the recent policy discussion. It suggests that the regulator should keep an eye on the total wealth invested in the banking sector since huge wealth will imply bigger incentives to innovate and a lower effect of regulation. This is also consistent with the hypothesis that global savings glut incentivized banks to perform innovation.

Note that if cost c is variable and proportional to λW_B then to attract more funds or to invest more in the risky project Bank B bears a larger cost. This implies that the incentives to innovate will be smaller since the bank needs to bear a larger cost to hide its exposure.

1.6 Conclusion

We build a tractable model of two competing banks offering contracts to HHs (as the investors) where one can only offer risk-free contracts due to its lack of access to risky asset technology, while the other offers a risky contract with a higher return. Risk-averse HHs solve a canonical portfolio optimization problem and invest their wealth within the financial sector to trade with a producer (as a representation of the real economy) at a later date. The producer thus faces a

stochastic demand and bears a capacity adjustment cost. Thus, by investing in the risky technology, the risk-taking bank causes an externality on the real sector through the contract it offers to HHs since a higher variance of ex-post wealth implies a more variable demand for the producer. We show that even when the risk-taking bank can completely evade the regulation by offering an innovative product; regulation is still relevant and effective through a composition effect. Thus, we contribute to the literature by offering a possible mechanism, through which policy matters. The model offers policy suggestions such as downsizing too big to fail institutions since the more funds that are attracted by the financial industry, the less effective the regulation is.

Chapter 2

Salience and Financial Fragility

Chapter Abstract

We show that salience theory can explain excess volatility of asset prices and the resulting fire-sales in periods of financial turmoil. We classify the risk into two types, idiosyncratic and systemic, and postulate that investors over-weigh the type of risk that is salient. Either component of risk (systemic or idiosyncratic) becomes salient when its realization deviates sufficiently from a reference point. We show that a change in salience – from one component (idiosyncratic) to the other (systemic) – will generate excess volatility. Interestingly, higher risk aversion generally exacerbates the excess volatility of prices.

2.1 Introduction

One of the proposed explanations for the last financial crisis is that investors, issuers and regulators¹ failed to fully grasp the scale of systemic risk in the financial system. This is often related to new, non-traditional securities, which were issued and bought in large volumes,² and which were believed to be safe. While insurance, tranching and sophisticated financial engineering provide wider idiosyncratic risk diversification, this does not necessarily result in less systemic risk. As a consequence, the financial system in principle can become even more vulnerable to systemic shocks. As occurred during the last financial crisis, with the arrival of negative news, both issuers and investors of these securities realized their true exposure to systemic risk (which had previously been ignored) and started fire-selling their assets, this time possibly worrying excessively about systemic risk.

In this paper, we propose a framework based on salience theory and show that focusing on one type of risk (idiosyncratic or systemic) can explain overpricing of securities ex-ante, and fire-sales at low prices during crisis periods. The central idea is that investors perceive the two types of risk disproportionately and focus primarily on the currently salient one. Either type of risk (systemic or idiosyncratic) is salient when its realization deviates sufficiently from a reference point. We show that assets are mispriced due to the salient thinking of investors, which can create excess price volatility. Considering the detrimental effect of this volatility on the real economy, understanding the driving mechanism is crucial for policy makers. It also provides a rationale for the natural question of why investors were prepared to pay excessively high prices for securities in the first place, only to fire-sell them during the crisis.³

¹Acharya et al. (2017) mentions that the regulatory framework before the crisis did not sufficiently focus on dealing with systemic risk.

²See Figure A.7 and Figure A.8 in Appendix.

³For example, Henderson and Pearson (2011) claim that it is difficult to rationalize the investors buying of overpriced complex securities.

The phenomenon of first ignoring one type of risk and later focusing too much on it can be explained by salience theory, which is based on agents' paying more attention to the salient attributes/states ([Bordalo et al., 2012, 2013](#)). We relate salience to two types of risk; idiosyncratic and systemic. Investors' risk perception is biased towards the type of risk that is currently salient based on prior beliefs or past data.⁴ We show that the diversification fallacy of the pre-crisis period, where seemingly safe assets were overpriced, can be explained by agents over-weighting idiosyncratic risk and ignoring systemic risk. In that sense, systemic assets are overpriced, while non-systemic ones are underpriced. Moreover, our model predicts that if a big systemic shock hits the financial system, salience changes and systemic risk becomes salient for investors. Due to over-weighting of systemic risk, the price for systemic assets falls sharply, similar to a fire-sales phenomenon. This relates to the observed fire-sales of assets during the recent financial crisis. The combination of overpricing ex-ante (during the pre-crisis period) and underpricing ex-post (during the crisis) leads to excess volatility in the market.

We also conduct comparative statics with respect to the model parameters. The most interesting result is that price volatility and mispricing increases with risk aversion. When risk aversion increases, due to mean-variance utility, investors expected utility becomes more sensitive to a change in perceived variance. Initially, when idiosyncratic risk is salient, agents are willing to pay a premium for assets that carry less idiosyncratic risk. The more risk-averse agents are, the more prevalent the mispricing of these assets. However, when a shock hits the market making systemic risk more salient, these investors will try to rebalance their portfolios towards assets that carry less systematic risk. The more risk-averse agents are, the more aggressively they will rebalance their portfolios, thus creating ex-

⁴As it has been discussed, considering the last financial crisis, the lack of availability of past historical default rates for non-traditional securities may have resulted in the neglect of systemic risk. Also see [Gennaioli et al. \(2015\)](#) as an application of the representativeness idea of [Kahneman and Tversky \(1972\)](#) where people overestimate the outcomes that occurred relatively more frequently in recent history.

cess volatility. In contrast, in the extreme case of risk neutrality, investors do not care about which type of risk they are facing - systemic or idiosyncratic - and focus solely on the expected returns. Thus, their demand is unaffected by which type of risk is salient. The policy implications of this paper involve both ex-ante and ex-post policies. In terms of ex-post interventions, the proposed model and results suggest that there is scope for interventions in financial markets such as Troubled Asset Relief Programmes (TARP). However, it should be noted that we do not conduct a welfare analysis, which means that the effectiveness of these interventions should be discussed and analysed in relation to a specific context. In terms of ex-ante policies the study suggests that investors and regulator should use better risk assessment technologies.

2.1.1 Literature Review

There is a vast literature based on the different types of explanations as to the market turmoil observed during the Global Financial Crisis. The contributions can broadly be classified as belonging to one of three strands of the literature. The first, incentive-based explanations mainly focus on the incentives of the banks during the build-up to the crisis - such as risk shifting and moral hazard generated by the originate-to-distribute banking model, (de)regulation and implicit government guarantees (Shleifer and Vishny, 2010; Rajan, 2006; Allen et al., 2015). Shleifer and Vishny (2010) explain the driving source of the instability in the banking system as the short-term extraordinary profit opportunities available to banks due to the huge demand for AAA securities. However, this prompts the question of why investors could not anticipate the fragility of the financial system. It has also been argued that securitization, as one of the possible main drivers of the crisis, created moral hazard (Gorton and Metrick, 2012; Acharya et al., 2013). This was mainly due to the fact that the holders of the liabilities and the issuers of the credits were different entities so that the issuers could sell off the risks

and did not need to worry about the payment ability of the borrowers.⁵ [Beccalli et al. \(2015\)](#) mention that prior to the financial crisis, securitization was generally assumed to have a positive role, mainly in dispersing credit risk and increasing efficiency in risk-sharing. However, with the outbreak of the financial crisis, it has been argued that securitization actually drove down the lending standards ([Greenlaw et al., 2008](#); [Altunbas et al., 2010](#); [Uhde and Michalak, 2010](#)). He claims: “Rather than dispersing the risk, securitization led to a concentration of the risk in the banking sector itself”.

The second type of argument is based on a “global savings glut”, global imbalances in capital flows and the deregulatory environment ([Bernanke et al., 2005](#); [Obstfeld and Rogoff, 2009](#); [Justiniano et al., 2014](#); [Ferrero, 2015](#); [Favilukis et al., 2012](#); [Aizenman and Jinjark, 2009](#)). The basic idea behind these arguments is that there was a search for higher yield worldwide, especially accumulated in the US financial markets. Moreover, the lax regulatory environment exacerbated the problems in the financial system. While the global savings glut explanation may help to understand global imbalances, it may not explain why the prices of these assets collapsed. Thus, the current paper complements this strand of the literature by also explaining a possible mechanism for fire-sales.

Our paper is more closely related to the last class of explanations, namely those models that consider behavioral biases as the main driver of the financial crisis. [Chernenko et al. \(2016\)](#) state that due to the unique feature of credit markets, as skewed payoffs, normal conditions do not convey much information. They further emphasise that “absence of negative experiences” may lead to over-optimism due to investors’ tendency to neglect downside risks ([Greenwood and Shleifer, 2014](#); [Gennaioli et al., 2012, 2015](#)). [Gennaioli et al. \(2012\)](#) show that due to the investor demand for securities that provide safe cash flows, financial intermediaries offer innovative securities that are perceived to be safe but in reality are exposed to neglected risk. This generates financial fragility when investors eventually recog-

⁵Several studies have examined the impact of securitization on monitoring incentives of the banks ([Fender and Mitchell, 2009](#); [Parlour and Plantin, 2008](#); [Plantin, 2011](#)).

nise the neglected state. However, such information neglect may not explain why this mispricing still prevails, even among sophisticated investors.⁶ [Klibanoff et al. \(1998\)](#) argues that the explanations based on the “over/under-reaction” of investors (or markets) can usually only justify ex post anomalies, not the driving mechanisms behind them. This paper offers a complementary idea to this behavioral literature, presenting a possible ex-ante theoretical mechanism based on the well-documented psychological phenomenon of salience (or information availability, as it is sometimes termed). Considering the existence of the diversification fallacy/illusion,⁷ especially during the boom period prior to the financial crisis, this paper builds upon the salience theory to explain the investors’ neglect of systemic risk.

Some studies have explicitly tested salience (or the availability heuristic) empirically ([Tsuji, 2006](#)) or experimentally ([Nelson et al., 2001](#)).⁸ [Andreassen \(1990\)](#) and [Andreassen and Kraus \(1990\)](#), for example, test the effect of the salience of price change information on investor forecasts and trading behaviour. Both papers find evidence in support of the hypothesis that investor forecasts vary as a function of the salience of the information. In focusing on the source of investor reaction in a non-laboratory setting, [Klibanoff et al. \(1998\)](#) test the hypothesis that more importance is assigned to more prominent (salient) news through the use of “closed-end country funds”. They find that salience plays a significant role in determining the magnitude of investor reaction to more prominent news. The current study offers a theoretical background in which the salience heuristic is based upon deeper information, namely risk.

Our paper is also related to the investor sentiment literature. According to classical asset pricing theory, stock prices (especially in the long-run) should reflect the fundamentals since rational investors (or arbitrageurs) will revert prices back

⁶See [Mendel and Shleifer \(2012\)](#) for a detailed discussion as to why the facts of the financial crisis of 2007-09 make the case more interesting in regards to the participants in derivative markets mostly being sophisticated investors.

⁷See [Thakor \(2015\)](#)

⁸Also see [Gärling et al. \(2009\)](#) for an extensive review and discussion of the psychology literature.

to fundamentals, even though in the short-run there may exist deviations. Thus, classical theory suggests that mispricing cannot exist in the long run. However, the vast literature on investor sentiment shows that sentiment has a long-lasting and non-negligible effect on prices. When the share of sentiment investors is significant, trading against sentiment may be costly and risky. As a result, arbitrageurs may not be able to afford to bet against the sentiment and prices may not reflect fundamentals (i.e., there are limits to arbitrage ([Shleifer and Vishny, 1997](#))). The history of financial crises - such as the Great Depression and the Dot.com bubble in the 1990s - validates the premise that prices can deviate from fundamentals for a long duration, especially considering the non-traditional asset markets (such as derivative markets). [Mendel and Shleifer \(2012\)](#) argue that the problems that occurred in the last financial crisis regarding derivative markets are particularly interesting, considering that the investors in these markets are mostly sophisticated investors. They present a model where uninformed rational traders, who only learn from prices, end up chasing the sentiment. Thus, they claim that large numbers of noise traders or large sentiment shocks are not always needed for the sentiment to matter as long as sophisticated investors constitute a large proportion of the traders. [Akhtar et al. \(2012\)](#) find evidence that when sentiment information is released, both futures markets and US stock markets react asymmetrically and that the reaction is more significant among the “stocks that are more salient to investors”. They posit that this can be explained through investors using the availability heuristic. [Baker and Wurgler \(2007\)](#) mention that speculative (riskier) stocks may have lower (higher) return during high (low) sentiment periods, implying that investors over (under) value the stocks when sentiment is high (low). This is perhaps especially true for non-traditional, difficult-to-value assets. As with the investors’ mispricing of stocks during different sentiment periods, in our model, when systemic risk is salient (not salient) investors under-(over-)value systemic stocks. The model offers a theoretical mechanism for how investors over- (under-) value the assets. In this sense, our model of investor salience bias as to perceived risk may fit the sentiment-related empirical evidence

(Palomino et al., 2009; Yu and Yuan, 2011; Akhtar et al., 2012; Stambaugh et al., 2012; Yu and Yuan, 2011; Baker et al., 2012; Da et al., 2014; Huang et al., 2014). To our knowledge, the application of salience bias by defining two types of risk, idiosyncratic and systemic, has not been considered in the sentiment literature.

The remainder of the paper proceeds as follows. The next section, 2.1.2 will briefly introduce the notion of salience in decision-making with the help of a simple example. Section 2.2 then presents the model of salience on different types of risk. Section 2.3 analyses and discusses the results for both the benchmark (rational investor) case and the salient investor case. Section 2.4 presents comparative statics. Section 2.5 discusses the results and mentions policy implications. Finally, Section 2.6 concludes and offers a number of possible future extensions.

2.1.2 Salience Theory

According to Taylor and Thompson (1982), salience refers to a phenomenon in which a decision-maker’s attention is attracted disproportionately by one portion of the environment, which will create disproportionate weighting in the decision-maker’s judgement. Similarly, Kahneman (2011) states that humans have a useful capacity to focus on the odd, different or unusual. This may be due to humans using a heuristic by concentrating their time and attention on salient information, considering their limited cognitive capacity (i.e., focusing on the salient state, attribute or condition provides a useful heuristic) (Akhtar et al., 2012). Based on the salience idea, Bordalo et al. (2012) present a model of choice under risk in which the decision-maker focuses on salient payoffs for a given state and thus evaluates a lottery by inflating the decision weight of the salient states. They emphasise that salience enables the development of a theory based on context-dependent choice aligned with a wide range of evidence. Bordalo et al. (2013) further apply a similar idea to a consumer choice context in which a consumer can pay more or less attention to some attributes of goods such as quality or price. A good’s attribute is salient when it stands out or is unusual as being the

most different from that of a reference good. To summarise this idea, let two attributes for N goods be price and quality.

Salience ranks the attributes of a good based on the reference good's attribute. An attribute is salient if it is significantly different from the average attribute. As an illustration, see the simple example below where there are only two goods, $N = 2$, and two attributes: quality and price:

Example 2.1 *Let the utility of a rational agent be $u_k = q_k - p_k$, where q represents quality and p represents the price of good k . Also, let the average (\bar{q}, \bar{p}) be the reference. To create a ranking between attributes, define a salience function as $S(a_k, \bar{a}) = \frac{|a_k - \bar{a}|}{\bar{a}}$ for $k \in \{1, 2\}, a \in \{q, p\}$. If $|q_1 - q_2| \gg |p_1 - p_2|$, then quality is salient for both goods, since $S(q_k, \bar{q}) > S(p_k, \bar{p})$ for both k .⁹*

If quality is salient, then the consumer will evaluate the utility as follows: $u_k = q_k - \delta p_k$, where $\delta < 1$. In other words, the consumer effectively inflates (deflates) the salient (non-salient) attribute (Bordalo et al., 2013).

2.2 The Model

We consider an OLG model where each generation lives for two periods. Agents (investors) start their lives with an endowment $W > 0$ and have mean-variance utility.¹⁰ The endowment W cannot be stored. They invest their endowment when young and consume when old. Each period, the young investors optimally choose their portfolio from N different risky assets acquired from the old gener-

⁹Note that when we have only two goods with two attributes, if one attribute is significantly different between the two goods compared to the other attribute, then it will stand out as the salient attribute for both goods due to symmetry.

¹⁰W.l.g let $W = \tilde{W} + \bar{c}$, and let us assume that the young generation consumes a fixed \bar{c} and invests the leftover. Thus, in a sense, we merely focus on the asset choice of the investors and isolate our discussion from the optimal consumption choice for the newborns. When a generation arrives in the next period (i.e. the old generation), they consume everything they have after selling their assets to the newborns.

ation, all assumed to be in fixed supply. We will consider $N = 3$.¹¹ Each asset has a deterministic dividend K , and a stochastic dividend. We assume that K is sufficiently large in order to ensure that in the optimisation problem of the investors (given below in (2.1)) the expected utility is non-negative in equilibrium, and prices of all assets are non-negative.

The stochastic dividend is formed of two parts, representing idiosyncratic risk and systemic risk respectively.

$$R_j = K + \varepsilon_j + \alpha_j \mu, \quad \varepsilon_j \sim (0, \sigma_j^2), \quad \mu \sim (0, \sigma_\mu^2) \quad j \in N$$

where ε is an idiosyncratic shock (ε_j are independently distributed and independent from μ) and μ is the systemic shock. Each asset thus has a stochastic part which is undiversifiable. If the coefficient α_j is large, then asset j entails more systemic risk. We do not assume any specific probability distribution for the stochastic parts but only assume a zero mean and finite variance. The time subscripts are omitted for R_j since the payout structure is identical for each period t .

For $N = 3$;

$$\begin{aligned} R_A &= K + \varepsilon_A + \alpha \mu, & E[\varepsilon_A] &= 0, & Var(\varepsilon_A) &= \sigma_A^2 \\ R_B &= K + \varepsilon_B + \beta \mu, & E[\varepsilon_B] &= 0, & Var(\varepsilon_B) &= \sigma_B^2 \\ R_C &= K + \varepsilon_C + \theta \mu, & E[\varepsilon_C] &= 0, & Var(\varepsilon_C) &= \sigma_C^2 \end{aligned}$$

We assume $\sigma_C^2 > \sigma_B^2 > \sigma_A^2$, $\alpha > \beta > \theta$ and $\sigma_A^2 + \alpha^2 \sigma_\mu^2 = \sigma_B^2 + \beta^2 \sigma_\mu^2 = \sigma_C^2 + \theta^2 \sigma_\mu^2$, so that $Var(R_A) = Var(R_B) = Var(R_C)$, i.e. all assets have identical variance. The assumption that all assets have identical variance is mostly for expositional convenience as it implies that if an asset has higher idiosyncratic risk, it must carry lower systemic risk. As a result, it is immaterial whether assets are ordered in terms of either type of risk.¹² For the rest of the paper, asset A , which carries

¹¹The reason we did not pick $N = 2$ is because when there are only 2 assets, their payouts are directly correlated due to the modelling of systemic risk. Thus, we need more than 2 assets to have a sensible systemic risk concept. Also our results carry through for the more generic case of $N > 3$.

¹²This simplification is to have a reasonable comparison between the prices of the assets in

the larger amount of systemic risk, will be called as the “systemic asset”.

The assets are in fixed supply. In each period t , the old generation (who were born at $t - 1$) will sell the assets at any non-negative price since they will consume anything they have invested. Thus, high or low demand will determine the prices at the end of each period. The young generation born at t choose the optimal portfolio by demanding x_A, x_B, x_C quantities of each asset (based on a prior regarding salience, which will be explained later).¹³ Prices are determined in any period, by market clearing.

