A MACRO-FINANCE TREATISE ON SYSTEMIC RISK

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Sujan Lamichhane

May 2018

A MACRO-FINANCE TREATISE ON SYSTEMIC RISK

Sujan Lamichhane, Ph.D.

Cornell University 2018

Systemic risk in the macro-finance context has garnered significant interest relatively recently and our understanding of it is limited. Systemic risk is a broad term used to describe economic and financial system breakdown. Its characterization can depend on the sources being explored. Fully understanding its nature is imperative, especially if we want to understand the causes and consequences of big economic meltdowns like the 2007-2009 financial crisis. Policies designed without proper understanding of systemic risk are likely to be either ineffective or have unforeseen ramifications. Thus, there is a need for a variety of models that explore different aspects of systemic risk.

The first chapter of this dissertation studies systemic risk as it relates to financial innovation in a stationary equilibrium. The second chapter studies the transition dynamics aspect. Since these two chapters are based on the same underlying economic model, they are jointly introduced and concluded. In these chapters a heterogeneous agents model in continuous time, driven by jump-diffusion processes, is developed. Methods from the theory of Lévy Processes and Feynman Path Integral are introduced. This approach allows for analytically exploring various properties of systemic risk. We derive explicit expressions of the financial sector's failure probability, its capital position at the random time of credit event, and the transition densities of the leverage and financial wealth distributions. We show that financial innovation can either increase or decrease systemic risk under some conditions. We characterize the notion of

a leverage trap – once the economy moves to high leverage systemic risk states, it tends to stay there. Financial innovation amplifies credit cycles. Transition speed increases (decreases) when the economy is leveraging up (deleveraging).

The third chapter studies how asset price bubbles, market liquidity, and trading constraints affect systemic risk. We build an equilibrium model with heterogeneous agents in which market liquidity is modeled as a stochastic quantity impact from trading on the price. We introduce a different framework for analyzing rational asset price bubbles, which are shown to exist in equilibrium due to heterogeneous beliefs, heterogeneous preferences, and binding trading constraints. Positive price bubbles are larger in illiquid markets and when trading constraints are more binding. A realization of systemic risk, defined as the risk of market failure due to an exogenous shock to the economy, results in a significant loss of wealth as agents are unable to meet their trading constraints and default. Systemic risk is shown to increase as: (i) the fraction of agents seeing an asset price bubble increases, (ii) as the market becomes more illiquid, and (iii) as trading constraints are relaxed.

BIOGRAPHICAL SKETCH

Sujan Lamichhane was born in Kathmandu, Nepal in January 1987. He came to the United States for his undergraduate studies. He earned his B.S. in Economics and a minor in Mathematics (*magna cum laude*) from Towson University in May 2011. A year after graduating college he began his graduate studies in the Department of Economics at Cornell University. He received his M.A. in Economics in February 2016 and will earn his Ph.D. in Economics in May 2018.

This document is dedicated to my late grandmother,

Nanda Kumari Lamichhane.

ACKNOWLEDGEMENTS

I am extremely grateful and deeply indebted to my Professors Robert Jarrow, Eswar Prasad, and Yongmiao Hong for their incredible guidance and support. They all have taught me, in their own unique ways, how to conduct rigorous research and be a good scholar. I am particularly grateful to Professor Robert Jarrow for the opportunity to co-author different papers with him. Working with him has been one of the most fruitful and memorable learning moments of my life as a student.

I am also grateful to Professor Pierre Patie (in the ORIE department at Cornell) for spending his valuable time discussing many mathematical ideas throughout the years. He genuinely inspired me to see beyond the apparent complexities of the equations and notations to ultimately appreciate the inherent beauty of mathematics. I would also like to thank Professor Csaba Csaki (in the Physics department at Cornell) for the unique learning experience in his graduate quantum mechanics class and many helpful discussions that boosted my confidence in applying quantum mechanics methods in my research.