At the beginning of each period, young investors solve the following problem:

$$\begin{aligned} \max_{x_{t,A}, x_{t,B}, x_{t,C} \in [0, W_t]} & E[W_{t+1}] - \gamma Var(W_{t+1}) \\ \text{s.t.} & \end{aligned} \quad (2.1)$$

,

$$x_{t,A}p_{t,A} + x_{t,B}p_{t,B} + x_{t,C}p_{t,C} = W \quad (2.2)$$

where $E[W_{t+1}] = (x_{t,A} + x_{t,B} + x_{t,C})K + W$. This is because investors ex-ante expect the prices to stay same, i.e. $E[p_{t+1,i}] = p_{t,i}$ since they do not predict salience to change. To see this point, note that $E[W_{t+1}] = E[\sum_j p_{t+1,i}x_{t,i} + \sum_j R_i x_{t,i}] = W_t + E[\sum_j R_i x_{t,i}] = W + (x_{t,A} + x_{t,B} + x_{t,C})K$

2.2.1 Salience

Salience works as follows; depending on which risk is salient, the perceived variance may differ. Hence, $Var(W_{t+1})$ will depend on investors’ perceptions of risk. The investor can perceive either idiosyncratic risk ε_j as salient or systemic risk μ as salient. Investors will over-weigh the salient risk and under-weigh the non-

equilibrium. When assets have different variances, due to variance size effect, the prices will be different. By assuming that all assets have the same total variance (total risk including systemic and idiosyncratic), one can say that the difference in prices directly results from a different combination of systemic and idiosyncratic risks.

¹³Having the demands as functions of prices and parameters, equilibrium prices will be determined in consideration that the supply is fixed.

salient risk.¹⁴ As with Bordalo et al. (2012, 2013), we will introduce a salience function in Section 2.3.2 which basically produces a ranking between two types of risk. Since we have only two elements to be ordered, the more salient element will be over-weighted and the other will be under-weighted.

According to salient decision-making, the stochastic dividends of the assets will be weighed as follows:

- When μ is salient, $R_j = K + \frac{1}{\delta}\varepsilon_j + \delta(\alpha_j\mu)$
- When ε is salient, $R_j = K + \delta\varepsilon_j + \frac{1}{\delta}\alpha_j\mu$

where $\delta > 1$ ¹⁵ and $j \in \{A, B, C\}$. This implies that the salient risk will be over-weighted.¹⁶ From a dynamic viewpoint, there can be four different cases: a) ε is salient at time t but μ becomes salient at time $t + 1$, b) μ is salient at time t but ε becomes salient at time $t + 1$, c) and d) ε or μ is salient throughout.

Depending on what is currently salient, investors will have different perceived variances. Formally, the variances as perceived by a rational agent, an agent who perceives ε as salient, and an agent who perceives μ as salient are as follows.

- Rational agents: $Var(W_2)^R = x_A^2\sigma_A^2 + x_B^2\sigma_B^2 + x_C^2\sigma_C^2 + (\alpha x_A + \beta x_B + \theta x_C)^2\sigma_\mu^2$
- ε -salient: $Var(W_2)^\varepsilon = \omega(x_A^2\sigma_A^2 + x_B^2\sigma_B^2 + x_C^2\sigma_C^2) + \frac{(\alpha x_A + \beta x_B + \theta x_C)^2\sigma_\mu^2}{\omega}$
- μ -salient: $Var(W_2)^\mu = \frac{(x_A^2\sigma_A^2 + x_B^2\sigma_B^2 + x_C^2\sigma_C^2)}{\omega} + (\alpha x_A + \beta x_B + \theta x_C^2\sigma_\mu^2)\omega$

where $\omega = \delta^2 > 1$.

We assume;

$$\frac{\sigma_A^2\omega}{\alpha(\alpha + \beta + \theta)} < \sigma_\mu^2 < \frac{\sigma_C^2}{\omega\theta(\alpha + \beta + \theta)} \quad (2.3)$$

Note that, by construction, $\sigma_A^2/\alpha < \sigma_C^2/\theta$. Hence, essentially we assume that

¹⁴This over/under weighing is aligned with the context dependency, as in Prospect Theory, since salience is also based on a reference point.

¹⁵Note that this modelling of δ generates salience bias in perceived risk. We can also generate the results in this paper qualitatively by modelling the bias as follows. $R_j = K + (1 - \delta)\varepsilon_j + \delta(\alpha_j\mu)$ where $1/2 < \delta < 1$.

¹⁶The model can be modified so that historical information forms the prior and the salience weight. One can build a dynamic model in which the historical realisation of risk may be modelled to increase the probability of being salient.

ω is not too large (i.e. salience is not extreme) and that the systemic variance is neither too large or too small relative to the idiosyncratic variances. This restriction is required due to the specific way the salience is modelled. If salience bias is too extreme; for example, consider arbitrarily large ω , then the investors perceive the variance of their portfolio arbitrarily large as well, independent of the assets they hold and independent of what is salient. This implies that, a systemic asset (A) with non-zero idiosyncratic variance could be in principle perceived as excessively risky even when idiosyncratic risk is salient.¹⁷

2.3 Results

2.3.1 Rational Investors

We start with the benchmark case, where rational young investors at t solve (2.1) using the actual variance $Var(W_2)^R$. Given the optimal portfolio allocation $(x_A^R(p_A^R, p_B^R, p_C^R), x_B^R(p_A^R, p_B^R, p_C^R), x_C^R(p_A^R, p_B^R, p_C^R))$ and solving for the equilibrium prices by assuming a fixed supply of 1 unit for each asset; at t ¹⁸

Proposition 2.1 *In the rational benchmark case prices for each asset are such that;*

$$p_i^R = \frac{W(K - 2\gamma(\alpha_i \sigma_\mu^2(\alpha + \beta + \theta) + \sigma_i^2))}{3K - 2\gamma(\sigma_\mu^2(\alpha + \beta + \theta)^2 + (\sigma_A^2 + \sigma_B^2 + \sigma_C^2))}$$

for $i \in \{A, B, C\}$, $\alpha_i \in \{\alpha, \beta, \theta\}$ and for any time period.¹⁹

Assuming that the total variance is the same for all assets, the systemic asset (A) has the lowest price, that is;

$$p_A^R < p_B^R < p_C^R.$$

Proof. For the benchmark case, rational investors solve (2.1) considering the

¹⁷Note that arbitrarily large ω may also imply negative prices.

¹⁸Time subscripts are ignored in the notation, in order to keep it simple.

¹⁹For the generic N asset case: $p_i^R = \frac{W(K - 2\gamma(\alpha_i \sigma_\mu^2(\sum_{j=1}^N \alpha_j) + \sigma_i^2))}{NK - 2\gamma(\sigma_\mu^2(\sum_{j=1}^N \alpha_j)^2 + \sum_{j=1}^N \sigma_j^2)}$ for $i \in N$ and for any time period, where α_i stands for the correlation with systemic risk for asset i

actual variance of the portfolio $Var(W_2)^R$. Then, having the demands for each asset and assuming a fixed 1 unit supply for each, we get the prices;

$$\begin{aligned} p_A^R &= \frac{W(K - 2\gamma(\sigma_\mu^2\alpha(\alpha + \beta + \theta) + \sigma_A^2))}{3K - 2\gamma(\sigma_\mu^2(\alpha + \beta + \theta)^2 + (\sigma_A^2 + \sigma_B^2 + \sigma_C^2))} \\ p_B^R &= \frac{W(K - 2\gamma(\sigma_\mu^2\beta(\alpha + \beta + \theta) + \sigma_B^2))}{3K - 2\gamma(\sigma_\mu^2(\alpha + \beta + \theta)^2 + (\sigma_A^2 + \sigma_B^2 + \sigma_C^2))} \\ p_C^R &= \frac{W(K - 2\gamma(\sigma_\mu^2\theta(\alpha + \beta + \theta) + \sigma_C^2))}{3K - 2\gamma(\sigma_\mu^2(\alpha + \beta + \theta)^2 + (\sigma_A^2 + \sigma_B^2 + \sigma_C^2))}. \end{aligned}$$

For K large enough, the expected utility is positive in equilibrium, which also implies both the numerator and denominator are positive for each price. Assuming assets have the same total variance, $\sigma_A^2 + \alpha^2\sigma_\mu^2 = \sigma_B^2 + \beta^2\sigma_\mu^2 = \sigma_C^2 + \theta^2\sigma_\mu^2$, and using the prices found above, we obtain the following; $p_A^R < p_B^R < p_C^R$. This holds since $\sigma_C^2 > \sigma_B^2 > \sigma_A^2$ and $\alpha > \beta > \theta$.²⁰ ■

The intuition of this proposition is simple. The first part basically says that if investors are rational, then the prices will be stable over time (i.e., they will be the same in any time period). Rational investors will perceive the same variance of assets in any period so that their demand will be the same in each period. Considering that the supply is fixed, this will imply stable prices over time. For the second part, the core of the intuition is that the idiosyncratic risk is partially diversifiable whereas the systemic is not. Thus, if the total variance is the same for all assets, in having higher systemic risk, asset A will have lower demand and consequently a lower price compared to other assets.

2.3.2 Salient Investors

This section provides the results for the case when investors have salience bias on the perceived risk. At any period, when the realized systemic risk is low enough, investors are defined as ε -salient since they will overweigh the idiosyncratic risk.

²⁰Note that the main results of the paper namely mispricing and the excess volatility can be also generated by using only two assets. In other words, Asset B can be removed from the analysis. But we need the third asset in order obtain unequal asset prices when investors are rational.

When the realised systemic risk is sufficiently large ($\mu > \hat{\mu}$), then systemic risk will become salient and the investors are defined as μ -salient. ε -salient investors consider $Var(W_2)^\varepsilon$ and μ -salient investors consider $Var(W_2)^\mu$ as the perceived risk for their optimal portfolio demands.

Lemma 2.2 *When idiosyncratic risk (ε) is salient, the systemic asset A is overpriced $p_A^\varepsilon > p_A^R$ and the non-systemic asset C is underpriced $p_C^\varepsilon < p_C^R$. Moreover, $p_A^\varepsilon > p_B^\varepsilon > p_C^\varepsilon$ for $\omega > \sqrt{\frac{\alpha + \beta + \theta}{\alpha + \beta}}$.*

Proof. For ε -salient investors, using $Var(W_2)^\varepsilon$ in (2.1) we get the demands for each asset. Assuming a fixed 1 unit supply for each asset we get the prices as follows;

$$\begin{aligned} p_A^\varepsilon &= \frac{W(K - 2\gamma(\frac{\sigma_\mu^2 \alpha(\alpha + \beta + \theta)}{\omega} + \omega \sigma_A^2))}{3K - 2\gamma(\frac{\sigma_\mu^2 (\alpha + \beta + \theta)^2}{\omega} + \omega(\sigma_A^2 + \sigma_B^2 + \sigma_C^2))}, \\ p_B^\varepsilon &= \frac{W(K - 2\gamma(\frac{\sigma_\mu^2 \beta(\alpha + \beta + \theta)}{\omega} + \omega \sigma_B^2))}{3K - 2\gamma(\frac{\sigma_\mu^2 (\alpha + \beta + \theta)^2}{\omega} + \omega(\sigma_A^2 + \sigma_B^2 + \sigma_C^2))}, \\ p_C^\varepsilon &= \frac{W(K - 2\gamma(\frac{\sigma_\mu^2 \theta(\alpha + \beta + \theta)}{\omega} + \omega \sigma_C^2))}{3K - 2\gamma(\frac{\sigma_\mu^2 (\alpha + \beta + \theta)^2}{\omega} + \omega(\sigma_A^2 + \sigma_B^2 + \sigma_C^2))}. \end{aligned}$$

For K large enough both the numerators and denominators are positive for each price. Using the assumption that $\sigma_C^2 > \sigma_B^2 > \sigma_A^2$ and $\alpha > \beta > \theta$, we get $p_A^\varepsilon > p_B^\varepsilon > p_C^\varepsilon$.

In order to show that $p_A^\varepsilon > p_A^R$ we need

$$W(K - 2\gamma(\frac{\sigma_\mu^2 \alpha(\alpha + \beta + \theta)}{\omega} + \omega \sigma_A^2)) > W(K - 2\gamma(\sigma_\mu^2 \alpha(\alpha + \beta + \theta) + \sigma_A^2))$$

which implies that;

$$\sigma_\mu^2 \alpha(\alpha + \beta + \theta) > \sigma_A^2 \omega. \quad (2.4)$$

However, condition (2.3) implies that (2.4) is satisfied.

Again, in order to show that $p_C^\varepsilon < p_C^R$ we need

$$W(K - 2\gamma(\frac{\sigma_\mu^2\theta(\alpha + \beta + \theta)}{\omega} + \omega\sigma_C^2)) < W(K - 2\gamma(\sigma_\mu^2\alpha(\alpha + \beta + \theta) + \sigma_C^2))$$

which implies that;

$$\sigma_\mu^2\theta(\alpha + \beta + \theta) < \sigma_C^2\omega \quad (2.5)$$

However, condition (2.3) implies that (2.5) is satisfied since $\omega > 1$. ■

Lemma 2.2 basically states that when idiosyncratic risk is salient, investors overweigh idiosyncratic risk and underweigh the systemic one. Thus the systemic asset, A, is overpriced whereas the non-systemic asset, C, is underpriced.

Lemma 2.3 *When systemic risk is salient, the systemic asset A is underpriced $p_A^\mu < p_A^R$, and the non-systemic asset C is overpriced $p_C^\mu > p_C^R$. Moreover $p_A^\mu < p_B^\mu < p_C^\mu$.*

Proof. For μ – salient investors, using $Var(W_2)^\mu$ in (2.1) we get the demands for each asset. Assuming a fixed 1 unit supply for each asset, we get the prices as follows;

$$\begin{aligned} p_A^\mu &= \frac{W(K - 2\gamma(\omega\sigma_\mu^2\alpha(\alpha + \beta + \theta) + \frac{\sigma_A^2}{\omega}))}{3K - 2\gamma(\omega\sigma_\mu^2(\alpha + \beta + \theta)^2 + \frac{(\sigma_A^2 + \sigma_B^2 + \sigma_C^2)}{\omega})}, \\ p_B^\mu &= \frac{W(K - 2\gamma(\omega\sigma_\mu^2\beta(\alpha + \beta + \theta) + \frac{\sigma_B^2}{\omega}))}{3K - 2\gamma(\omega\sigma_\mu^2(\alpha + \beta + \theta)^2 + \frac{(\sigma_A^2 + \sigma_B^2 + \sigma_C^2)}{\omega})}, \\ p_C^\mu &= \frac{W(K - 2\gamma(\omega\sigma_\mu^2\theta(\alpha + \beta + \theta) + \frac{\sigma_C^2}{\omega}))}{3K - 2\gamma(\omega\sigma_\mu^2(\alpha + \beta + \theta)^2 + \frac{(\sigma_A^2 + \sigma_B^2 + \sigma_C^2)}{\omega})}. \end{aligned}$$

For K large enough both the numerators and denominators are positive for each price. Using the assumption that $\sigma_C^2 > \sigma_B^2 > \sigma_A^2$ and $\alpha > \beta > \theta$, we get $p_A^\mu < p_B^\mu < p_C^\mu$.

In order to show that $p_A^\mu < p_A^R$ we need

$$W(K - 2\gamma(\omega\sigma_\mu^2\alpha(\alpha + \beta + \theta) + \frac{\sigma_A^2}{\omega})) < W(K - 2\gamma(\sigma_\mu^2\alpha(\alpha + \beta + \theta) + \sigma_A^2))$$

which implies that;

$$\sigma_\mu^2 \alpha (\alpha + \beta + \theta) > \frac{\sigma_A^2}{\omega} \quad (2.6)$$

However, condition (2.3) implies that (2.6) is satisfied since $\omega > 1$.

Again, in order to show that $p_C^\mu > p_C^R$ we need

$$W(K - 2\gamma(\omega\sigma_\mu^2\theta(\alpha + \beta + \theta) + \frac{\sigma_C^2}{\omega})) < W(K - 2\gamma(\sigma_\mu^2\alpha(\alpha + \beta + \theta) + \sigma_A^2))$$

which implies that;

$$\sigma_\mu^2\theta(\alpha + \beta + \theta) < \frac{\sigma_C^2}{\omega} \quad (2.7)$$

However, condition (2.3) implies that (2.7) is satisfied. ■

Lemma 2.3 basically states that when systemic risk is salient, investors overweigh systemic risk and underweigh the idiosyncratic one. Thus, the systemic asset, A, is underpriced whereas the non-systemic asset, C, is overpriced. The intuition is simple. In the benchmark case, asset A has the lowest price due to its highest systemic correlation coefficient. When systemic risk is salient, the demand for asset A is even smaller when compared to that of the benchmark case.

Proposition 2.4 *Assuming $\text{Var}(R_A) = \text{Var}(R_B) = \text{Var}(R_C)$, we obtain $p_A^\varepsilon > p_A^\mu$ and $p_C^\varepsilon < p_C^\mu$:*

- *If at time t , ε is salient and μ becomes salient at $t + 1$, then the price p_A of the systemic asset (A) decreases and the price p_C of the less systemic asset (C) increases.*
- *If at time t , μ is salient and ε becomes salient at $t + 1$, then the price p_A of the systemic asset (A) increases and the price p_C of the less systemic asset (C) decreases.*
- *In both cases, price volatility is higher than in the benchmark case of full rationality.*

Proof. The proof directly follows from the Lemma 1 and Lemma 2 above. For

the benchmark case, since prices stay the same over time, any price deviation after a change in salience implies excess volatility in prices. ■

If investors start as ε -salient at t , there are two cases that can occur in the subsequent period. If investors remain ε -salient, the price will not change since the young investors will have the same perceived variance. In contrast, if they switch to μ -salience, the young investors will use $Var(W_{t+1})^\mu$ instead of $Var(W_{t+1})^\varepsilon$, which will generate price volatility.

The main implication of this result is as follows. In taking no stance on how investors form their prior, we conclude that when the salient feature of the risk (idiosyncratic or systemic) changes, investors change their perceived risk for the assets, which results in price volatility in equilibrium. The main insight is that due to the bias in perceived risk, there is excess volatility in the asset markets.

In order to be more specific on the change in salience and how salience may arise, define a salience function $S(a_k, a_j, \bar{a})$ where $a \in \{\varepsilon, \mu\}$, $k, j \in \{A, C\}$ and \bar{a} represents the reference point.²¹ The role of this function is to produce a ranking between the two types of risk, idiosyncratic and systemic. In that sense, whenever $S(\varepsilon_A, \varepsilon_C, \bar{\varepsilon}) < (>) S(\mu_A, \mu_C, \bar{\mu})$, then systemic (idiosyncratic) risk will be salient for assets $j \in \{A, C\}$.

In the model, upon the realization of the stochastic components of dividends, similar to an information arrival, investors will update their perceived variance. For example, if the realized systemic part μ is bigger than some value $\hat{\mu}$, investors will focus on the systemic risk and the systemic part will be salient.

Considering the result in Proposition 2.4 above, the implication of the above salience function is that salience changes based on the realized values of ε and μ . If at time t ,

- ε is salient, after a systemic shock ($\mu > \hat{\mu}$) μ becomes salient at $t+1$. Then, the price of the systemic asset p_A decreases and the price of the non-systemic asset p_C , increases.
- μ is salient, after an idiosyncratic shock ε becomes salient at $t+1$. Then, the

²¹The reference point can be defined as the average of all assets or can be a specific asset.

price of the systemic asset p_A increases and the price of the non-systemic asset p_C decreases.