I wish to express my sincere gratitude to Professor Matthew Baron for taking the time to discuss many aspects of my research and tremendously helping me refine and sharpen it. I also thank Professor David Ng for various helpful discussions.

Last but not least, I want to thank my family and friends, both in Nepal and in the U.S., for their support.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Systemic Risk and Financial Innovation

Chapter 2 studies systemic risk as it relates to financial innovation in a stationary equilibrium. Chapter 3 studies the transition dynamics aspect. Since these two chapters are based on Lamichhane (2018) with the same underlying economic model, we jointly introduce and conclude them.

Since the late 1990s we have seen exponential growth of various complex financial derivatives products, such as Credit Default Swaps (CDS), equity options, and interest rate/currency swaps (as seen in Figures (A.3) and (A.4)). These financial innovations are designed to fulfill various (complex) risk management needs in the economy (e.g. credit, liquidity, counterparty, currency, interest rate risks), thereby allowing for efficient risk-sharing. For example, in an incomplete market, having one additional risk sharing product (say a CDS contract) to insure against the risky states of the world is a step towards market completion and can be beneficial to those agents insuring against the risk. Thus, these innovations are often thought of as being beneficial as they improve risk sharing in the economy. But one open question still is, are these innovations always beneficial? Or, could they also destabilize the entire economy? If so, how?

It has been argued that complex financial derivatives such as CDS were the root causes of the recent financial crisis, after which the idea of systemic risk started receiving a lot of attention. Systemic risk is a broad term used to describe economic and financial system breakdown (Billio et al. (2010))[1]. These two chapters explore two closely related questions. First, what is the relation between systemic risk and financial innovation? Second, how do we understand and characterize the evolution of systemic risk over time?

We now briefly motivate our focus on the financial innovation role of the financial sector. Two unique features of financial innovation relate it to systemic risk: *counterparty risk* and moral hazard risk. What are the economy-wide implications when the financial sector, the insurer/counterparty to all the risks in the economy, becomes insolvent? This is the counterparty risk. The effects of initial (small) credit shocks can amplify if the counterparty insuring those risks fails. Further, the ability to insure against credit risks might make lenders supply even more credit which could lead to an excessive amount of outstanding debt in the economy, thereby affecting the financial stability. This is the added risk from moral hazard. Thus, as a first-order effect, financial innovation definitely serves a useful purpose of fulfilling various risk management needs. However, the second-order effect due to the failure possibility of the financial sector, the counterparty to all the risk in the economy, could exacerbate the effects of initial risks rather than contain it. Thus, the debt/leverage and financial innovation channels together can create subtle systemic risk dynamics/implications for the broader economy[2]. The depth of the theoretical/analytical characterizations in

[1]Bisias et al. (2012) argue that we must continuously adapt the systemic risk definition and related measures given that the financial/economic system is continuously evolving. Hence, there is a need for a variety of models that emphasize different aspects of systemic risk. Excellent references that discuss systemic risk from different conceptual frameworks are the two recent NBER conference volumes: Quantifying Systemic Risk (2013) and Risk Topography: Systemic Risk and Macro Modeling (2014).

[2]This is reminiscent of the recent financial crisis when a large portion of the U.S. households had a huge debt overhang going into the crisis. As the financial conditions tightened, they were forced into rapid deleveraging, leading to a large-scale defaults on mortgages, followed by a sharp cut-back in consumption spending. Mian and Sufi (2011a, 2011b) and Eggertsson and Krugman (2012) also emphasize similar mechanisms. Further, the initial shock from defaults was amplified by the failure (or near failure) of large counterparty firms like Lehman Broth-

this regard is still wanting in the literature. The first two chapters make contributions in this direction.

We build a heterogeneous agents model in continuous time driven by jump-diffusion processes and explore the full equilibrium dynamics of an incomplete market endowment economy with financial frictions (borrowing constraint and default/credit risk). There are three classes of agents: heterogeneous borrowers and lenders, where both are heterogeneous in their income, and a representative financial sector. Default/credit risk is modeled via a jump in borrowers' income processes. We introduce methods from the theory of Lévy Processes and Feynman Path Integral from quantum mechanics. This approach allows us to *analytically* explore various properties of systemic risk, including its transition dynamics, while generating a variety of new testable predictions.

The *first distinctive feature* of these chapters is that they focus on the financial innovation role of the financial sector. This is a special case of the financial sector which is still largely unexplored in the literature as opposed to the usual intermediary role that facilitates lending and borrowing – in our model, the financial sector sells credit risk protection to the lenders, insuring them against the default risk of the borrowers. Thus, the financial sector is the *counterparty* to all the risks in the economy (analogous to large counterparty firms like AIG before the crisis)[3]. It receives regular premiums from lenders, which forms its capital position, and it agrees to pay the lenders if borrowers default. Thus, lenders can insure a certain fraction of their loans through the financial sec-

ers and AIG. They could not shoulder the burden of assumed/insured risks, rendering the perceived benefits of financial innovations merely an illusion. Additional details on financial innovation are provided in the appendix Additional Details on Financial Innovation

[3]Gennaioli, Shleifer and Vishny (2012) also view securitization and financial innovation as important forms of risk sharing, driven by market demand for certain patterns of cash flow. This view was also propounded by Ross (1976) and Allen and Gale (1994).

tor, and the parameter that captures this fraction is used to model the usage of financial innovation. A higher fraction means increasing use of financial innovation, i.e. this parameter captures the degree of financial innovation. Note that we introduce financial innovation exogenously because we are interested in understanding the systemic risk implications *given* that financial innovation exists (and is used). Thus, by varying this parameter, we can trace out various degrees of systemic risk implications associated with increasing use of financial innovation.

Thus, the first contribution of these chapters is that we consider a different modeling framework for exploring the financial innovation role of the financial sector. This modeling innovation relies on using some powerful mathematical methods from the theory of Lévy Processes. This approach helps us analytically characterize different aspects of systemic risk vis-à-vis counterparty and moral hazard risks associated with financial innovation.

The *second distinctive feature* of these chapters is that we also make some additional methodological contributions in terms of showing how to *analytically characterize transition dynamics* in the class of equilibrium heterogeneous agent models. This represents the evolutionary dynamics of systemic risk. This is done by introducing *Feynman Path Integral (FPI)* method from Quantum Mechanics, i.e. the probabilistic phenomenon of quantum particles. From a mathematical perspective, Quantum Mechanics is largely concerned with characterizing the probabilistic phenomenon of quantum particles, i.e. the strange behaviors of quantum particles. We focus on FPI as it is simple/intuitive and particularly suitable for our model[4]. The rich set of rigorous mathemat-

[4]Developing path integral allowed Feynman (1948) to completely side step the standard formalism of quantum mechanics via Hamiltonians or the spectral theory of operators in Hilbert Space.

ical/probabilistic methods explored in Quantum Mechanics to deal with uncertainties (associated with quantum particles) make them applicable in economics. In particular, an appealing property of FPI is that it satisfies the *Chapman-Kolmogorov equation* for Markov processes, making it useful for a wide range of economic applications[5].

We explore two closely related dimensions of systemic risk, as we define it in terms of: (i) the *failure probability* of the financial sector, i.e. its capital position going negative, and (ii) the tail of the *equilibrium leverage distribution* which represents the total mass of borrowing households that are at or around their borrowing constraint. With our methods, we are able to derive four explicit expressions that sharply characterize systemic risk: the financial sector's failure probability (which is equivalent to the distribution of first time such an event occurs), its capital position *at the random time* of credit event when borrowers default, and the transition densities of the leverage and financial wealth distributions. These expressions deliver various insights. Thus, we now briefly highlight three key takeaways from our results and some policy implications.