2.4 Comparative Statics

The results from the previous section show that there is excess volatility in prices due to investor salience. We now conduct comparative statics analysis on price volatility. Let $P(\varepsilon\text{-salient})$ and $P(\mu\text{-salient})$ be the probabilities that in a given time period, ε is salient or μ is salient, respectively. W.l.g assume that investors are ε -salient at t . Then we can proxy the price volatility of each asset $j \in \{A, B, C\}$ by $|p_{t,j}^\varepsilon - E[p_{t+1,j}]|$,²² where $E[p_{t+1,j}] = P(\varepsilon\text{-salient})p_{t+1,j}^\varepsilon + P(\mu\text{-salient})p_{t+1,j}^\mu$. However, since $P(\varepsilon\text{-salient}) = 1 - P(\mu\text{-salient})$ and since the price will not change unless salience changes at $t + 1$ (i.e. $p_{t,j}^\varepsilon = p_{t+1,j}^\varepsilon$), we have;

$$|p_{t,j}^\varepsilon - E[p_{t+1,j}]| = P(\mu\text{-salient})|p_{t+1,j}^\mu - p_{t,j}^\varepsilon|.$$

We will focus on the difference $|p_{t+1,j}^\mu - p_{t,j}^\varepsilon|$ without imposing a structure on the probabilities and look at the comparative statics w.r.t to the model parameters γ (risk aversion), σ_μ^2 (systemic risk), σ_j^2 (idiosyncratic risks), ω (salience weight) and K . We are thus implicitly assuming that these exogenous parameters do not effect the salience function (and thus the probabilities). Time subscripts are omitted for simplification since $p_{t,j}^\varepsilon = p_{t+1,j}^\varepsilon$ and $p_{t,j}^\mu = p_{t+1,j}^\mu$ for all assets and all periods t .

For ease of comparison, we restrict $Var(W_2)^\varepsilon$ and $Var(W_2)^\mu$ to have the same value. This implies $\sigma_A^2 + \sigma_B^2 + \sigma_C^2 = (\alpha + \beta + \theta)^2 \sigma_\mu^2$. In other words, we are considering the case where salience does not affect the investors' perception of the total risk that they face. This restriction merely normalizes perceived variances

²²We use absolute difference rather than $(p_{t,j}^\varepsilon - E[p_{t+1,j}])^2$ since the results do not change qualitatively.

$Var(W_2)^\varepsilon$ and $Var(W_2)^\mu$ so that they are comparable.²³

Proposition 2.5 *Price volatility $|p_j^\varepsilon - p_j^\mu|$ increases with γ for all assets $j \in \{A, B, C\}$.*

Proof. The assumption that $\sigma_A^2 + \sigma_B^2 + \sigma_C^2 = (\alpha + \beta + \theta)^2 \sigma_\mu^2$ simplifies the denominator of $|p_j^\varepsilon - p_j^\mu|$ as $3K - 2\gamma(\omega\sigma_\mu^2(\alpha + \beta + \theta)^2 + \frac{(\sigma_A^2 + \sigma_B^2 + \sigma_C^2)}{\omega})$. Consider asset A . As shown in Proposition 2.4, $p_A^\varepsilon > p_A^\mu$. Thus, use $|p_A^\varepsilon - p_A^\mu| = p_A^\varepsilon - p_A^\mu$.

$$p_A^\varepsilon - p_A^\mu = \frac{2W\gamma(\sigma_\mu^2\alpha(\alpha + \beta + \theta) - \sigma_A^2)\frac{\omega^2-1}{\omega}}{3K - 2\gamma(\omega\sigma_\mu^2(\alpha + \beta + \theta)^2 + \frac{(\sigma_A^2 + \sigma_B^2 + \sigma_C^2)}{\omega})}. \quad (2.8)$$

Note that for $\sigma_\mu^2 > \frac{\sigma_A^2}{\alpha(\alpha + \beta + \theta)}$ (i.e., for systemic risk not very small, which is satisfied due to condition (2.3)), when γ increases, the numerator increases whereas the denominator decreases. Thus, $\frac{d|p_A^\varepsilon - p_A^\mu|}{d\gamma} > 0$.

Formally;

$$\frac{d(p_A^\varepsilon - p_A^\mu)}{d\gamma} = \frac{6KW\omega(\omega^2 - 1)(\sigma_\mu^2\alpha(\alpha + \beta + \theta) - \sigma_A^2)}{((2\gamma\omega^2(\sigma_A^2 + \sigma_B^2 + \sigma_C^2) - 3K\omega) + 2\gamma\sigma_\mu^2(\alpha + \beta + \theta)^2)^2} > 0$$

for $\sigma_\mu^2 > \frac{\sigma_A^2}{\alpha(\alpha + \beta + \theta)}$.

Consider now asset C . As shown in Proposition 2.4, $p_C^\varepsilon < p_C^\mu$. Thus, use $|p_C^\varepsilon - p_C^\mu| = p_C^\mu - p_C^\varepsilon$.

$$p_C^\mu - p_C^\varepsilon = \frac{2W\gamma(\sigma_C^2 - \sigma_\mu^2\theta(\alpha + \beta + \theta))\frac{\omega^2-1}{\omega}}{3K - 2\gamma(\omega\sigma_\mu^2(\alpha + \beta + \theta)^2 + \frac{(\sigma_A^2 + \sigma_B^2 + \sigma_C^2)}{\omega})}. \quad (2.9)$$

Note that for $\sigma_C^2 > \sigma_\mu^2\theta(\alpha + \beta + \theta)$ (i.e., for the idiosyncratic risk of asset C not being very small, which is satisfied due to condition (2.3)), when γ increases, the numerator increases whereas the denominator decreases. Thus $\frac{d|p_C^\varepsilon - p_C^\mu|}{d\gamma} > 0$.

The proof is analogous for asset B . ■

²³Investors have mean-variance utility, which implies that as variance increases the demand decreases. Thus, to have a meaningful comparison between ε - and μ -salient cases, the variances should be comparable. Given that there is a fixed supply of assets, this restriction allows us to compare p_j^ε and p_j^μ by allowing $Var(W_2)^\varepsilon = Var(W_2)^\mu$.

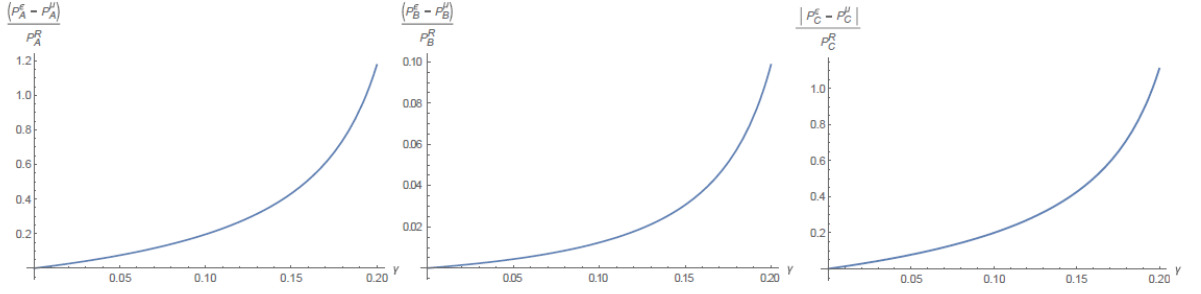


Figure 2.1: Price volatility w.r.t risk aversion γ , normalized by dividing into the rational (benchmark) prices p_j^R . The parameters used for these graphs ensure that the expected utility is non-negative for the maximization problem (2.1).

Thus, price volatility increases with risk aversion γ , which at first glance may seem counter-intuitive.²⁴ The intuition for this result lies within how we model mispricing. We can understand this by focusing on asset A as an illustration. Investors overprice the systemic asset (A) when idiosyncratic risk is salient and underprice it when systemic risk is salient (i.e. $p_A^\epsilon > p_A^\mu$). When risk aversion is high, overpricing and underpricing are more significant due to the investors' expected utility being more sensitive to risks. This implies that $|p_A^\epsilon - p_A^\mu|$ increases with risk aversion. To better grasp this intuition, consider the scenario of $\gamma = 0$. When investors are risk neutral, they do not care about the difference between the assets as long as their expected returns are the same. Thus, no risk aversion implies no price deviation, i.e. $p_A^\epsilon = p_A^\mu$.

This result holds for any asset. When risk aversion increases, due to mean-variance utility, investors' expected utility becomes more sensitive to a change in perceived variance. Initially, when idiosyncratic risk is salient, agents are willing to pay a premium for assets that carry less idiosyncratic risk. The more risk averse agents are, the more prevalent the mispricing of these assets. However, when a shock hits the market making systemic risk more salient, these investors will try to rebalance their portfolios towards assets that carry less systematic risk. The more risk averse agents are, the more aggressively they will rebalance their

²⁴Note that this is not a size effect. Actually size effect works in the opposite direction. When risk aversion increases, investors have a smaller demand and this results in smaller prices.

portfolios, thus creating excess volatility.

Also, see Figure 2.1 for a graphical illustration of this result, whereby the price volatilities of the assets are normalized by dividing into the benchmark prices p_j^R . This normalization helps us to understand that for asset A (left figure) and asset C the price volatility is significant as a percentage of rational prices. For asset B (middle figure) the change is not large (%10 at maximum) since asset B is like an average asset.

Lemma 2.6 *Price volatility $|p_j^\varepsilon - p_j^\mu|$ increases with σ_μ^2 for all assets $j \in \{A, B, C\}$*

Proof. This result follows from (2.8). Again, given $\sigma_\mu^2 > \frac{\sigma_A^2}{\alpha(\alpha+\beta+\theta)}$ (condition (2.3)), when σ_μ^2 increases the numerator of (2.8) increases whereas the denominator decreases. Thus, $\frac{d|p_A^\varepsilon - p_A^\mu|}{d\sigma_\mu^2} > 0$.

The proof is analogous for assets B and C. ■

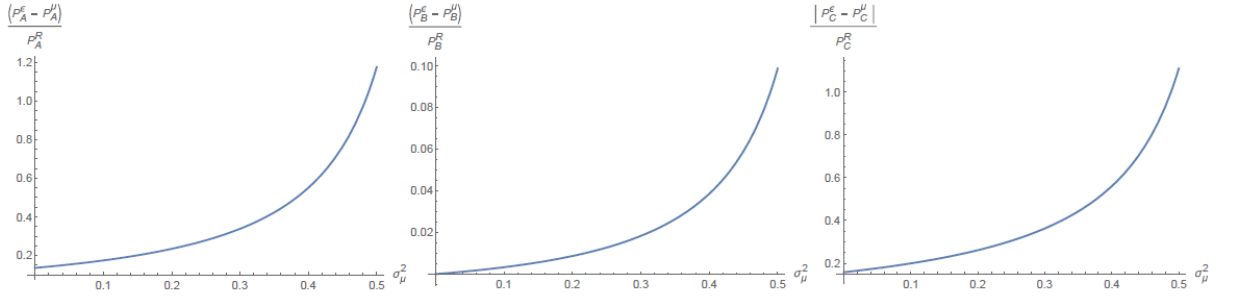


Figure 2.2: Price volatility with respect to systemic risk σ_μ^2 , normalized by dividing into the rational (benchmark) prices p_j^R .

As systemic risk σ_μ^2 increases, the price of assets will have more volatility between two different salient cases. Due to investors ignoring systemic risk, when it is not salient, mispricing is more detrimental when σ_μ^2 is bigger. The intuition is that when systemic risk is larger, the gap between ε -salient and μ -salient cases is more significant since the perceived systemic parts will be more different. This can be observed from Figure 2.2 as well, where again to gain an idea as to the size of the volatility we divide them by the rational prices. Again, note that the volatility is significant for assets A and C compared to the benchmark prices.

Note that in both Figures 2.1 and 2.2, the price volatility of asset A is larger (in terms of % of rational price). This implies that, A, being the systemic asset experiences larger volatility when either the systemic risk or the risk aversion parameter increases. This result may produce interesting policy implications. For example regulators may want to monitor the systemic assets/products more intensely.

Lemma 2.7 *For asset C, price volatility $|p_C^\varepsilon - p_C^\mu|$ increases in σ_C^2 . For asset A, price volatility $|p_A^\varepsilon - p_A^\mu|$ can increase or decrease in σ_A^2 depending on parameter values.*

Proof. For asset C, the proof is straightforward. From (2.9), it is obvious that the numerator is increasing with σ_C^2 whereas the denominator is decreasing. Thus, $\frac{d|p_C^\varepsilon - p_C^\mu|}{d\sigma_C^2} = \frac{dp_C^\mu - p_C^\varepsilon}{\sigma_A^2} > 0$.

Now consider asset A. Using equation (2.8) and given $p_A^\varepsilon > p_A^\mu$, we obtain;

$$\frac{d|p_A^\varepsilon - p_A^\mu|}{d\sigma_A^2} = \frac{2\gamma W((\omega^2 - 1)(\omega(2\gamma\omega(\sigma_B^2 + \sigma_C^2) - 3K) + 2\gamma\sigma_\mu^2(\alpha + \beta + \theta)(\alpha\omega^2 + \alpha + \beta + \theta))}{(2\sigma_\mu^2\gamma(\alpha + \beta + \theta)^2 + \omega(-3K + 2(\sigma_A^2 + \sigma_B^2 + \sigma_C^2)\gamma\omega))^2}$$

Thus, $\frac{d|p_A^\varepsilon - p_A^\mu|}{d\sigma_A^2} > 0$ as long as σ_μ^2 is not small, $\sigma_\mu^2 > \frac{\omega(3K - 2\gamma\omega(\sigma_B^2 + \sigma_C^2))}{2\gamma(\alpha + \beta + \theta)(\alpha\omega^2 + \alpha + \beta + \theta)}$. ■

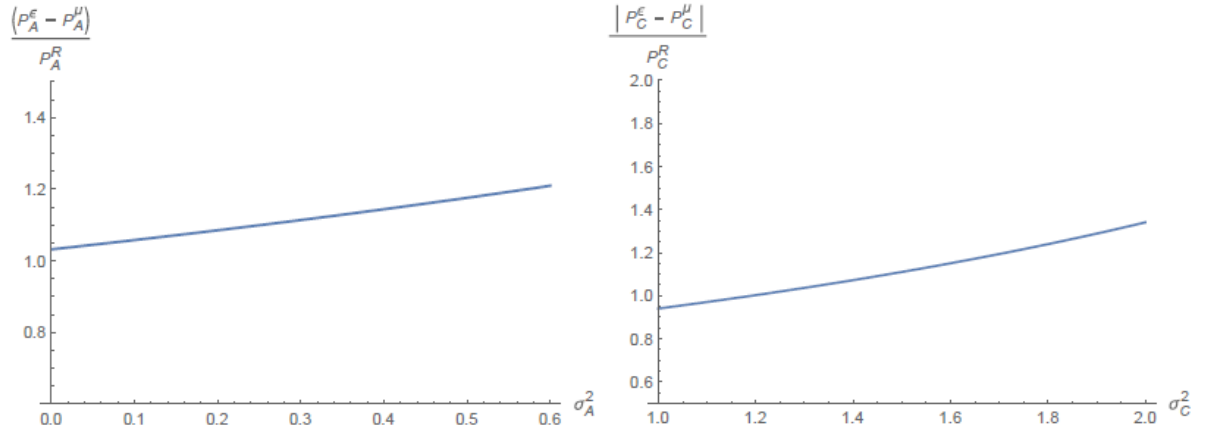


Figure 2.3: Price volatility with respect to idiosyncratic risks

Figure 2.3 shows that as σ_C^2 increases, price volatility increases as well for asset C. The intuition of this lies within the weighing parameter. Since salient

risk is over-weighted and the other risk is under-weighted, a larger idiosyncratic variance implies a higher gap between the perceived variances of the assets for the different salient cases. For asset A the result is ambiguous since a higher σ_A^2 reduces overpricing when ε is salient.

Lemma 2.8 *Price volatility $|p_j^\varepsilon - p_j^\mu|$ decreases with K for all assets $j \in \{A, B, C\}$.*

Price volatility $|p_j^\varepsilon - p_j^\mu|$ increases with ω , for all $\omega > 1$ and for all assets $j \in \{A, B, C\}$.

Proof. The proof directly follows from (2.8). The denominator of (2.8) increases in K , implying that the price volatility is decreasing for asset A.

The numerator of (2.8) is increasing in ω , whereas given $\sigma_A^2 + \sigma_B^2 + \sigma_C^2 = (\alpha + \beta + \theta)^2 \sigma_\mu^2$, the denominator is decreasing in ω , thus $\frac{d|p_A^\varepsilon - p_A^\mu|}{d\omega} > 0$ for all $\omega > 1$.

The proof is analogous for assets B and C. ■

The intuition of this lemma is quite obvious. As K , the expected dividend, becomes larger, the difference between systemic and non-systemic assets becomes less significant. Thus, the deviation of price between two different salient cases becomes smaller. The second part of the lemma is obvious considering that ω represents the over/under weighing parameter. As it becomes bigger, the mispricing becomes more significant.

2.5 Discussion and Policy Implications

Price volatility can be a detrimental element that undermines financial stability. Considering the soundness of the financial markets, price stability should have a high priority in policy making. As the last financial crisis has manifested, the costs of such financial crises can be severe and detrimental to the real economy. Thus, policy makers should try to understand the possible drivers of excess asset price volatility so that policies can be tailored to address the structural problems in the financial system.

The policy implications of this paper can be considered in terms of both ex-ante and ex-post policies. The results suggest that for ex-ante policies, investors should use better risk assessment technologies and regulators should try to develop early warning signals for the build up of an extreme systemic risk within the system. In terms of ex-post interventions, the model and results suggest that there is scope for interventions in financial markets - such as Troubled Asset Relief Programs (TARP). However, note that we do not conduct a welfare analysis so the effectiveness of these interventions should be discussed in regards to specific contexts. Specifically, the result that higher risk aversion creates more volatility in prices can help policy makers to tailor their policy to reverse the cyclical behaviour of investors, especially during a financial downturn. For instance, if the risk appetite of investors changes during times of financial stress, intervention policies (such as TARP) not only help to stabilise prices but may also decrease agents mispricing if these policies are able to revert the risk appetites.

Several assumptions of our model need to be discussed. The structure of the model presented here has an explicit focus on the change of salience in order to illustrate the main point in a simple and straightforward manner. However, the model does not mean to explain why investors initially neglected systemic risk.

2.6 Conclusions and Future Work

Although there is a vast literature on explanations for the last financial crisis in terms of the supply side (i.e. the bank side), the investor side has been overlooked. The main puzzle as to what happened prior to the crisis is the extent to which investors neglected systemic risk. In addressing the question of why complex securities are significantly overpriced, this paper looks at the demand side (i.e., investors) by building a tractable model based on salience theory (Bordalo et al., 2012, 2013). Classifying the risk into two components, idiosyncratic and systemic, and assuming investors over-(under-)weigh the risk that is salient (non-salient), we show that investors will misprice the assets and that this will

generate excess volatility when there is an extreme systemic shock. This work thus complements the existing supply side explanations. From the fact that systemic risk was ignored prior to the financial crisis, the model shows how systemic assets were over-priced. Under an extreme systemic shock, systemic risk becomes salient and investors dump systemic assets. This results in fire-sales and excess volatility. The excess volatility in prices could potentially create externalities for the real sector.

Our theory can be tested in a lab through the conduction of an experimental study. The participants may be asked to participate in an investment decision where they are instructed with the definition of two types of risk. In order to make systemic risk more intuitive to understand, we can instruct the students that there is a system-wide (macro) risk which cannot be diversified by choosing any of the assets. Without entering into a discussion as to whether the participants, as decision makers, will choose the optimal portfolio suggested by the classical asset pricing theory, we can test whether they change their investment decision when faced with either a very small systemic risk or a big systemic risk.

Another possible way of testing our theory is by measuring participants' risk aversion in an experimental setting. Then one can ask the subjects to form portfolios of risky assets and then look at whether the more risk averse participants rebalance their portfolio more aggressively (i.e., change the composition of assets more significantly) after a systemic shock.