First, we show that financial innovation can either increase or decrease systemic risk *under some conditions*, as we highlight the interplay between counterparty and moral hazard risks. Thus, even though financial innovation is designed to and does mitigate risks, sometimes it can also magnify risk for the entire economy. Second, we introduce a new notion of a *leverage trap* – once the economy moves to high leverage systemic risk states, it tends to stay there.

[5]We provide a brief review of these methods relevant to these chapter (with many accessible references) in Theory of Lévy Processes and Feynman Path Integral appendices. The benefits of using these new methods are immense. This is reflected in the proofs of various results, many of which are straightforward applications of these methods. These are also valuable additions to economists' arsenal of tools. Further, the FPI formalism is largely independent of stochastic calculus methods. This means combining these two methods can allow for a natural emergence of new ideas and insights.

Stated differently, the deeper the economy moves into high leverage systemic risk states, the lower the chances of moving out of such states, essentially trapping the economy in the high leverage states. Third, we show that greater use of financial innovation when credit and the economy are already contracting can slow the recovery process. This result is the combination of two intermediate results. First, we show that financial innovation amplifies credit cycles (or roughly speaking, financial innovation is pro-cyclical). This means the total credit supplied can increase (decrease) with financial innovation when the economy is doing well (not as well). Second, the transition speed (which captures the time to recovery), of moving from some initial stationary equilibrium to another new one, increases (decreases) when the economy is leveraging up (deleveraging).

Our results suggest that policy should not require greater use of credit risk protections when credit and the economy are already contracting. Such a policy will, first, induce even lower credit supply in the economy. Second, such a policy can also make the recovery process slower (i.e. transition speed here) as lower credit supply generally means the economy is not able to leverage up, making the recovery process slower. Thus, the effect of requiring lenders to purchase more credit risk protection is not clear as the outcome can depend on whether the economy is in stationary equilibrium or in a transition phase.

The first two chapters draws from three different strands of literature that have been largely evolving separately: (i) macroeconomics literature on financial frictions and its evolution to the recent macro-finance literature[6], (ii) litera-

[6]Bernanke and Gertler (1989), Kiyotaki and Moore (1997) and Bernanke, Gertler, and Gilchrist (1999) are some of the early papers that study the effects of financial frictions (eg. borrowing/collateral constraints) on the macroeconomy, highlighting channels through which temporary financial shocks can have persistent and/or amplified effects. Many papers since the financial crisis appeal to this framework, such as Eggertsson and Krugman (2012), Jermann and

ture on heterogeneous agent models, and (iii) literature on credit risk modeling. In particular, the first two chapters are related to Brunnermeier and Sannikov (2014, BS-2014 hereafter) and He and Krishnamurthy (2014, HK-2014 hereafter) with regards to exploring macro-finance linkages. From the methods side, exploring the rich structure of heterogeneous agents model in continuous time, the first two chapters are related to Achdou, Han, Lasry, Lions and Moll (2015) and Gabaix, Lasry, Lions and Moll (2016). From a credit risk modeling approach, these chapters are related to Jarrow and Turnbull (1995) and Lando (1998). We contribute to the macro-finance and heterogeneous agent modeling literatures.

Like BS-2014 and HK-2014, the first two chapters features similar continuous time methodology to study equilibrium dynamics of an economy with financial frictions. However, there are important structural differences. In HK-2014 model, intermediaries raise equity and debt from households to fund capital and housing purchases. Intermediaries face equity capital constraints, and systemic risk arises when capital constraints are binding. BS-2014 highlight nonlinear effects during crisis periods when small shocks can easily be amplified. Systemic risk is understood as the total volatility of the value of productive capital. We explore a full heterogeneous agents model, where systemic risk and its amplifications can come from different channels (e.g. default intensity, risk aversion, use of financial innovation, state/drift of the economy etc.) that eventually affect the measure of the equilibrium leverage distribution across borrowers (around the high leverage systemic risk states), and the failure probability of the financial sector. Additionally, we show the existence of a Dirac measure, i.e. point mass, *at* the borrowing constraint.

Quadrini (2012), He and Krishnamurthy (2012, 2013, 2014), Brunnermeier and Sannikov (2014). For a literature review on financial frictions and macroeconomics, see Brunnermeier, Eisenbach and Sannikov (2012).

We also employ a heterogeneous agents modeling paradigm which has been primarily used for studying income/wealth inequality issues. Bewley (1986), Huggett (1993), Aiyagari (1994), and Krusell and Smith (1998) are early papers studying heterogeneous household economies[7]. While we do not study inequality issues, we do explore *financial* wealth and leverage distributions. Further, the literature mainly studies stationary equilibrium and numerical computations are used almost exclusively for studying transition dynamics (see Gabaix et al. (2016) for references). As discussed above, in addition to analyzing stationary equilibrium, we also *analytically characterize transition dynamics*, which represents the evolutionary dynamics of systemic risk. Thus, relative to the literature where many results rely on numerics, we show all of our results analytically, highlighting some new economic phenomenon during the transition phase (e.g. the notion of a *leverage trap*).

Achdou et al. (2015) and Gabaix et al. (2016) are notable exceptions. Achdou et al. (2015) highlight the advantages of the theories/methodology available in continuous time and showcase efficient algorithms for numerical computations (relatively few papers study equilibrium models with heterogeneous agents in continuous time). We closely follow the numerical techniques in Achdou et al. (2015), using finite difference method, to compute stationary equilibrium. However, our results do not depend on these computations, and we only compute to (visually) illustrate some of our results. Recently, Gabaix et al. (2016) analytically characterized the transition dynamics of income distribution, though in a reduced form setting. They focus on transition speed of an economy not necessarily at the steady state, but transitioning towards it, as op-

[7]Hopenhayn (1992) studies heterogeneous firms. Benhabib, Bisin, and Zhu (2011, 2013, 2014) and Zhu (2010) have used the such models to explore economic forces that give rise to the observed income/wealth distributions. Heathcote, Storesletten and Violante (2009) and Guvenen (2012) provide excellent surveys on macroeconomics with heterogeneous households.

posed to most existing papers that study the dynamics of an economy already in the steady state. We also analytically characterize transition speed, but in context of an equilibrium model. Further, we explore various other transition dynamics/evolutionary properties, such as *leverage trap*, exit-time distribution, pro-cyclicality of financial innovation etc.

In particular, we also employ tools of partial differential equations and operators as in Gabaix et al. (2016). Interestingly, spectral theory of operators (in Hilbert space) explored in Gabaix et al. (2016) is used in the standard approach to quantum mechanics via Hamiltonian operator (or the *Schrödinger* equation). We emphasize the FPI approach for transition dynamics as it allows for a simplified characterization of transition dynamics in our model[8]. Nevertheless, due to mathematical equivalence, the tools introduced in these two chapters complement those in Gabaix et al. (2016).

Merton (1974) introduced the structural approach to credit risk modeling. This lead to the development of reduced form approach by Jarrow and Turnbull (1995, 1997), Lando (1998), and Jarrow, Lando, and Yu (2005) that addressed the limitations of the structural approach. For analytic traction and efficient modeling of credit default risk, we follow the reduced form approach, where default arrives at some random (stopping) time. This time is characterized by jump in the income process of the borrowers. For more details on credit risk models see the review by Jarrow (2009).

[8] Ait-Sahalia (1999) also developed a method for closed-form approximations to the transition function based on Hermite expansion of density around normal density function. This is advantageous for asset pricing applications. But the main issue is that the reliability of approximation diminishes when the time horizon is large, say more than six months or so. The power of FPI is that regardless of the time interval, the transition density function remains valid.