Another important dimension worth exploring is the interaction between the "supply side" and the demand. Consider an issuer (bank), who can exploit the bias in variance to extract more rent. The issuer may intentionally design/package the securities in a more complex structure which may affect the misperception parameter ω . In other words, ω increases with the opaqueness/complexity of the security, which may imply a higher profit opportunity for the issuer. However, in considering that a higher misperception will increase price volatility, a larger degree of mispricing may be detrimental to the financial sector and consequently to the economy. This extension is quite interesting in regards to what we observed

in the last global financial crisis. Issuers strategically built more complexity into the securities and investors who ignored systemic risk consequently, bought these opaque securities in huge volumes without really understanding the actual risk they faced.

Chapter 3

Subprime Mortgages and Credit Fuelled Bubbles

Chapter Abstract

We consider a model with two types of households; the poor with no initial endowment and the rich with positive endowment; and two types of assets; properties in a poor area and properties in a rich area. In the model, poor agents need credit to buy an asset whereas the rich can draw from their endowment. We show that credit-fuelled housing bubbles sometimes may improve welfare, making the poorer individuals better-off. More precisely, there exist two types of equilibria in both markets: One is a bubble equilibrium, and the other is an equilibrium where asset prices are stable over time. While the poor always obtain a positive surplus in the bubble equilibrium, this is not necessarily true for the rich. Our results suggest that there may be scope for market interventions aimed at sustaining the value of assets held by credit-constrained agents after the burst of a credit bubble.

3.1 Introduction

The years preceding the financial crisis have witnessed unprecedented access to credit for low income individuals with little or no credit history. This was partly facilitated by government sponsored institutions like Fannie Mae and Freddie Mac, whose main task was to incentivize the banks to extend the credit base by buying securitized loans from commercial banks, especially subprime loans. Later, after the crisis, the US Government has been criticised for this policy, which was considered to be one of the fundamental reasons for the credit fuelled housing bubble (especially the subprime bubble). This paper aims to show that sometimes credit fuelled bubbles - where the price of an asset is above the fundamental - may make poor agents (subprime borrowers) better-off and improve welfare. The argument relies on the idea that asset price bubbles can improve the intergenerational allocation of resources in the presence of financial frictions such as borrowing constraints. In principle, the bubble may make the poor better-off through various channels. One such channel is short-term borrowing. When the house owned by a poor gains value, it can be used as collateral to start a business. For example, consider a scenario where the poor with access to good investment opportunities require a short-term credit. If house prices are high (overvalued), the poor may be able to collateralize the house and borrow, whereas if their houses are significantly underpriced they may not be able to access credit at all. This argument is similar to [Tirole \(1985\)](#). He argues that even if bubbles crowd out total investment, still they can improve the flow and allocation of funds through relaxing the borrowing constraint for the investors with a good investment opportunity, i.e., the bubble can lead to a Pareto improvement.

Another channel is mobility. With the possibility of selling their houses at a high price, poor households can enjoy higher socio-economic and geographic mobility. Consider, for instance, a scenario where a house owner may want to move to a better neighbourhood. If his property is over-valued, it can be sold or collateralized to get credit in order to purchase a property in the new neighbour-

hood. However, mobility is instead hindered, if the assets owned by the poor are significantly underpriced. A branch of the mobility literature has shown that mobility is negatively affected when house prices fall due to negative equity. Being trapped in negative equity due to a significant decrease in house prices is called a “lock-in effect” (Ferreira et al., 2010).

We consider a model with two types of households; the poor (with no initial endowment) and the rich (with some endowment) and two types of assets; a house in a poor area (a poor asset) and a house in a rich area (a rich asset). In the model, poor agents need credit to buy an asset whereas the rich can draw from their endowment. In other words, the poor need a 100% loan-to-value (LTV) mortgage (can be thought of as subprime loans) whereas the rich can provide some downpayment. Thus, the equilibrium price for the poor asset is determined by the availability of credit. We show that there exist two types of equilibria for houses in the poor region. One of them is a bubble equilibrium, which exists if credit growth is sufficiently large, i.e. a credit-fuelled bubble. The other is an equilibrium where the asset price is stable over time but the asset is significantly underpriced. Under the bubble scenario, prices grow fast enough so that the poor who purchased the asset not only can pay back their debt by selling the asset when old, but also enjoy a positive surplus. Under the no-bubble scenario, credit growth is small implying that price growth is not large enough to enable the agents to pay back their debt by selling the asset in the future. Thus, in the no-bubble case, the equilibrium price will be zero, i.e. the asset is significantly underpriced. The implication of this result is that the bubble scenario makes the poor better-off and improves welfare.¹ For the rich asset market, there are also two types of equilibria; a bubble equilibrium and a no bubble equilibrium with a price at the fundamental value. In the rich asset market, the price is bounded below by the fundamental value since the rich can always purchase the asset by paying from their endowment. This, in contrast, is not true for the poor asset,

¹However, there is a risk that if the bubble bursts, there may be widespread defaults among the poor. We will comment on this issue in the extension section.

which is always underpriced in the absence of bubbles.

In order to analyse welfare, we consider an extension where both poor and rich assets depreciate through time. We consider a supplier who can build new houses, whenever the price of the house is bigger than the cost. In that case, when the poor asset is underpriced (zero price), the supplier will not find it profitable to build new houses in the poor neighbourhood. Thus, over time, the total stock of poor assets will decrease due to depreciation. In contrast, if there is a bubble in the poor market, the supplier will build new houses, keeping the total stock of assets stable. Comparing these two scenarios, the bubble brings a welfare improvement by stabilizing the total stock of the poor assets.

One of the important empirical implications of the model is that a bubble grows faster in the poor market than a bubble in the rich market. Moreover, if at any point of time the 100% LTV policy is abandoned, there will be a mass default among the poor. The intuition of both results is due to poor agents having zero initial wealth and their dependence on credit.² Figures 3.1 and 3.2 show the Case-Shiller Home prices indexes for both High- and Low-tier classes since 1994 for San Francisco and Miami respectively.³ As can be easily seen from these graphs, the bubbles in the low-tier home markets were much more pronounced compared to the high-tier.⁴ These figures provide anecdotal evidence to motivate our result that housing bubbles may be more pronounced for assets owned by the poor due to poor home buyers' greater dependence on credit growth. Consequently, the downturn also affected the low-tier housing market more than the high-tier, i.e. the price volatility in the poor housing market is higher. In terms of policy, our results suggest that there may be scope for market interventions aimed at sustaining the value of the assets held by the credit-constrained agents

²As will be discussed in the conclusion, the zero wealth assumption is just to make the main point clearer to illustrate. Otherwise, one can consider a model where the poor have some wealth but significantly a smaller amount than the wealth of the rich. That model will also bring similar results qualitatively as long as the poor and the rich differ from each other significant enough.

³See Figures A.9, A.10 and A.11 in Appendix for Los-Angeles, New York and Tampa.

⁴For San Francisco and Miami, the difference between the low and high classes seems to be much more significant.

after the burst of a credit bubble.

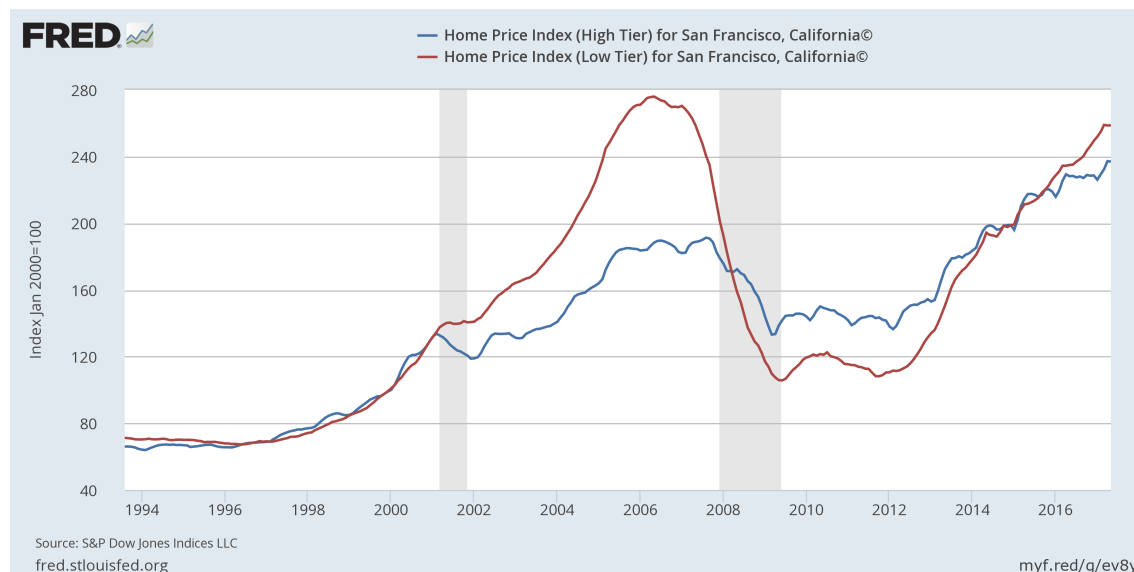


Figure 3.1: Case-Shiller Home Price Index (High vs Low Tier) for San Francisco, California

3.1.1 Related Literature

Bubbles have been studied extensively in the literature and have attracted the attention of economists, both academics and the policy makers, due to their consequences on the allocation of resources. This section will give an overview of the vast literature without presenting a detailed discussion. Even though there is no consensus among economists on the definition of the term ‘bubble’, still one can define a bubble as the sustained mispricing of an asset. Not every mispricing can be considered as a bubble though. The term bubble refers to a period where investors believe that the price growth will continue so that they hold the asset at the ongoing price - even though it seems to be overvalued - since they believe that the asset can be sold at a higher price in the future.⁵

In the literature there are various theoretical explanations for bubbles.⁶ One strand of models is rational expectations models where agents have identical

⁵This type of explanation for the bubbles is termed as the “greater-fool theory of bubbles” (Barlevy, 2015).

⁶See Brunnermeier (2008) and Brunnermeier and Oehmke (2013) for an overview.

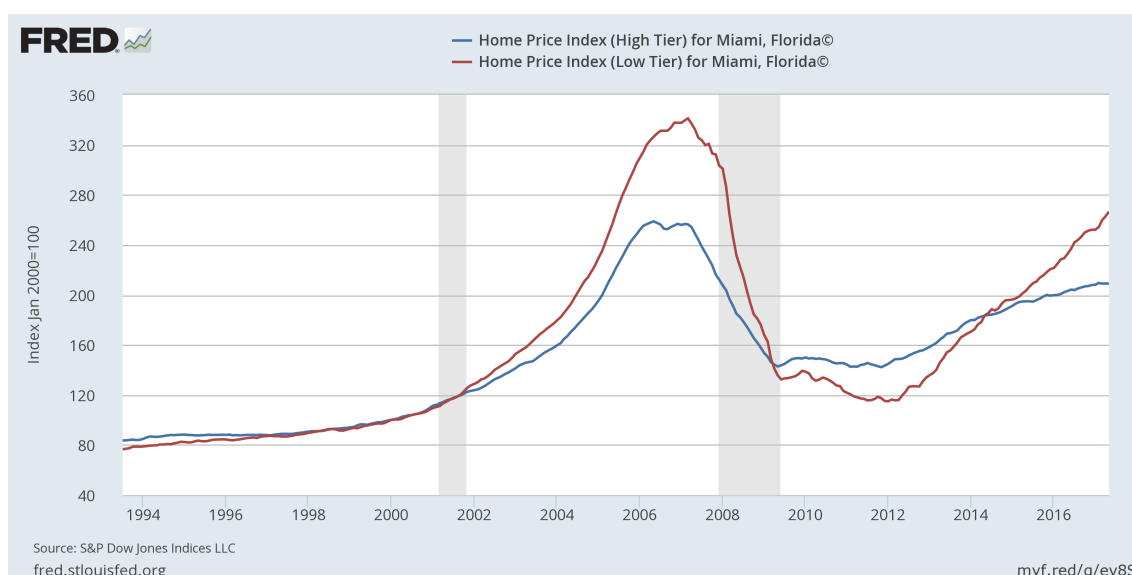


Figure 3.2: Case-Shiller Home Price Index (High vs Low Tier) for Miami, Florida

information (Martin and Ventura, 2012; Galí, 2014). Considering the possibility of speculation when traders are assumed to have rational expectations, Tirole (1982) derives the conditions under which bubbles can be ruled out. He shows that at least one of the following four conditions must be violated in order to sustain a bubble (Barlevy, 2015): The number of potential traders is finite. All traders are assumed to be rational, which is common knowledge. Traders should hold common prior beliefs about the environment. And lastly, resources are allocated efficiently ex-ante, before the trade. Rational bubbles can exist under restrictive theoretical conditions. Blanchard (1979) shows that it is consistent to have bubbles followed by market crashes under rational expectations. He also claims that it is detecting these bubbles seems to be quite hard. Diba and Grossman (1988) show that due to free disposal negative rational bubbles are ruled out. They also show that a positive bubble can only start at the first trading day of a stock and a burst rational bubble cannot restart again. This study complements the literature in a sense that it does not focus on the conditions under which a rational bubble exists but it focuses on the welfare implications.

Another strand of the literature considers asymmetric information bubbles (or heterogeneous belief models) where agents have different information, but still

based on a common prior distribution ([Brunnermeier, 2008](#); [Allen et al., 1993](#); [Conlon, 2004, 2015](#); [Harrison and Kreps, 1978](#); [Scheinkman and Xiong, 2003](#)). In these types of models, prices reflect information but, contrary to the symmetric information case, the existence of a bubble may not be common knowledge.⁷ Another strand of the literature considers the interaction between well-informed sophisticated investors and behavioral investors, who have psychological biases ([De Long et al., 1990](#); [Shleifer and Vishny, 1997](#); [Abreu and Brunnermeier, 2002, 2003](#)). In their seminal paper [Shleifer and Vishny \(1997\)](#) show that there are limits to arbitrage. Even though sophisticated investors understand the overpricing, they may not be able to trade against the bubble. This “limits to arbitrage” argument implies that bubbles can exist since arbitrageurs cannot drive the prices to the fundamental.

This study is more related to the class of “credit bubble” models such as [Allen and Gale \(2000\)](#). They present a model built on a risk shifting argument. Investors, having limited liability, borrow from the banks and bid up the asset prices. When the value of their investment turns out to be low, they simply default and walk away. [Barlevy \(2014\)](#) develops a credit-driven bubble model in order to investigate the possible empirical patterns that can be used as indicators of bubbles. He suggests that rapid price appreciation together with high turnover rates and speculative trading are more likely to take place when assets are overvalued. This study complements the aforementioned literature focusing on the policy and welfare implications of housing bubbles, rather than the destructive consequences of bubbles or the explanations of why bubbles form. Our results can explain why the value of assets owned by the poor may depend more on credit growth. More importantly, we offer a channel through which bubbles may improve the welfare of the poor.

⁷As [Brunnermeier \(2008\)](#) mentions, it can be the case that all agents are aware of the overvaluation of an asset but not everybody knows that all other investors also know of this fact. “This lack of higher-order mutual knowledge” allows for the possibility of the existence of finite bubbles. See also [Allen et al. \(1993\)](#).

3.2 The Model

We consider an OLG model where each generation lives for two periods. Agents are heterogeneous in their initial endowment. There are two types of agents; the rich with an initial endowment $A > 0$ and the poor with no initial endowment. In the economy, there are also two types of assets (a durable consumption good like a residential property) whose consumption generate a positive utility. More precisely, agent $i \in \{R, P\}$ ($R = Rich, P = Poor$) derives a utility $u_i(a_j)$ from the consumption of asset $a_j \in \{a_R, a_P\}$ when young. There is a continuous mass m of poor agents and n of rich agents. Type a_R asset represents a residential property in a rich area whereas a_P represents a residential property in a poor area. Consider a segregated city where the rich live in one neighbourhood and the poor in the other. We assume that the rich have no intrinsic utility from the consumption of a property in the poor area. This is just for simplification and, as long as the rich value the poor asset less than the poor do, our results would follow. Formally, we have the following assumption.

Assumption 3.1 $u_R(a_R) > u_P(a_R) > u_P(a_P) > u_R(a_P) = 0$

A.3.1 basically states that rich agents value the rich asset more than the poor agents and rich agents have no value for the poor asset.⁸ In contrast, poor agents derive some utility from the poor asset, albeit lower than the utility they derive from the rich asset.

The assets are in fixed supply. There is a continuum S^P of poor assets and S^R of rich assets. In any period t , the old generation retires and those who hold an asset can sell it to a member of the young generation at some price p_t^j , $j = R, P$.

Each agent, either poor or rich, cannot purchase more than one house, either as consumers or investors. Moreover, the agents who cannot buy a house obtains zero utility.

⁸The assumption, $u_R(a_P) = 0$, is just for simplification. It does not affect our results qualitatively as long as the above utility ranking holds.

Assumption 3.2 $m > S^P > n > S^R$

A.3.2 will help us to identify the equilibrium prices. Specifically, there are more poor(rich) agents than the supply of poor(rich) asset so that agents compete to buy the assets. The assumption that the supply of poor assets is larger than the mass of rich agents ensures that rich agents will not drive the prices in the poor asset market, even if they may be interested in buying a poor asset for speculative reasons. This will be discussed during the characterization of the equilibrium in the next section.

To simplify the illustration, assume all agents have a discount factor $\frac{1}{r} < 1$ per period.⁹

3.3 Characterization of Equilibria

We define the fundamental value of asset a_j , $j \in \{R, P\}$ as

$$F^j \equiv \frac{u_i(a_j)}{r - 1}. \quad (3.1)$$

As will become clear, in equilibrium the value of asset a_j , is entirely determined by the demand of type j individuals. We thus simplify the notation by setting $u_P(a_P) \equiv u_P$ and $u_R(a_R) \equiv u_R$.

Consider an equilibrium where banks offer contracts with no initial down payment (100% loan to value ratio), type i young agents borrow B_t^j from the bank to buy type j asset from the old agents, and the banks demand a repayment rB_t^j , $r \geq 1$. Type $i = R, P$ agent will buy a property when the following holds;

$$\frac{u_i(a_j) + p_{t+1}^j}{r} \geq p_t^j. \quad (3.2)$$

⁹ Assume that banks borrow from depositors whose discount rate is β and there are enough depositors so that the funding cost of banks will be determined by the depositor's discount factor. Considering that the banking sector is competitive, banks will charge $r = \frac{1}{\beta}$. This will imply that the opportunity cost of borrowers in the economy is $\frac{1}{r}$.

Consider the following environment:

- Banks are competitive, but assume that for each borrower there is an upper bound for the credit available \bar{B}_t . This constraint is relaxed at a rate $r_2 > 1$, i.e. $\bar{B}_{t+1} = r_2 \bar{B}_t$. This borrowing constraint may reflect the total amount of credit available in the economy¹⁰ and will be crucial for the determination of equilibrium. Whenever the borrower is expected to default, we also assume that banks do not lend, so that $\bar{B}_t = 0$. This implies that if a bank expects that $p_{t+1}^j < rB_t^j$, so that the anticipated price in the next period is not enough to cover the repayment, then $\bar{B}_t = 0$.
- In each period t , since there are more young than old who own an asset¹¹, the young compete to buy the asset.

Given this environment, consider first the poor asset a_P . The poor have no initial wealth so they pay a price such that $p_t^P = B_t^P$. Then (3.2) becomes;

$$\frac{u_P(a_P) + B_{t+1}^P}{r} \geq B_t^P. \quad (3.3)$$

Also the young poor cannot pay an amount greater than the maximum borrowing:

$$p_t^P \leq \bar{B}_t. \quad (3.4)$$

Lemma 3.1 *For the poor asset market, if there exists a borrowing equilibrium with $B_t > 0$, then price grows at a rate r or higher in order to avoid default, i.e. $p_{t+1}^P \geq rp_t^P$. This implies that agents obtain a positive surplus, i.e. (3.2) does not*

¹⁰We can think that this assumption is related to the credit-fuelled bubble which resulted in the last financial crisis. In that sense, due to either government policies (as was the case in the US) or simply due to the banks extending their lending base, there will be some credit growth which is captured by the parameter r_2 . We do not put any restriction on the parameter so that any scenario where credit is squeezed can be also captured with parameter $0 < r_2 < 1$. However, in the equilibrium characterization, we will discuss the cases for $r_2 > 1$ since our focus is shedding light on credit fuelled bubbles.

¹¹One can also think of the basic assumption of the classical OLG model, where population grows each period at a fixed rate. However, this assumption will be redundant in our model since, in every period only some young will get the asset and in the next period when they become old, the new generation will keep competing for the asset considering that the mass of asset is always less than the young population in each period. Thus, we simply assume that the population is fixed each period.

bind.

Proof. In the model there are two constraints for a young poor agent; the borrowing constraint $p_t^P \leq \bar{B}_t$ (3.4) and the willingness-to-buy constraint (WTB_P) (3.2). Due to competition among the young, at any period at least one of these two constraints should bind (otherwise, if both are slack, one can outbid others by paying $p_t^P + \epsilon < \bar{B}_t$ where (3.2) still holds.). The agents do not default in an equilibrium since they cannot access credit if they are not expected to pay back their debt, i.e. $\bar{B}_t = 0$ if banks expect that $p_{t+1}^j < rB_t^j$. Thus, no-default requires $p_{t+1}^P \geq rB_t = rp_t$. This implies that agents can only borrow if the price of the poor asset grows at a rate r or higher. In turn, this implies that agents will obtain a positive surplus since WTB_P does not bind. To see this point, note that a binding WTB_P would imply that agents cannot pay back their debt:

$$\frac{u_P(a_P) + p_{t+1}^P}{r} = p_t^P \Rightarrow p_{t+1}^P = rp_t^P - u_P(a_P) < rp_t^P. \quad (3.5)$$

(3.5) then implies that in any equilibrium with positive borrowing WTB_P does not bind. ■

Lemma 3.2 *Suppose A.3.1 and A.3.2 hold. Then, in the poor asset market;*

(i) *If $r_2 \geq r$, there exists an equilibrium with prices $p_t^P = \bar{B}_t$ and $p_{t+1}^P = \bar{B}_{t+1} = r_2\bar{B}_t$ in any consecutive periods. Poor agents who hold an asset obtain a positive surplus, (3.2) holds with strict inequality and there are no defaults in equilibrium.*

(ii) *If $r_2 < r$, there is a unique equilibrium with price $p_t^P = 0$ in every period.*

Comparing two equilibria (i) and (ii), the poor are better off under (i).

Proof.

Lemma 3.1 shows that in equilibrium, (3.4) must bind due to competition among the young (since WTB_P is slack). Thus, in case (i), the young must borrow up to the maximum in equilibrium. Then, equilibrium prices are $p_t^P = \bar{B}_t$ and $p_{t+1}^P = \bar{B}_{t+1} = r_2\bar{B}_t$. Since $r_2 \geq r$, we have $p_{t+1}^P = r_2\bar{B}_t \geq r\bar{B}_t$ implying that

agents can pay back their debt and there are no defaults in equilibrium.

A slack in (3.2) implies that the young who hold an asset obtain a positive surplus;

$$\frac{u_P}{r} + \frac{p_{t+1}^P}{r} - p_t^P = \frac{u_P}{r} + \frac{(r_2 - r)\bar{B}_t}{r} > 0. \quad (3.6)$$

Note that there may also exist no bubble equilibria with prices $p_t^P = 0, \forall t$, if agents expect $p_{t+1}^P = B_{t+1} < rB_t = rp_t^P$ at any point in time. The proof is similar to the one provided below.

We now check whether the rich have an incentive to buy the poor asset when $r_2 \geq r$. Even though they derive no utility, if $r_2 > r$ the rich may also buy the poor asset to speculate on the price growth. But Assumption 3.2 indicates that they cannot drive the prices since the mass of the poor asset is bigger than the mass of rich agents. Thus, the marginal buyer is a poor and the equilibrium is as established above.

For case (ii); again as proved above, if there exists a borrowing equilibrium $B_t^P > 0$, it has to be $B_t^P = \bar{B}_t = p_t^P$. However, considering that $r_2 < r$, $p_{t+1}^P = r_2\bar{B}_t < r\bar{B}_t$ implying that the old cannot pay back their debt and default. But, then expecting a default, banks will never give credit, i.e. $\bar{B}_t = 0$. Thus, the only equilibrium involves $p_t^P = 0, \forall t$.

Now suppose the rich buy the poor asset. Again since $S^P > n$, the marginal buyer is a poor agent and thus the previous argument applies.¹²

Comparing the two cases of (i) and (ii), in (i) the surplus from purchasing an asset is bigger than that of case (ii). The reason for this is that, when there is a bubble, agents enjoy an excess surplus, on top of the utility they get from the consumption, considering that price growth is bigger than the repayment to the bank. Formally, the surplus under case (i) is $\frac{u_P}{r} + \frac{(r_2 - r)\bar{B}_t}{r}$, where the gain from the price growth is positive $\frac{(r_2 - r)\bar{B}_t}{r} > 0$ for $r_2 > r$. Whereas the surplus

¹²Note that, when $r_2 < r$, the rich do not have speculative incentives to buy the poor asset since $u_R(a_P) = 0$ whereas the poor derive a positive utility even though the price is zero. Thus, even without the assumption $S^P > n$, the marginal buyer is a poor agent.

under case (ii) is $\frac{u_P}{r}$ since $p_t^P = p_{t+1}^P = 0$. Then, comparing the surpluses we get;

$$\frac{u_P}{r} + \frac{(r_2 - r)\bar{B}_t}{r} > \frac{u_P}{r}.$$

■

Note that here not all young obtain an asset. Thus it should be the case that there will be some credit rationing.

Figures 3.3 and 3.4 provide a graphical illustration of the arguments. Figure 3.3 represents the constraints for the poor when $r_2 \geq r$. The red lines represent the two main constraints; WTB_P (3.2) and the borrowing constraint (3.4). Rearranging (3.2) we get $p_{t+1} \geq rp_t - u_P$.¹³ The area above this red line represents the willingness-to-buy constraint for the poor (WTB_P). The area on the left-side of \bar{B}_t represents the borrowing(resource) constraint. Note that, as proved in Lemma 3.1, in equilibrium WTB_P (3.2) does not bind considering the no-default condition. Thus, the equilibrium must lie in the area above the blue-dotted line rp_t , which represents the no default condition. Due to the competition among the young, an equilibrium has to be on the dashed-yellow/red line, i.e. borrowing constraint binds, $p_t^P = \bar{B}_t$ (since WTB_P is slack). But this means for the next period, the borrowing constraint also binds as well $p_{t+1}^P = \bar{B}_{t+1}$. Then, the equilibrium is at point E^* .

Note that the (0,0) point at Figure 3.3 can be also sustained as a no bubble equilibrium as long as agents expect the next generation not to borrow (i.e. if agents do not expect any positive price in the next period) and the price will stay at zero.

¹³Note that for the poor, the price equals to the borrowing $p_t = B_t$ and $p_{t+1} = B_{t+1}$. Thus, (3.2) can be also represented in the graph by $B_{t+1} \geq rB_t - u$.

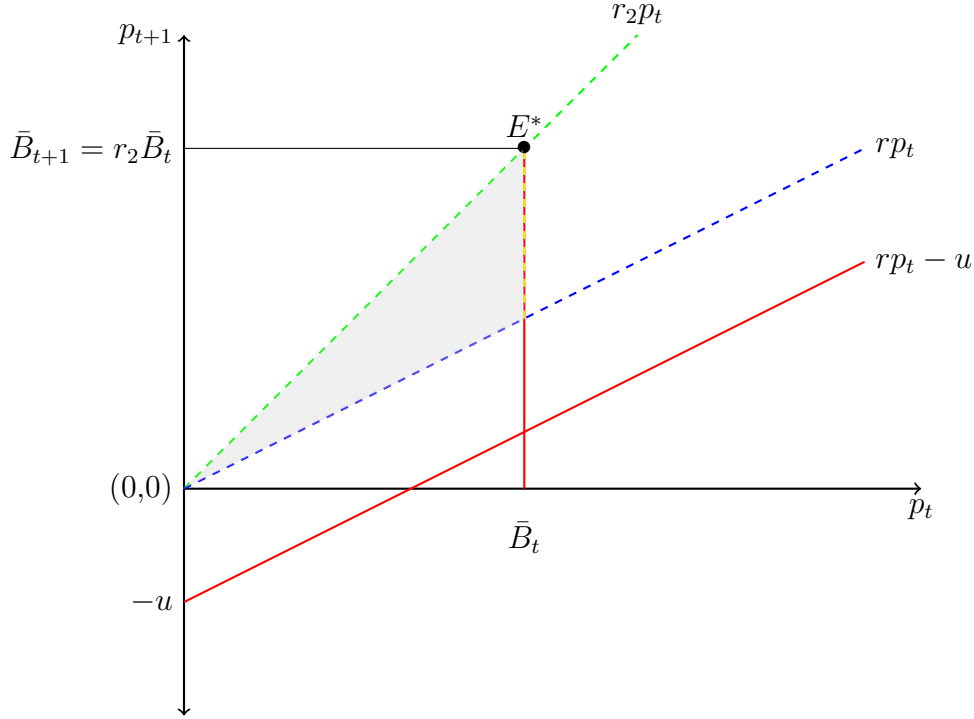


Figure 3.3: Poor asset market equilibrium for $r_2 \geq r$ where $p_t = B_t$ and $p_{t+1} = B_{t+1}$.

Figure 3.4 shows the case when $r_2 < r$. The no-default region is shown by the shaded grey area. And the red lines, the binding constraints (3.2) and (3.4), are in the default region. Thus, the equilibrium is at $p_t = p_{t+1} = 0$. To see this is indeed the unique equilibrium, consider a point C on the shaded region. But C is off the WTB_P , i.e. $p_{t+1} > r p_t - u$, so that borrowers obtain a positive surplus. In turn, competition among the borrowers implies that the borrowing constraint must be binding. Otherwise, borrowers would have chosen to borrow more to secure the asset. However, any $p_t = \bar{B}_t > 0$ is outside the shaded region, so that default would have occur. But then, expecting a default, banks will not supply any credit. As a result, $\bar{B}_t = 0$ and the only equilibrium involves $p_t = p_{t+1} = 0$, $\forall t$.

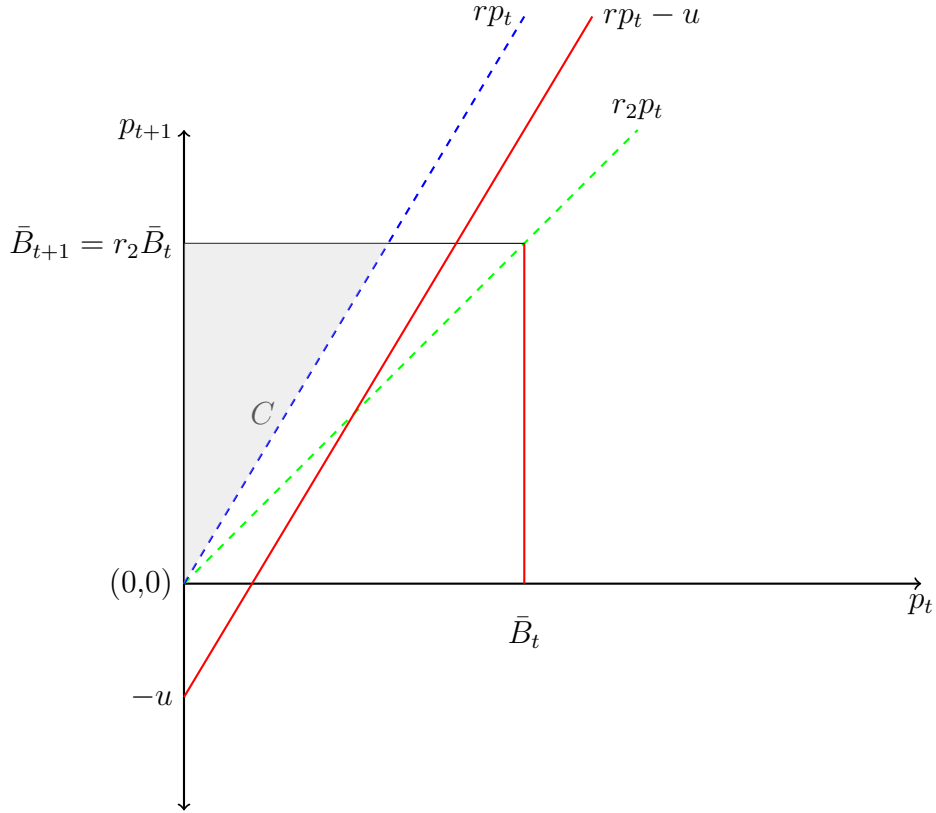


Figure 3.4: Poor asset market equilibrium for $r_2 < r$

As stated above in Lemma 3.2, there are two main cases in the poor asset market:

- i. For $r_2 \geq r$, the young bid the price up to the borrowing limit $p_t^P = \bar{B}_t$ and $p_{t+1}^P = \bar{B}_{t+1} = r_2 \bar{B}_t$, thus implying that price grows at a rate r_2 every period. Under this equilibrium, young poor individuals can pay back their debt since price growth is fast enough. Thus, there is a bubble equilibrium and for the individuals who got the house (3.2) holds with strict inequality. The price fetched by the asset in the second period of the agents' life is used to repay the principal plus the interest rate of the loan borrowed $r \bar{B}_t$. Banks break even and there are no defaults. Under this parametric case, rich agents may also have an incentive to buy the poor asset - even though they do not drive any intrinsic utility from that - since price growth is large

enough for them to speculate on the bubble. As discussed in the above proof, rich agents' demand does not drive the prices since the mass of rich agents is smaller than the mass of poor assets available in the economy. As a result, the marginal buyer is a poor agent.

- ii. For $r_2 < r$, the borrowing constraint is relaxed tighter than the cost of borrowing r , thus implying that even if the poor agent buys the asset by borrowing up to the maximum, still the price growth is not large enough to cover the repayment in the next period; i.e. the price fetched by the asset in the second period of life $r_2 \bar{B}_t$ is smaller than the repayment, $r \bar{B}_t$. This implies that if an agent borrows and purchases a house, the agent will default when old. So, the poor asset is not traded at a positive price and there will be an equilibrium where the asset price is below the fundamental, namely $p_t^P = 0$, in every period. In this scenario, no rich agent will have an incentive to buy the asset.

The above discussion implies that there are two types of equilibria for the poor housing market depending on the credit available in the economy. If there is strong credit growth in the economy (case *i.*), there will be a bubble equilibrium where agents who buy the house obtain positive surplus ($WTB_P(3.2)$ holds with strict inequality). If there is not enough credit growth in the economy (case *ii.*), then in equilibrium the poor asset is significantly underpriced (the equilibrium price is below the fundamental).

- **Now, consider the rich asset, a_R :**

Different from the poor, rich agents have an endowment A , which can be used to buy a house.

The rich again have two constraints; the willingness-to-buy constraint (WTB_R)

$$\frac{u_R(a_R) + p_{t+1}^R}{r} \geq p_t^R, \quad (3.7)$$

and the resource constraint;

$$p_t^R \leq \bar{B}_t + A. \quad (3.8)$$

The rich can pay a downpayment D_t and borrow B_t to buy the house, so that the price of the rich asset is $p_t^R = B_t + D_t$. Below, we will focus on the case where the downpayment is the same in each period, $D_t = D_{t+1} = D$ ¹⁴, so that the price is;

$$p_t^R = B_t + D. \quad (3.9)$$

Then, the willingness-to-buy constraint (WTB_R) (3.7) can be written as;

$$\frac{u_R(a_R) + B_{t+1} + D}{r} \geq B_t + D \quad (3.10)$$

and when the maximum is borrowed $\bar{B}_{t+1} = r_2 \bar{B}_t$ we obtain¹⁵

$$D \leq \frac{u_R}{r-1} + \bar{B}_t \frac{r_2 - r}{r-1} = F^R + \bar{B}_t \frac{r_2 - r}{r-1}. \quad (3.11)$$

Firstly, we need to check whether there can be an equilibrium where rich agents do not borrow, $B_t = 0$ and just pay from their initial endowment so that there is no bubble, $p_t^R = p_{t+1} = D$.

Lemma 3.3 *If $A \geq F^R = \frac{u_R}{r-1}$, then there exists an equilibrium without borrowing, $p_t^{R*} = p_{t+1}^{R*} = F^R$ and $B_t^* = 0$. If $A < F^R$, then in equilibrium it must be $B_t^* > 0$ (whenever $\bar{B}_t > 0$).*

Proof. In any equilibrium, at least one of the constraints (3.7) or (3.8) must bind due to the competition among the young (as buyers). Now, consider an equilibrium where the young do not borrow and pay a downpayment D so that in each period the price equals to the downpayment, $p_t^{R*} = p_{t+1}^{R*} = D$. However,

¹⁴Note that there can be other equilibria where $D_t \neq D$ and it grows each period. We restrict attention to a case where under a bubble the rich competes for the asset and after some point due to increase in price they have to pay their all endowment. Thus, even though at the initial periods $D_t < D$ after some period the young rich need to pay D to ensure to buy the asset.

¹⁵Note that we denote $u_R = u_R(a_R)$ for the rest of the paper.

this means that the resource constraint (3.8) is slack (since agents do not borrow). Thus, (3.7) must bind, which implies

$$\frac{u_R(a_R) + D}{r} = D \Rightarrow D = \frac{u_R(a_R)}{r - 1} = F^R.$$

The second part of the lemma follows from the fact that if $A < F^R$, then the WTB_R does not bind so that the resource constraint must bind. ■

Lemma 3.3 basically states that, if the endowment A is big enough, there exists an equilibrium where the rich asset is traded at the fundamental value without any bubble (and the price is stable). If the initial endowment is small then, similar to the poor market, the young borrow to buy the asset, i.e. $B_t^* > 0$ provided that $\bar{B}_t > 0$.¹⁶

Consider now an equilibrium with a positive amount of borrowing. The rich have resources A , which can be used to buy the asset. They already paid $D \leq A$ as a downpayment when they were young. Then, when they are old, the no default condition is $(A - D) + p_{t+1}^R \geq rB_t$. Since $p_{t+1}^R = B_{t+1} + D$, rearranging the no default condition we get

$$B_{t+1} \geq rB_t - A. \quad (3.12)$$

Rearranging (3.10)(WTB_R) we get

$$B_{t+1} \geq rB_t + D(r - 1) - u_R. \quad (3.13)$$

- Below, Figures 3.5 and 3.6 show these two constraints of (3.12) and (3.13) for the rich when $r_2 \geq r$ and $r_2 < r$ respectively. In both graphs, the dotted blue lines represent the no default conditions and the grey-shaded area presents the pairs (B_t, B_{t+1}) where no default occurs. The red and red dotted lines represent the WTB_R for $D > F^R$ and $D = F^R$ respectively. Note that (3.13) (WTB_R)

¹⁶Note that still there may be a no-bubble equilibrium even if $A < F^R$. This will be proved in Lemma 3.4.

partially lies within the shaded area for the case $D \geq F^R$ (red line).

In Figure, 3.5 the maximum borrowing (point E^*) is an equilibrium, though not unique, where the resource constraint binds and the agents obtain a positive surplus. Note that there can be other equilibria (which will be discussed below) such as a point on the WTB_R when $D = F^R$ and $B_{t+1} = rB_t$ so that agents obtain zero surplus (since WTB_R binds).