1.2 Asset Price Bubbles, Market Liquidity and Systemic Risk

The third chapter, based on Jarrow and Lamichhane (2018), studies the effects of financial markets on the real macroeconomy. In the process, we introduce a different framework for analyzing rational asset price bubbles in equilibrium, which should prove useful for formulating policy prescriptions. In particular, we study how asset price bubbles, market liquidity, and trading constraints affect systemic risk. Indeed, many economic crises have been associated with the failure of market liquidity and the bursting of asset price bubbles. The two most recent examples are the liquidity crisis of 1998 due to the failure of LTCM and the 2007 credit crisis due to bursting of the housing price bubble. Government intervention was necessary in both cases to ensure continued market liquidity and (allegedly) to avoid market failure (see Brunnermeier and Pedersen (2008) for more details on these crises).

The sheer size of financial markets evidences its potential for impacting the real economy. In this regard, consider some US household asset allocation data from the Federal Reserve's quarterly release of the Financial Accounts of the United States (Z.1), which includes flow of funds, balance sheet, and integrated macroeconomic accounts. The June 2017 release shows that the net-worth of US households and nonprofits rose to $94.8 trillion during the first quarter of 2017. For this time period, total asset values were about $109.98 trillion and liabilities were about $15.15 trillion. Of this, financial assets were about $77.11 trillion while the market value of equity shares were $26.88 trillion. This means that financial assets constituted about 70% of total asset value and equities about 24.44%. These numbers document the significance of financial assets as a percent of total assets, and the importance of equity. In contrast, for the same time period, non-financial assets constituted about \$32.87 trillion and among such assets, real estate accounted for about \$26.86 trillion, which is 81.71% of non-financial assets or 24.44% of total assets. Thus, equity and real estate represent roughly same percent of US households' total assets.

Given these magnitudes, it should come as no surprise that shocks to prices in financial and related asset markets can easily affect the real macroeconomy. The bursting of price bubbles are an undeniable phenomena in asset markets, be it equities, real estate/housing, or commodities[9]. One of the first recorded price bubbles was the Dutch tulip mania of the mid 1600s. Other notable bubbles were related to the South Sea company in 1720 (Garber 1989,1990), the equity price bubble preceding the Great Depression in the U.S. (White 1990), the dot com bubble (Brunnermeier and Nagel, 2004), and the US housing price bubble before 2007 (Clark and Coggin, 2011). It has been argued that the Great Recession of 2007-2009 was precipitated by massive declines in both equity and real estate prices – the bursting of price bubbles in these markets.

Market liquidity is another dimension of asset markets that interacts with asset price bubbles to affect the macroeconomy. The liquidity of asset markets can affect both the magnitude and severity of asset price bubbles. To understand this interaction, consider housing versus equity markets, in particular the equity futures market. According to the recent U.S. Census Bureau report, the median price of new homes sold in the U.S. in June 2017 was \$311,600. Compare that to the E-mini S&P 500 Index futures, one of the most liquid index futures contracts, which traded around 2500 points (rounding up) during the same period. The

[9] All these markets are connected to the financial sector that intermediates the flow of funds within the economy. We will use the terms financial sector or markets interchangeably. This should cause no confusion as asset markets essentially operate through the financial sector, be it lending, borrowing, or insurance.

notional value of one contract is 2500 times 50 (the multiplier) or $125,000. So three futures contracts total $375,000 in notional; about the same value as the median home.

Now assume that an agent wants to sell their home. One can not ignore various frictions associated with the sale, such as commissions and search cost. These costs reflect the illiquid nature of the housing market. In contrast, assume that an agent wants to sell three S&P 500 futures position. The process is literally just one click of a button (or two depending on the broker), and the transaction costs, in every aspect are negligible. Most brokers charge about $2.50 to $3 per-contract to open or close the trade, which includes commissions and various fees. Consequently, one suspects that price bubbles should be more prevalent in the S&P futures markets.