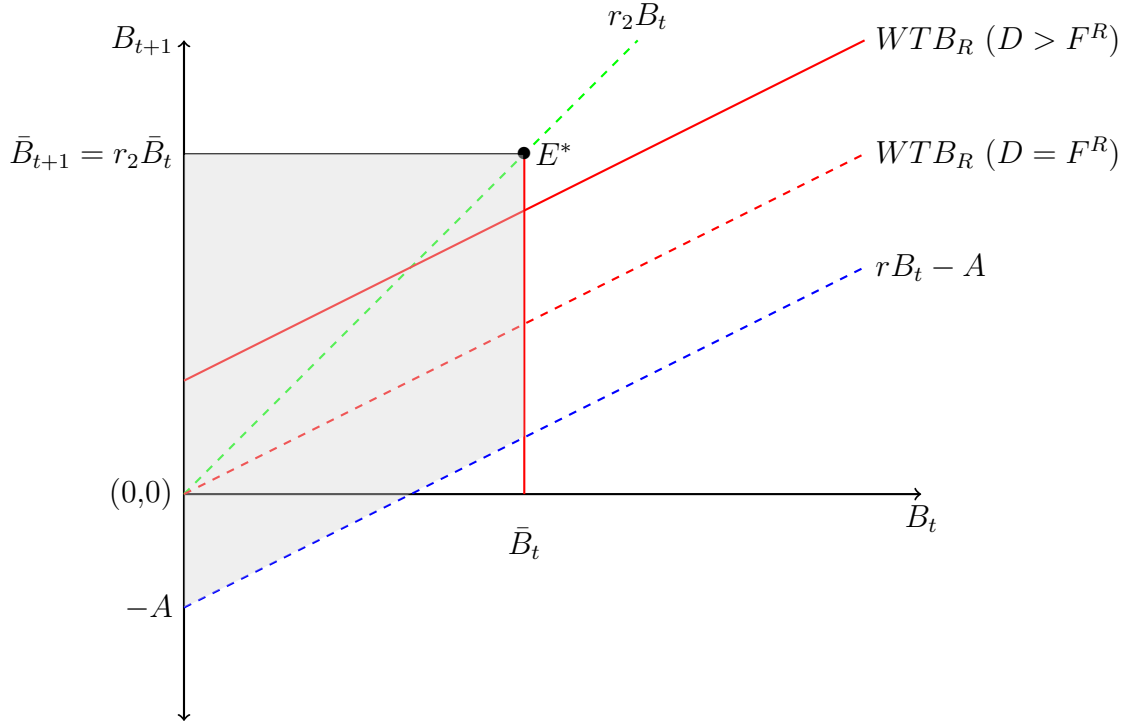


Figure 3.5: Constraints for the rich when $r_2 \geq r$

When the wealth is smaller than a specific value $\underline{A} = \frac{u}{r} < F^R$, then the downpayment has to be small as well since $D \leq A$, so that the red line WTB_R will lie below the default line (the blue-dotted line). This case is similar to the poor asset market where agents are wealth constrained. Thus, in order to have a meaningful difference between the poor and the rich, we assume that A is large enough. Lemma 3.9 in Appendix considers the case when this assumption does not hold.¹⁷

¹⁷Basically, as shown in the Appendix, if the initial wealth of the rich is smaller than this

Assumption 3.3 $A > \underline{A}$ where $\underline{A} = \frac{u_R}{r} < F^R$.

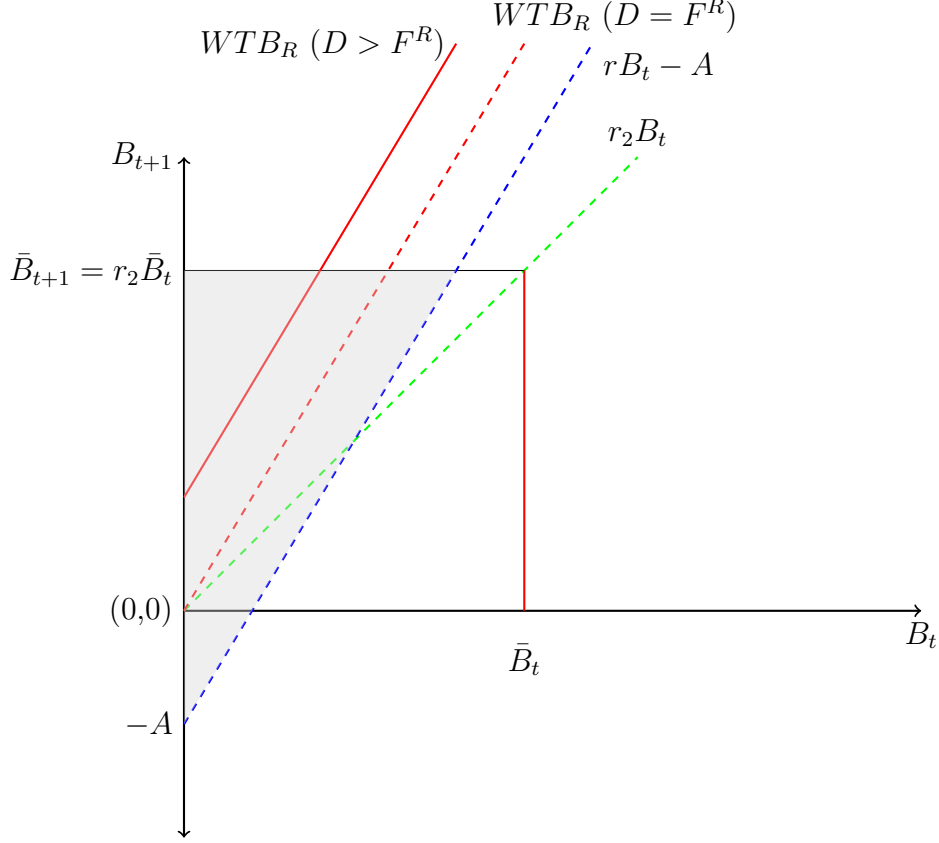


Figure 3.6: Constraints for the rich when $r_2 < r$

Figure 3.6 represents the constraints that the rich face when the credit growth rate is smaller than the borrowing rate, i.e. $r_2 < r$. Considering a candidate bubble equilibrium, there are two possible cases, one where WTB_R binds and the other where the no-default (blue dotted line) binds (when A.3.3 is violated). However, the borrowing level in both of these cases (where the equilibrium is on these lines and moving up along the lines) cannot be sustainable since the credit growth rate r_2 is smaller than the slope of these lines r . Thus, there can only exist steady state equilibria (with no bubble) where $B_t = B_{t+1} = B^*$ and $D_t = D_{t+1} = D^*$.

cut-off \underline{A} , as is similar to the poor market, the rich asset will be traded below fundamental, albeit at a positive price.

Note that there may be other equilibria where B_t and/or D_t grow at rates that are not constant. But we restrict attention to balanced growth paths.

Let \bar{B}_0 be the initial borrowing limit in period 0.

Lemma 3.4 *Suppose A.3.1, A.3.2 and A.3.3 hold. Then, in the rich asset market,*

- (i) *For the case $r_2 \geq r$;*
 - (a) *if $A < F^R = \frac{u_R}{r-1}$ or*
 - (b) *if $A \geq F^R$, $r_2 > r$ and \bar{B}_0 is sufficiently large,*
then there exist balanced bubble equilibria such that the maximum amount
is borrowed $B_t = \bar{B}_t, \forall t$, and rich agents who hold an asset obtain a
positive surplus ($WTB_R(3.13)$ is slack). The price is $p_t^{R} = \bar{B}_t + A$,*
implying that the resource constraint (3.8) binds .
 - (c) *if $A \geq F^R$, there also exist bubble equilibria where the rich agents*
obtain no surplus in all periods, i.e. $WTB_R(3.13)$ binds.
 - (d) *also in all cases, there exist equilibria with no bubble where $p_t^{R*} =$*
 $F^R = \frac{u_R}{r-1}, \forall t$.
- (ii) *If $r_2 < r$, there is no balanced growth path involving a bubble; i.e. there*
exists a continuum of equilibria without a bubble (a steady state) where the
asset is traded at the fundamental price $p_t^{R} = F^R, \forall t$ such that:*
 - (a) *If $A \geq F^R$, then (3.11) (WTB_R) binds and $p_t^{R*} = p^* = B^* + D^* = F^R$*
for $D^ \in [0, F^R]$ and $B^* \in [0, F^R]$.*
 - (b) *If $\underline{A} < A < F^R$, then (3.11) (WTB_R) binds and $p_t^{R*} = p^* = B^* + D^* =$*
 F^R for $D^ \in [F^R - \frac{A}{r-1}, A]$ and $B^* \in [F^R - A, \frac{A}{r-1}]$ where $\underline{A} = \frac{u}{r}$.*

Proof.

We can rewrite the constraints for the rich.

$$B_t \leq \bar{B}_t.$$

$$D_t \leq A.$$

$$WTB_R: \quad \frac{u_R + B_{t+1} + D_{t+1}}{r} \geq B_t + D_t.$$

If WTB_R holds with inequality, then competition implies $B_t = \bar{B}_t$ and $D_t = A$ so that $p_t^R = \bar{B}_t + A$, i.e. the resource constraint binds.

Then we can rewrite WTB_R as follows: $\frac{u_R + r_2 \bar{B}_t + A}{r} > \bar{B}_t + A \Rightarrow$

$$u_R > \bar{B}_t(r - r_2) + A(r - 1). \quad (3.14)$$

However, we know that $\bar{B}_{t+1} = r_2 \bar{B}_t$. Taking the limit for $t \rightarrow \infty$ shows that (3.14) can only hold for all t if $r_2 \geq r$. If also $A < F^R = \frac{u_R}{r - 1}$, then the resource constraint always binds in every bubble equilibrium (and the rich who hold an asset experience positive surplus).

Consider then the case where $A \geq \frac{u_R}{r - 1}$. We need to look at the initial conditions. If $r_2 > r$ and \bar{B}_0 is large enough such that WTB_R is slack (i.e., (3.14) holds at time zero), then we are back to the case above.

If \bar{B}_0 is small or $r_2 = r$ so that

$$u_R < \bar{B}_0(r - r_2) + A(r - 1)$$

then either $B_0 < \bar{B}_0$ or $D_0 < A$ or both.

Whenever $A \geq \frac{u_R}{r - 1}$, bubble equilibria where WTB_R binds are also possible.

We now construct a bubble equilibrium where WTB_R binds.

Suppose that $D_t = D$ is constant and B_t grows as follows

$$B_{t+1} = rB_t + (r - 1)D - u_R \quad (3.15)$$

then the rich make zero surplus in all periods. Consider a special case where $D = \frac{u_R}{r - 1}$. Then B_t grows at a rate $r < r_2$ in every period so that the borrowing constraint (and the resource constraint as well) never binds.

Note however that there may be other equilibria where B_t and/or D_t grow

at rates that are not constant and satisfy (3.15). But we restrict attention to balanced growth paths.

For the case (ii), when $r_2 < r$:

Note that due to the resource constraint,¹⁸ the price ratio $\frac{p_{t+1}^R}{p_t^R}$ is bounded above by r_2 in the limit for $t \rightarrow \infty$. Since $r_2 < r$,

$$\lim_{t \rightarrow \infty} \frac{p_{t+1}^R}{p_t^R} \leq r_2 < r.$$

$WTB_R(3.7)$ implies $\frac{p_{t+1}^R}{p_t^R} \geq r - \frac{u_r}{p_t^R}$ so that even though we have binding WTB_R ,

$$\lim_{t \rightarrow \infty} \frac{p_{t+1}^R}{p_t^R} \geq r > r_2.$$

Since we obtain a contradiction, this means that for $r_2 < r$ there cannot exist a balanced bubble equilibrium.

Thus, consider an equilibrium where the price is constant, $p_t^R = p^{R*}$. Then, from WTB_R , it must be $p^{R*} \leq \frac{u_R}{r-1}$. However, due to competition, $p^{R*} < \frac{u_R}{r-1}$ cannot be an equilibrium, since otherwise a young agent can pay $p^{R*} + \epsilon$ and obtain the asset for sure. Thus, the equilibrium must be such that WTB_R binds (agents obtain no surplus). The price is thus equal to the fundamental value in all periods.

$$p^{R*} = \frac{u_R}{r-1} = F^R.$$

The only difference between (ii)(a) and (ii)(b) is the maximum borrowing considering the default cut-off. Formally, at steady state, no default becomes $B^* \geq rB^* - A$,

$$B^* \leq \frac{A}{r-1}. \quad (3.16)$$

Note that agents have a continuum of choices between borrowing and down-payment. This is due to rich agents having two choice variables and the cost of

¹⁸Recall that the resource constraint is $p_t \leq \bar{B}_t + A$ and $p_{t+1} \leq \bar{B}_{t+1} + A = r_2 \bar{B}_t + A$

borrowing being the same as the opportunity cost of consumption in the model. In that sense, agents are indifferent between borrowing less and consuming less today but repaying a lower amount tomorrow vs. borrowing more and consuming more today and repaying a higher amount. Since we are interested in the equilibrium price, this does not affect our argument.¹⁹ ■

Considering the different possible cases for r_2 , the intuition of the Lemma 3.4 is as follows. Whenever the credit growth rate is larger than the cost of borrowing, there exists a bubble-equilibrium where prices grow steadily and the agents who hold an asset enjoy a positive surplus. This is due to the fact that the price fetched in the second period of life (when agents are old) is higher than the repayment to the bank.

Whenever the credit growth rate is smaller than the borrowing rate, $r_2 < r$, the borrowing is not sustainable. Since the price growth would be smaller than the borrowing rate,²⁰ a bubble could not be sustained since agents would not be able to pay back their debt. Thus, the only equilibrium is a steady state where the rich asset is traded at the fundamental price.

It is also worth noting that when the initial credit limit is small and the endowment A is large, there exist bubble equilibria where the rich experience zero surplus.

Lemma 3.5 *As a special case, when $r_2 = r$, in the rich asset market any borrowing level $0 \leq B_t \leq \bar{B}_t$ with prices $p_t^{R*} = B_t + D^*$ and $p_{t+1}^{R*} = B_{t+1} + D^*$ is an*

¹⁹In order to understand this point clearly and formally. Consider an equilibrium for case (ii)(a). For the given equilibrium price $p_t^R = B_t + \hat{D} = F^R$ consider two strategies where a young pays $0 \leq \hat{D} < F^R$ at time t and borrows, $B_t = F^R - \hat{D}$ vs. a young pays $D = F^R$ from the endowment and borrows nothing $B_t = 0$ so that still pays same price $p_t^R = F^R$. Both of these strategies are equivalent in terms of the utility generated. Since (3.8) does not bind and willingness-to-buy constraint (3.7) binds. When \hat{D} is paid, (3.7) can be written as $\hat{D} = \frac{u}{r} + \frac{p_{t+1} - r\bar{B}_t}{r} \Rightarrow \hat{D} + \bar{B}_t = \frac{u}{r} + \frac{p_{t+1}}{r}$ whereas when A is paid and $\bar{B}_t - (A - \hat{D})$ is borrowed (3.7) can be written as $A = \frac{u}{r} + \frac{p_{t+1} - r(\bar{B}_t - (A - \hat{D}))}{r} \Rightarrow \hat{D} + \bar{B}_t = \frac{u}{r} + \frac{p_{t+1}}{r}$. Thus as long as the resource constraint, (3.8) does not bind, there exist a continuum of equilibria, $p_{t+1}^{R*} = \bar{B}_{t+1} + \hat{D}$ where $D \in [\hat{D}, A]$.

²⁰Actually, in the rich asset market, the price growth rate should be smaller than the growth rate of credit.

equilibrium where $B_{t+1} = rB_t$ and

- $D^* = F^R$ for $A \geq F^R$ and WTB_R binds, i.e. the young who hold an asset get no excess surplus.
- $D^* = A$ for $A < F^R$ and WTB_R is slack, i.e. the young who hold an asset get excess surplus.

Proof. The poof is similar to case (i) in Lemma 3.4, the only difference is that WTB_R binds for $A \geq F^R$. ■

Proposition 3.6 Suppose A.3.1, A.3.2 and A.3.3 hold. An equilibrium of this economy consists of price pairs (p_t^{j*}, p_{t+1}^{j*}) and borrowing levels B_t^{j*} for both asset markets $j \in \{R, P\}$ and a downpayment D^* for the rich such that,

- (I) if credit growth is fast, i.e., $r_2 \geq r$;
- (a) in the poor asset market there exists an equilibrium where agents borrow up to the maximum $p_t^P = \bar{B}_t$ and $p_{t+1}^P = \bar{B}_{t+1} = r_2 \bar{B}_t$. Poor agents who hold an asset obtain a positive surplus.
 - (b) in the poor asset market there also exists a steady state with price $p_t^P = 0$ in all periods.
 - (c) in the rich asset market, if $A < F^R = \frac{u_R}{r-1}$ or if $A \geq F^R$, $r_2 > r$ and \bar{B}_0 is sufficiently large, then there exist balanced bubble equilibria such that the maximum amount is borrowed $B_t = \bar{B}_t, \forall t$, and rich agents who hold an asset obtain a positive surplus ($WTB_R(3.13)$ is slack). The price is $p_t^{R*} = \bar{B}_t + A$, implying that the resource constraint (3.8) binds.
If $A \geq F^R$, then there also exist bubble equilibria where the rich agents obtain no surplus in all periods, i.e. $WTB_R(3.13)$ binds.
 - (d) in the rich asset market there also exists a steady state with price $p_t^{R*} = p_{t+1}^{R*} = F^R, \forall t$.
- (II) if credit growth is smaller than the borrowing rate $r_2 < r$, then
- (a) in the poor asset market there exists a unique steady state with price $p_t^{P*} = 0, \forall t$.

(b) in the rich asset market there exists a steady state (no bubble) where the asset is traded at the fundamental: i.e. $p_t^{R*} = p_{t+1}^{R*} = F^R$, $\forall t$ and (3.11) (WTB_R) binds

Proof. See proof of the previous Lemmas 3.2 and 3.4. ■

When $r_2 > r$, in the poor asset market, the only alternative to a bubble equilibrium is an equilibrium where the asset is severely underpriced, i.e. $p_t^P = 0$ in all periods. In contrast, in the rich asset market there exist equilibria where the price is equal to the fundamental. If the young rich do not expect any price growth, they will pay $p_t = F^R$. Thus, expectations are crucial in the equilibrium selection.

An interesting observation following from Proposition 3.6 is that, while the poor always obtain positive surplus in a bubble equilibrium (even though in the steady state), this is not generally true for the rich.

Note also that when the credit growth is fast, the rich agents also have an incentive to purchase the poor asset in order to speculate on the price growth (even if they do not derive any intrinsic utility from the poor asset). However, considering A.3.2, the rich cannot drive the prices in poor asset market. As for the rich asset market, since the rich can rely on their endowment and value the rich asset more than the poor, they can always price out poor agents since the poor cannot provide a downpayment.²¹

Comparing the two types of equilibria (I).(a) and (I).(b), the poor are better off when there is a bubble in the poor asset market.

Proposition 3.7 *For a given credit growth rate $r_2 > r$, in any bubble equilibrium such that the resource constraint binds, the bubble grows faster in the poor market than when compared to the rich market.*

²¹Note that this depends on the nature of the contract for the rich. If the contract requires a downpayment then the poor can never buy a rich house. If the contract does not require any downpayment, then we need to consider a scenario where the price of the poor asset grows faster than that of the rich asset. For example, if the price of the rich asset grows at r_2 (for the case $r_2 > r$) and the price of the rich asset grows at rate r , at some point the poor asset will become more expensive. Then, we need to assume that the price of the poor asset at any period is bounded by the price of the rich asset.