This chapter constructs an equilibrium model to study the impact of asset price bubbles and market liquidity on systemic risk. The model is an extension of Jarrow (2017b). The setting is discrete time with a finite horizon. The economy is populated by heterogeneous agents/households facing trading constraints. Market liquidity costs are modeled as a stochastic quantity impact from trading on price, where the size of the impact depends on the trade size. Traded are two assets: a bond/money market account and a risky asset/stock. Borrowing and lending occurs where agents can only borrow up to a certain fraction of their equity value. Short sales are allowed but margin must be posted to insure coverage of the short position at a future date[10]. All agents are risk averse,

[10]Note that our trading constraint has a very general structure. First, it limits direct borrowing by requiring collateral, a widely used approach in the macro literature. Second, it limits indirect borrowing by restricting the magnitude of short positions. Although short sale restrictions are widely used in the finance literature, this constraint is largely ignored in macro. Our formulation jointly considers both of these constraints, and hereafter the term *trading constraint* will be used to describe both of these restrictions.

maximizing the expected utility of terminal consumption.

In this economy, we show that asset price bubbles can arise endogenously in a rational equilibrium due to heterogeneous beliefs, heterogeneous preferences, and binding trading constraints. A bubble is defined as the difference between the actual/observed market price of the asset and the fundamental value that agents assign given their own beliefs, preferences, and optimal (constrained) trading strategy. Due to this economic structure, some agents may see bubbles while others may not. We define systemic risk as the risk of market failure due to an exogenous shock to the economy that results in funding illiquidity, which is the conjunction of market illiquidity (i.e. liquidity risk) and binding trading constraints. In our setting this is equivalent to the shock resulting in the inability of agents to meet their trading constraints, leading to default and the nonexistence of an economic equilibrium. Such a market failure implies a large loss of wealth in the economy.

Consistent with the previous intuition, we show the following:

- Due to borrowing constraints, negative price bubbles exist (assets are undervalued), and they are larger (smaller) in more liquid markets (illiquid markets).
- Due to short sale constraints, positive price bubbles exist (assets are overvalued), and they are smaller (larger) in more liquid markets (illiquid markets).
- The percentage of agents in the economy who view negative asset price bubbles increase as a market becomes more liquid.
- The percentage of agents in the economy who view positive asset price

bubbles decreases as a market becomes more liquid.

- The magnitude of a bubble increases when trading constraints are more restrictive.
- Systemic risk increases as the percentage of agents who see bubbles increases.
- Systemic risk increases as the market becomes more illiquid.
- Systemic risk decreases (increases) when the trading constraints are more restrictive (relaxed).

After the financial crisis of 2007-2009, many scholarly papers in macro-finance have qualitatively discussed the interaction among systemic risk, market liquidity and asset price bubbles, generally understood to be the difference between the market price and a unique fundamental price (for example, see Brunnermeier and Oehmke (2012), Hall (2011) and references therein.). The implication, of course, is that everyone sees (or should see) the same price bubble. In contrast, in the context of our heterogeneous agents economy, agents can differ both in whether a price bubble exists and if it does, its magnitude[11]. The systemic risk implications of agents seeing bubbles is important because the existence of bubbles increases systemic risk, i.e. massive agent defaults and market failure, which leads to a significant loss of wealth in the economy.