Proof. Compare the two bubbles in the poor and the rich asset markets. For the poor asset, $p_t^P = \bar{B}_t$ and $p_{t+1}^P = \bar{B}_{t+1} = r_2 \bar{B}_t$. This implies that the bubble grows at the rate r_2 . For the rich market $p_t^R = \bar{B}_t + D$ and $p_{t+1}^R = \bar{B}_{t+1} + D = r_2 \bar{B}_t + D$. However, this means the growth rate of the bubble is less than r_2 in the rich market since $D > 0$. ■

The intuition of this result is simple. Since the rich can pay from their endowment, when there is a bubble in the rich market the price grows slower than when compared to the poor market. This result matches the empirical observations given in Figures 3.1 and 3.2.

Several assumptions of the model need to be discussed. First, the zero wealth assumption for the poor just allows us to present the main idea of this paper in a starker manner. A model where the poor also have some initial wealth, albeit significantly smaller than the wealth of the rich would generate similar results.

Second, when $r_2 > r$, i.e. the credit growth rate is larger than the cost of borrowing, there exist ever growing bubbles in both markets. Considering the limited resources, these bubbles cannot be sustained forever, since at some point the bubble will become so big that all the credit of the economy will be allocated to the bubble.

3.4 Extensions

3.4.1 Depreciation

In order to illustrate some possible consequences of underpricing, consider an environment where the stock of houses depreciate over time at rate $\lambda \in (0, 1)$. More precisely, if no new houses are built in period t ,

$$S_{t+1}^j = \lambda S_t^j.$$

Suppose also that there is a producer of houses, who can replenish the stock provided that the market price is above the marginal cost. We assume that the producer is a price taker (i.e. can only sell at market price) and has a constant marginal cost c^j of producing a type j house, $j \in \{P, R\}$. Assume

Assumption 3.4 $0 < c^P < c^R \leq F^R$.

For simplicity, we restrict attention to the case where the total supply of houses cannot exceed an upper bound \bar{S}^j . For instance, this might be the case if there are building restrictions that limit the amount of land available for building. Similar to the previous sections, we assume

$$m > \bar{S}^P > n > \bar{S}^R.$$

Clearly enough, whatever the price at time t , an old agent selling his property will only receive λp_t^j . We compare two types of equilibrium; a bubble where the borrowing limit is binding and an equilibrium with no bubble. In general, the surplus generated by the transactions in market $j \in \{P, R\}$ at time $t + 1$ is

$$W_{t+1} = \underbrace{S_{t+1}^j [u_j + \lambda p_{t+1}^j - r p_t^j]}_{\text{Surplus of asset holders}} + \underbrace{\Delta [p_{t+1}^j - c^j]}_{\text{Producer Surplus}}. \quad (3.17)$$

where $\Delta \geq 0$ denotes production at time $t + 1$. (Note that the total surplus (3.17) does not include the banks' surplus, since credit is modelled in a reduced form. This is however consistent with the competitive case where banks make zero profits).

Restrict attention to the poor asset market and consider a bubble equilibrium where the producer supplies $(1 - \lambda)\bar{S}_t^P$ in every period, so that $S_t^P = \bar{S}_t^P$ for all t . Assume also that λ is not too small so that $\lambda r_2 > r$. This implies that price growth allows old agents to repay their debt with the proceeds of the house sale.

Then, (3.17) becomes

$$W_t = \bar{S}^P[u_P + (\lambda r_2 - r)\bar{B}_t] + (1 - \lambda)\bar{S}^P[r_2\bar{B}_t - c^P]. \quad (3.18)$$

At the other extreme, consider now an equilibrium with no bubble (so that $p_t^P = 0, \forall t$ and no new houses enter the market). In this case, (3.17) becomes

$$W_t = \lambda S_t^P u_P. \quad (3.19)$$

Clearly enough, (3.18) is always larger than (3.19). Moreover, in the second case $\lim_{t \rightarrow \infty} W_t = 0$, so that the total surplus would converge to zero in the long run. It is also clear that, while a bubble equilibrium may also generate a higher surplus in the rich market, so long as $c^R \leq F^R$, the supply of rich assets will never go to zero in the long run (so that $\lim_{t \rightarrow \infty} W_t > 0$). This is because the price of the rich asset is bounded below by the fundamental F^R .²²

Proposition 3.8 *Suppose Assumption 3.4 holds. Then comparing an equilibrium without a bubble ($r_2 < r$) to an equilibrium with a bubble ($\lambda r_2 > r$) welfare is higher under the bubble scenario.*

3.4.2 Change in LTV Policy

Now, consider a scenario where the banks no longer supply 100% LTV loans. This would mean that the poor cannot access credit, given that they have no endowment. When this change is announced there will be widespread default among the poor since the young generation cannot receive any credit, implying that the old cannot find anybody to sell their house to. Thus, the old at the time of the announcement are unable to pay back their debt. It is possible to enrich the model by endogenizing lending in order to generate endogenous credit freezes.

²²Note that part of our welfare result is due to the fact that the interest rate r is fixed and does not respond to changed conditions.

3.5 Policy Implications and Conclusion

Credit-fuelled bubbles, where the price is above the fundamental, may sometimes make the poor agents better-off compared to a scenario without a bubble, where the poor asset would be significantly underpriced. For example, consider a scenario where the poor with access to good investment opportunities need to take a short-term loan. If the house prices are high (overvalued), the poor may be able to collateralize the house and borrow, whereas if their houses are significantly underpriced they may not be able to access credit at all. This argument is similar to [Tirole \(1985\)](#). He argues that even if bubbles crowd out total investment, still they can improve the flow and allocation of funds through relaxing the borrowing constraint for the investors with a good investment opportunity, i.e., the bubble can lead to a Pareto improvement.

In terms of policy, our results suggest that there may be scope for market interventions aimed at sustaining the value of the assets held by credit-constrained agents after the burst of a credit bubble.

Appendices

Appendix

Appendix to Ch.2 and Ch.3

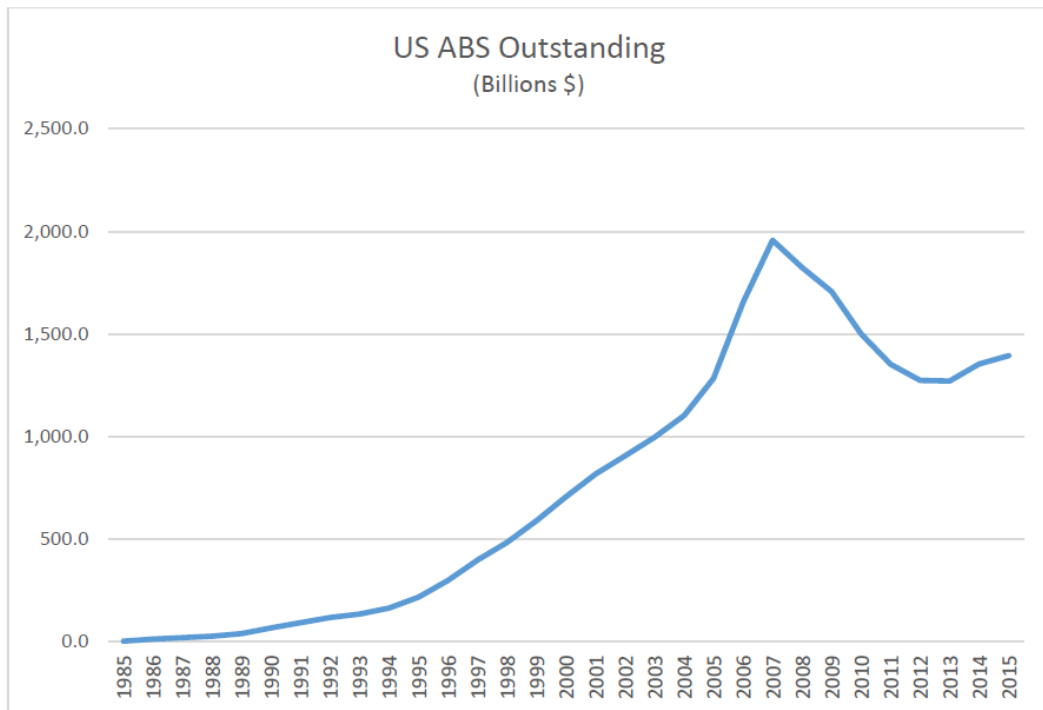
Proofs

Lemma 3.9 *Suppose A.3.3 is violated $A < \underline{A}$, then the equilibrium in the rich market is similar to that of the poor market.*

- (i) *If $r_2 \geq r$, then there exists an equilibrium $p_t^{R*} = \bar{B}_t + D^*$ where $D^* = A < F^R$. WTB_R (3.11) does not bind and (3.8) binds, i.e the rich obtain a positive surplus from purchasing the asset.*
- (ii) *If $r_2 < r$, then the rich borrow in equilibrium $B_t^* = B_{t+1}^* = \frac{A}{r-1}$, and equilibrium price is $p_t^{R*} = p_{t+1}^{R*} = p^* = \frac{Ar}{r-1}$.*

The proof is similar to Lemma 3.2. For part (ii) the only difference from the poor is any price smaller than the endowment can be satisfied as equilibrium, and no agent borrows. The intuition is similar to Lemma 3.2, when the credit growth is smaller than the cost of borrowing, rich agents have an upper limit for the borrowing which is determined by the binding no default condition. Thus, the asset is traded below fundamental value since $p^* = \frac{Ar}{r-1} < \frac{\underline{A}r}{r-1} = F^R$ since $\underline{A} = \frac{u}{r}$.

Figures and Graphs



Source: Securities Industry and Financial Markets Association

Figure A.7: Outstanding Asset Backed Securities in US

Quarterly issuance of traditional and nontraditional securitizations

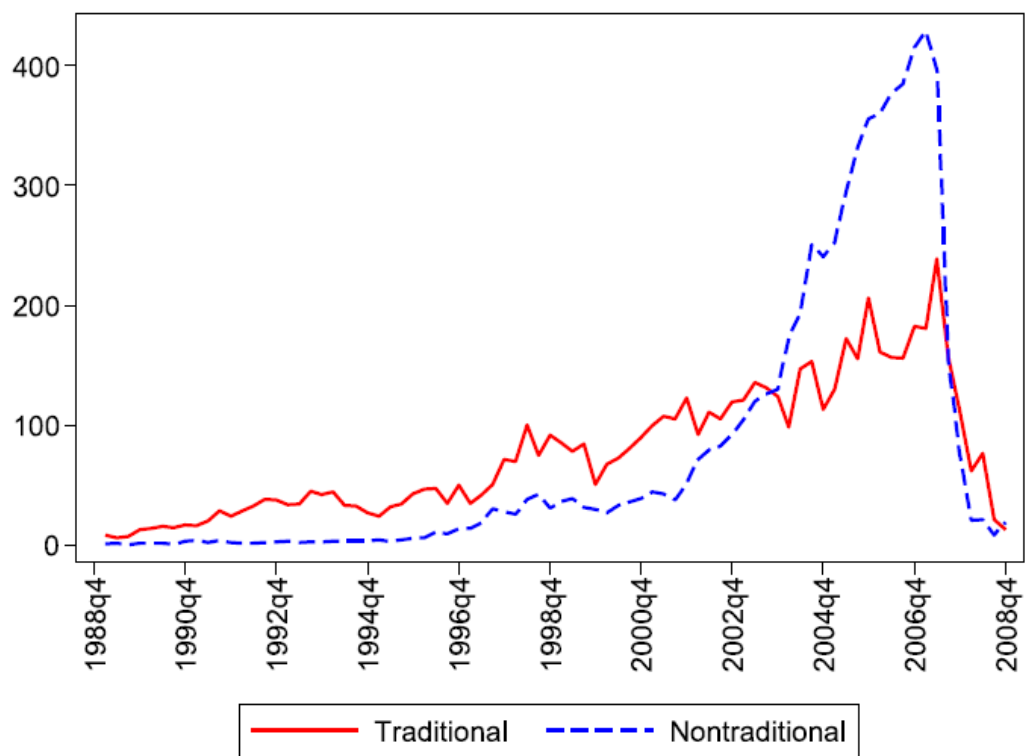


Figure A.8: Issuance of Securitization in billion \$

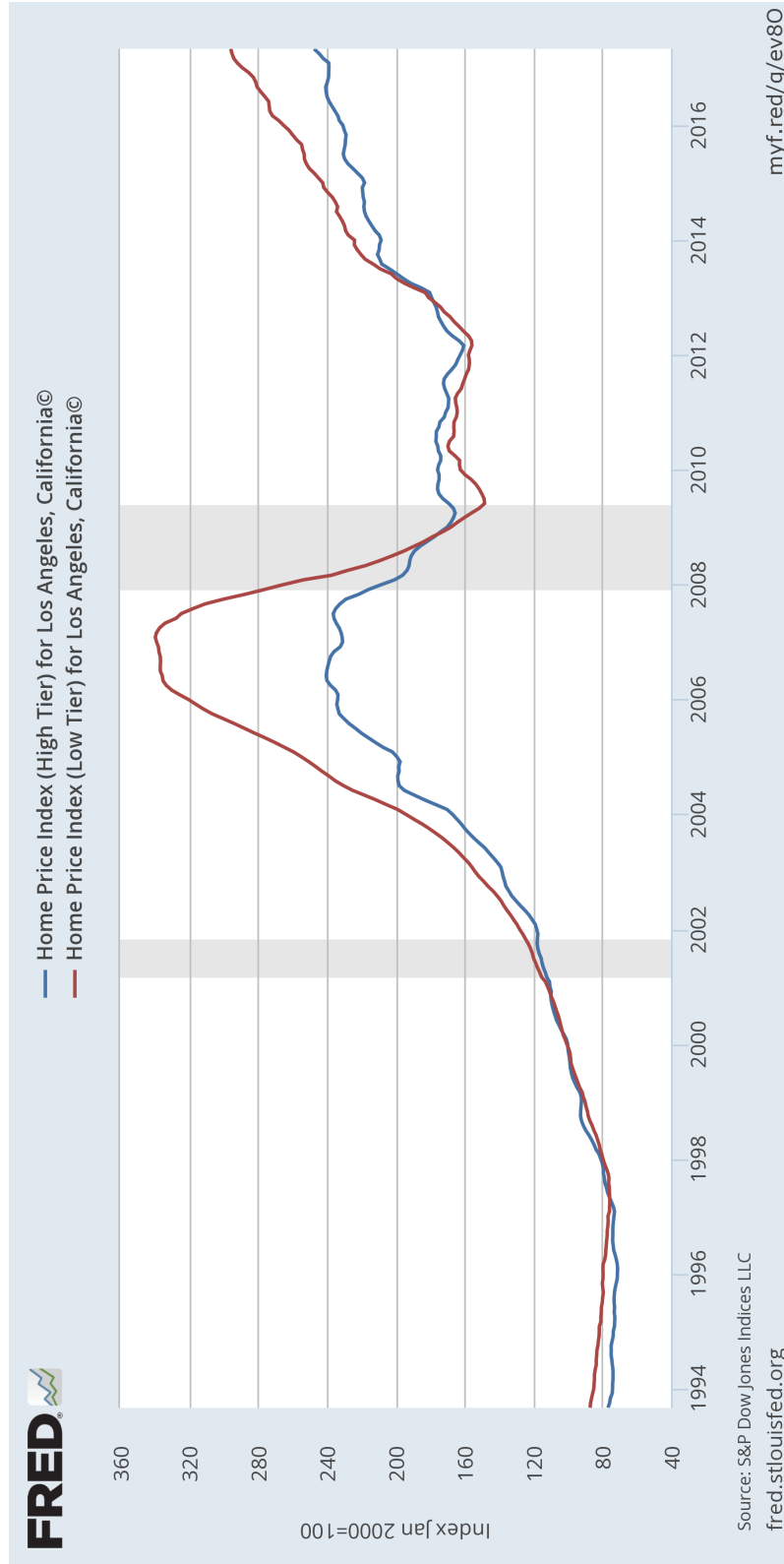


Figure A.9: Case-Shiller Home Price Index (High vs Low Tier) for Los Angeles, California

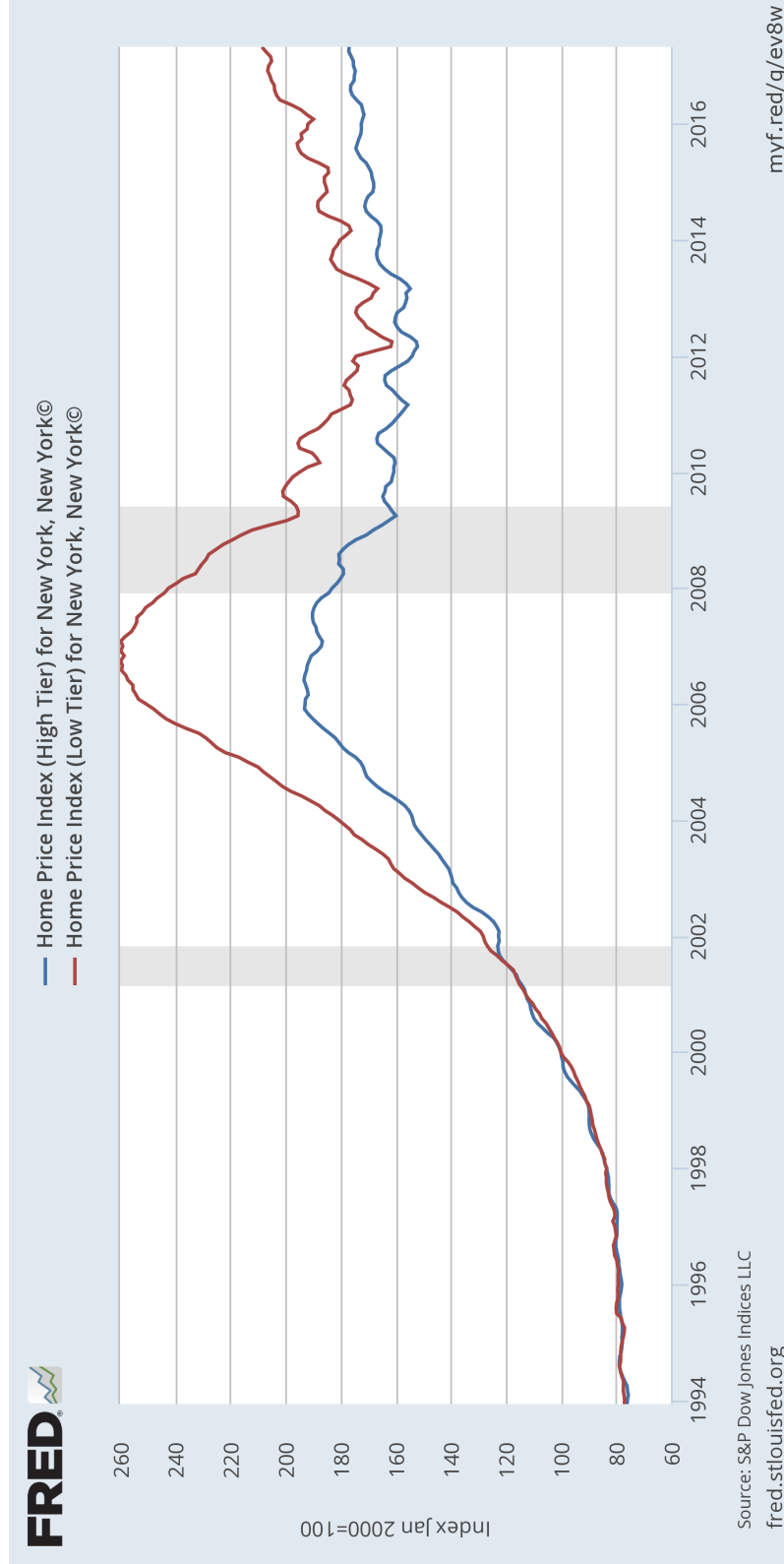


Figure A.10: Case-Shiller Home Price Index (High vs Low Tier) for New York, New York

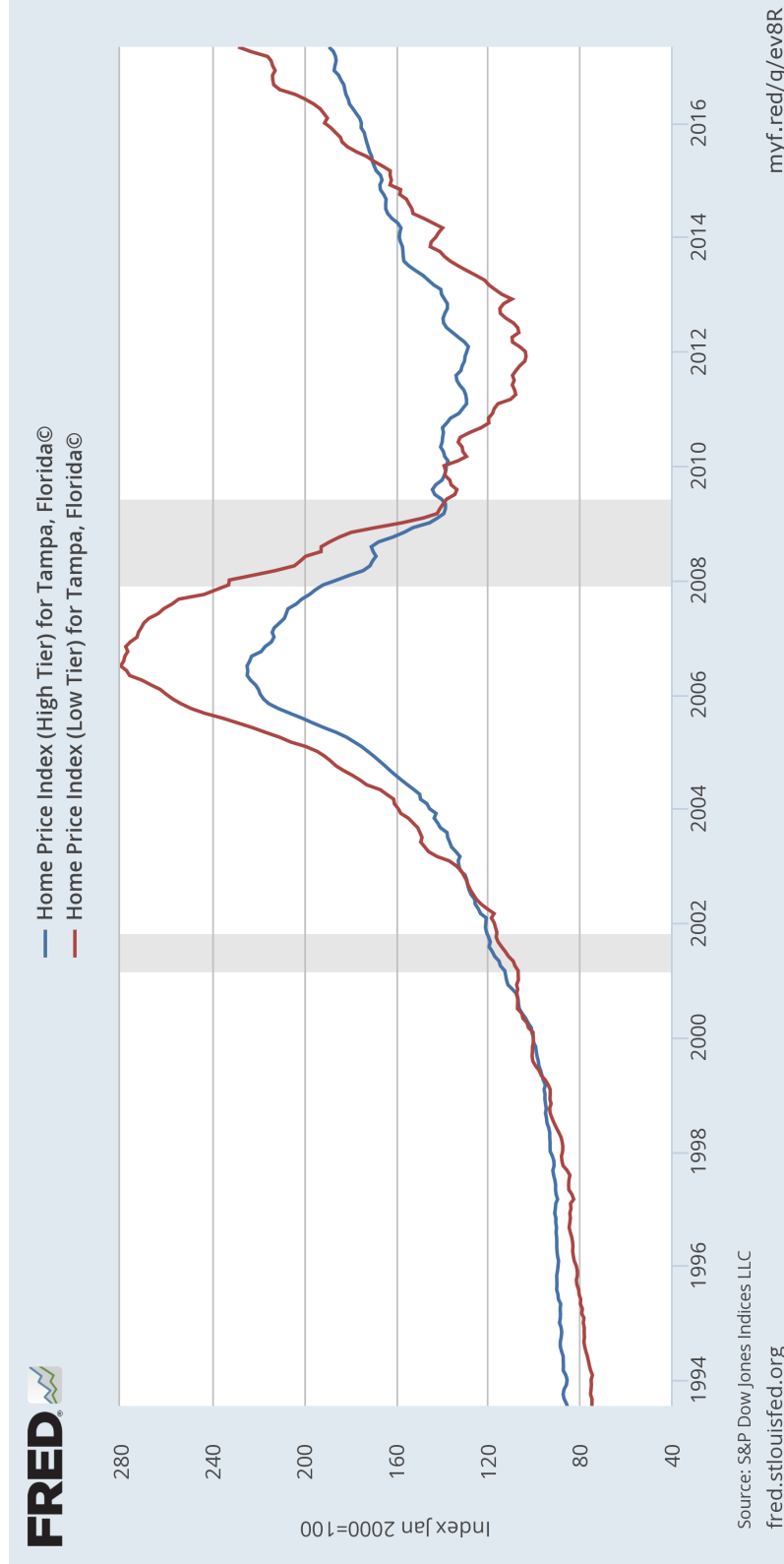


Figure A.11: Case-Shiller Home Price Index (High vs Low Tier) for Tampa, Florida

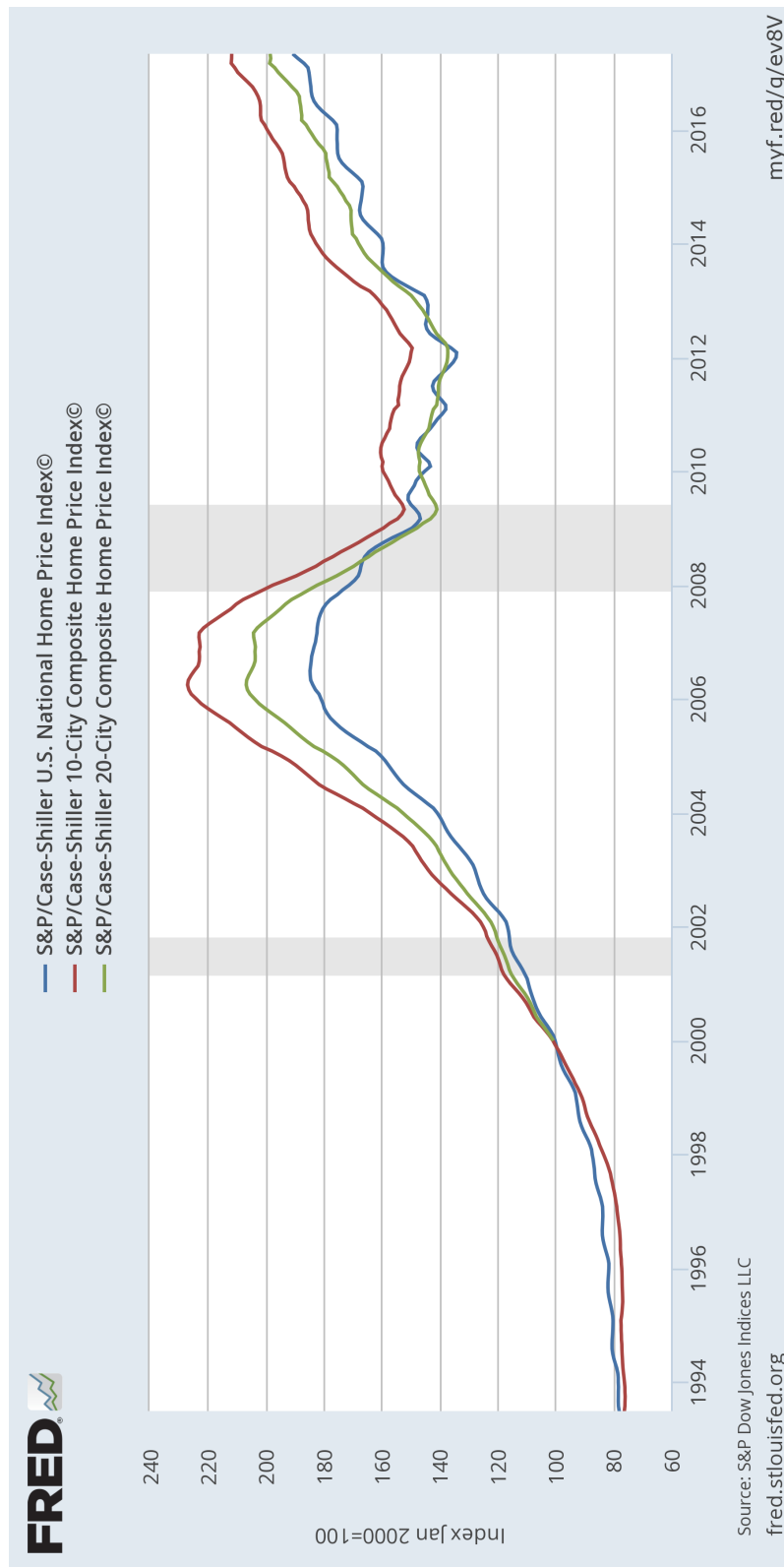


Figure A.12: Case-Shiller National, 10-City and 20-City Home Price Indices for the USA

Bibliography

- Abreu, D. and Brunnermeier, M. K. (2002). Synchronization risk and delayed arbitrage. *Journal of Financial Economics*, 66(2):341–360.
- Abreu, D. and Brunnermeier, M. K. (2003). Bubbles and crashes. *Econometrica*, 71(1):173–204.
- Acharya, V. V., Pedersen, L. H., Philippon, T., and Richardson, M. (2017). Measuring systemic risk. *Review of Financial Studies*, 30(1):2–47.
- Acharya, V. V., Schnabl, P., and Suarez, G. (2013). Securitization without risk transfer. *Journal of Financial Economics*, 107(3):515–536.
- Admati, A. R., DeMarzo, P. M., Hellwig, M. F., and Pfleiderer, P. C. (2011). Fallacies, irrelevant facts, and myths in the discussion of capital regulation: Why bank equity is not expensive. *MPI Collective Goods Preprint*, (2010/42).
- Aizenman, J. and Jinjara, Y. (2009). Current account patterns and national real estate markets. *Journal of Urban Economics*, 66(2):75–89.
- Akhtar, S., Faff, R., Oliver, B., and Subrahmanyam, A. (2012). Stock salience and the asymmetric market effect of consumer sentiment news. *Journal of Banking & Finance*, 36(12):3289–3301.
- Allen, F., Carletti, E., Goldstein, I., and Leonello, A. (2015). Moral hazard and government guarantees in the banking industry. *Journal of Financial Regulation*, page fju003.
- Allen, F. and Gale, D. (1998). Optimal financial crises. *The Journal of Finance*, 53(4):1245–1284.
- Allen, F. and Gale, D. (2000). Financial contagion. *Journal of Political Economy*, 108(1):1–33.

- Allen, F. and Gale, D. (2003). Capital adequacy regulation: in search of a rationale. *Economics for an imperfect world: Essays in honor of Joseph Stiglitz*, pages 83–109.
- Allen, F., Morris, S., and Postlewaite, A. (1993). Finite bubbles with short sale constraints and asymmetric information. *Journal of Economic Theory*, 61(2):206–229.
- Altunbas, Y., Gambacorta, L., and Marques-Ibanez, D. (2010). Bank risk and monetary policy. *Journal of Financial Stability*, 6(3):121–129.
- Andreassen, P. B. (1990). Judgmental extrapolation and market overreaction: On the use and disuse of news. *Journal of Behavioral Decision Making*, 3(3):153–174.
- Andreassen, P. B. and Kraus, S. J. (1990). Judgmental extrapolation and the salience of change. *Journal of Forecasting*, 9(4):347–372.
- Baker, M. and Wurgler, J. (2007). Investor sentiment in the stock market. *The Journal of Economic Perspectives*, 21(2):129–151.
- Baker, M., Wurgler, J., and Yuan, Y. (2012). Global, local, and contagious investor sentiment. *Journal of Financial Economics*, 104(2):272–287.
- Barlevy, G. (2014). A leverage-based model of speculative bubbles. *Journal of Economic Theory*, 153:459–505.
- Barlevy, G. (2015). Bubbles and fools. *Economic Perspectives*, 39(2):54–77.
- Barth, J. R., Caprio Jr, G., and Levine, R. (2004). Bank regulation and supervision: what works best? *Journal of Financial Intermediation*, 13(2):205–248.
- Beccalli, E., Boitani, A., and Di Giuliantonio, S. (2015). Leverage procyclicality and securitization in us banking. *Journal of Financial Intermediation*, 24(2):200–230.
- Bernanke, B. S. et al. (2005). The global saving glut and the us current account deficit. Technical report.
- Blanchard, O. J. (1979). Speculative bubbles, crashes and rational expectations. *Economics letters*, 3(4):387–389.
- Bordalo, P., Gennaioli, N., and Shleifer, A. (2012). Salience theory of choice

- under risk. *The Quarterly Journal of Economics*, 127(3):1243–1285.
- Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, 121(5):803–843.
- Brunnermeier, M. K. (2008). Bubbles. In *New Palgrave Dictionary of Economics*, volume 8.
- Brunnermeier, M. K. and Oehmke, M. (2013). Bubbles, financial crises, and systemic risk. In *Handbook of the Economics of Finance, Chapter 18*, volume 2, pages 1221–1288. Elsevier.
- Calomiris, C. W. (2009). Financial innovation, regulation, and reform. *Cato J.*, 29:65.
- Chernenko, S., Hanson, S. G., and Sunderam, A. (2016). Who neglects risk? investor experience and the credit boom. *Journal of Financial Economics*, 122(2):248–269.
- Conlon, J. R. (2004). Simple finite horizon bubbles robust to higher order knowledge. *Econometrica*, 72(3):927–936.
- Conlon, J. R. (2015). Should central banks burst bubbles? some microeconomic issues. *The Economic Journal*, 125(582):141–161.
- Da, Z., Engelberg, J., and Gao, P. (2014). The sum of all fears investor sentiment and asset prices. *The Review of Financial Studies*, 28(1):1–32.
- De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, 98(4):703–738.
- Diamond, D. W. and Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *The Journal of Political Economy*, pages 401–419.
- Diba, B. T. and Grossman, H. I. (1988). The theory of rational bubbles in stock prices. *The Economic Journal*, 98(392):746–754.
- Favilukis, J., Kohn, D., Ludvigson, S. C., and Van Nieuwerburgh, S. (2012). International capital flows and house prices: Theory and evidence. In *Housing and the Financial Crisis*, pages 235–299. University of Chicago Press.
- Fender, I. and Mitchell, J. (2009). Incentives and tranche retention in securitisation: a screening model. *BIS Working Paper*, 289.

- Ferreira, F., Gyourko, J., and Tracy, J. (2010). Housing busts and household mobility. *Journal of Urban Economics*, 68(1):34–45.
- Ferrero, A. (2015). House price booms, current account deficits, and low interest rates. *Journal of Money, Credit and Banking*, 47(S1):261–293.
- Galí, J. (2014). Monetary policy and rational asset price bubb. *The American Economic Review*, 104(3):721–752.
- Gärling, T., Kirchler, E., Lewis, A., and Van Raaij, F. (2009). Psychology, financial decision making, and financial crises. *Psychological Science in the Public Interest*, 10(1):1–47.
- Gennaioli, N., Shleifer, A., and Vishny, R. (2012). Neglected risks, financial innovation, and financial fragility. *Journal of Financial Economics*, 104(3):452–468.
- Gennaioli, N., Shleifer, A., and Vishny, R. (2015). Neglected risks: The psychology of financial crises. *The American Economic Review*, 105(5):310–314.
- Gennaioli, N., Shleifer, A., and Vishny, R. W. (2013). A model of shadow banking. *The Journal of Finance*, 68(4):1331–1363.
- Goldstein, I. and Razin, A. (2013). Three branches of theories of financial crises. Technical report, National Bureau of Economic Research.
- Gorton, G. and Metrick, A. (2012). Securitized banking and the run on repo. *Journal of Financial Economics*, 104(3):425–451.
- Greenlaw, D., Hatzius, J., Kashyap, A. K., and Shin, H. S. (2008). Leveraged losses: lessons from the mortgage market meltdown. In *Proceedings of the US Monetary Policy Forum*, pages 7–59.
- Greenspan, A. (2011). Regulators must risk more, and intervene less. *Financial Times*, 26 July.
- Greenwood, R. and Shleifer, A. (2014). Expectations of returns and expected returns. *Review of Financial Studies*, 27(3):714–746.
- Harrison, J. M. and Kreps, D. M. (1978). Speculative investor behavior in a stock market with heterogeneous expectations. *The Quarterly Journal of Economics*, 92(2):323–336.

- Henderson, B. J. and Pearson, N. D. (2011). The dark side of financial innovation: A case study of the pricing of a retail financial product. *Journal of Financial Economics*, 100(2):227–247.
- Huang, D., Jiang, F., Tu, J., and Zhou, G. (2014). Investor sentiment aligned: A powerful predictor of stock returns. *The Review of Financial Studies*, 28(3):791–837.
- Justiniano, A., Primiceri, G. E., and Tambalotti, A. (2014). The effects of the saving and banking glut on the us economy. *Journal of International Economics*, 92:S52–S67.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Kahneman, D. and Tversky, A. (1972). Subjective probability: A judgment of representativeness. pages 25–48.
- Kane, E. (2012). The inevitability of shadowy banking. *Available at SSRN*.
- Kane, E. J. (1981). Accelerating inflation, technological innovation, and the decreasing effectiveness of banking regulation. *The Journal of Finance*, 36(2):355–367.
- Kisin, R. and Manela, A. (2016). The shadow cost of bank capital requirements. *Review of Financial Studies*.
- Klibanoff, P., Lamont, O., and Wizman, T. A. (1998). Investor reaction to salient news in closed-end country funds. *The Journal of Finance*, 53(2):673–699.
- Martin, A. and Ventura, J. (2012). Economic growth with bubbles. *The American Economic Review*, 102(6):3033–3058.
- Mendel, B. and Shleifer, A. (2012). Chasing noise. *Journal of Financial Economics*, 104(2):303–320.
- Nelson, M. W., Bloomfield, R., Hales, J. W., and Libby, R. (2001). The effect of information strength and weight on behavior in financial markets. *Organizational Behavior and Human Decision Processes*, 86(2):168–196.
- Obstfeld, M. and Rogoff, K. (2009). Global imbalances and the financial crisis: products of common causes.
- Palomino, F., Renneboog, L., and Zhang, C. (2009). Information salience, in-

- vestor sentiment, and stock returns: The case of british soccer betting. *Journal of Corporate Finance*, 15(3):368–387.
- Parlour, C. A. and Plantin, G. (2008). Loan sales and relationship banking. *The Journal of Finance*, 63(3):1291–1314.
- Plantin, G. (2011). Good securitization, bad securitization. Technical report, Institute for Monetary and Economic Studies, Bank of Japan.
- Plantin, G. (2014). Shadow banking and bank capital regulation. *Review of Financial Studies*.
- Rajan, R. G. (2006). Has finance made the world riskier? *European Financial Management*, 12(4):499–533.
- Scheinkman, J. A. and Xiong, W. (2003). Overconfidence and speculative bubbles. *Journal of Political Economy*, 111(6):1183–1220.
- Shleifer, A. and Vishny, R. W. (1997). The limits of arbitrage. *The Journal of Finance*, 52(1):35–55.
- Shleifer, A. and Vishny, R. W. (2010). Unstable banking. *Journal of Financial Economics*, 97(3):306–318.
- Stambaugh, R. F., Yu, J., and Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, 104(2):288–302.
- Stein, J. C. (2010). Securitization, shadow banking & financial fragility. *Daedalus*, 139(4):41–51.
- Taylor, S. E. and Thompson, S. C. (1982). Stalking the elusive “vividness” effect. *Psychological Review*, 89(2):155.
- Thakor, A. V. (2012). Incentives to innovate and financial crises. *Journal of Financial Economics*, 103(1):130–148.
- Thakor, A. V. (2015). The financial crisis of 2007–2009: Why did it happen and what did we learn? *Review of Corporate Finance Studies*, 4(2):155–205.
- Tirole, J. (1982). On the possibility of speculation under rational expectations. *Econometrica: Journal of the Econometric Society*, pages 1163–1181.
- Tirole, J. (1985). Asset bubbles and overlapping generations. *Econometrica: Journal of the Econometric Society*, pages 1071–1100.

- Tsuji, C. (2006). Does investors' sentiment predict stock price changes? with analyses of naive extrapolation and the salience hypothesis in japan. *Applied Financial Economics Letters*, 2(6):353–359.
- Uhde, A. and Michalak, T. C. (2010). Securitization and systematic risk in european banking: Empirical evidence. *Journal of Banking & Finance*, 34(12):3061–3077.
- Wall, L. D. (2014). Notes from the vault two drivers of financial innovation.
- Yu, J. and Yuan, Y. (2011). Investor sentiment and the mean–variance relation. *Journal of Financial Economics*, 100(2):367–381.