This insight implies that what is relevant for macro/monetary policy is gauging the total *fraction of agents that see price bubbles* and *not* whether the level

[11]This link between bubbles and heterogeneous beliefs was started by Harrison and Kreps (1978). In this literature, bubbles arise because agents disagree about an asset's fundamental value and they are trading/short sale constrained. See the review papers by Brunnermeier and Oehmke (2012) and Xiong (2013) for more details.

of market prices is too large[12]. In fact, it may be counterproductive for policy makers to seek a universal price bubble that all agents agree on. To take action based on the level of market prices, policy makers need to verify its existence and magnitude, which is a difficult task[13]. Yet, with heterogeneous beliefs, policy makers need to only focus on the fraction of agents that see bubbles, monitoring market sentiment using the financial press, surveys, and other relevant borrowing and short sale interest data, and act when that fraction becomes too large. Because bubbles are affected by market liquidity and trading constraints, as detailed above, policy makers can affect the fraction of agents seeing bubbles by making markets either more liquid or by making trading constraints more restrictive. This indirect channel can be very effective in changing agents' beliefs regarding price bubbles, perhaps even more effective than interest rate monetary policy.

This chapter also relates to the macroeconomics literature studying the impact of financial frictions. As previously discussed, the classic papers in this area include Bernanke and Gertler (1989), Kiyotaki and Moore (1997), and Bernanke, Gertler, and Gilchrist (1999). Some recent papers include He and Krishnamurthy (2012, 2013, 2014) and Brunnermeier and Sannikov (2014). This chapter is closest to Brunnermeier and Sannikov (2014), He and Krishnamurthy (2014), and Lamichhane (2018) because these papers study systemic risk with a financial sector. This chapter is also related to the papers by Huggett (1993), Aiyagari (1994), Benhabib, Bisin and Zhu (2011, 2014), and Achdou et al. (2015). The model in this chapter has heterogeneous beliefs across agents, as opposed to

[12]This statement formally applies only if we restrict ourselves to equilibrium prices. In disequilibrium it is possible that all agents see the same uniform price bubble. This would occur, for example, in a complete and arbitrage-free market, see Jarrow (2015). Characterizing systemic risk in terms of the fraction of economic agents was first introduced in Lamichhane (2018).

[13]For example, see the discussion by the president of the Federal Reserve Bank of Minneapolis Neel Kashkari (Kashkari, 2017) on monetary policy and bubbles.

heterogeneous income or productivity shocks considered in these papers.

With respect to finance, this chapter relates to the literature studying an investor's optimal trading strategy with liquidity risk (see Cetin and Rogers (2007), Vath, Mnif, and Pham (2007), Chebbi and Soner (2013), and Pennanen (2014)). We use a similar formulation as contained in Pennanen (2014). In this literature, there are two papers studying dynamic Radner equilibrium. These are the overlapping generations model of Acharya and Pedersen (2005) and the discrete time model of Jarrow (2016b). Acharya and Pedersen (2005) include no short sales and liquidity risk is characterized by a fixed but stochastic transaction cost that is independent of trade size. Jarrow (2016b) includes a stochastic liquidity cost that depends on trade size, but there are no trading constraints. Standard transaction costs are a special case of our formulation (see Jarrow and Protter (2008) for a detailed explanation). Our notion of market liquidity is related to but more general than the market and funding liquidity as considered in Brunnermeier and Pedersen (2008).

CHAPTER 2

SYSTEMIC RISK AND FINANCIAL INNOVATION: A STATIONARY EQUILIBRIUM ANALYSIS

2.1 Model

In this section, we first describe our model which forms the basis for the analysis in first two chapters of this dissertation. Then we will re-caste the problem in a dynamic programming framework to efficiently solve the model. First we briefly summarize the overall model structure.

Borrowers consume and borrow from the lenders and face a certain borrowing/credit constraint. But they might default on their loan. Default is modeled by making borrowers' income follow a jump-diffusion process, where jump in income translates into defaults. Similarly, lenders also consume and save, and their savings takes the form of loans made out to the borrowers. Since borrowers might default, lenders can insure certain fraction of their loans through the financial sector. So, the parameter that captures the total fraction of loans insured is used to model the usage of financial innovation. Higher fraction means increasing use of financial innovation. The financial sector sells credit risk protection to the lenders, insuring them against the default risk of the borrowers. It receives regular premium from lenders, which forms its capital position, and it agrees to pay the lenders if borrowers default.

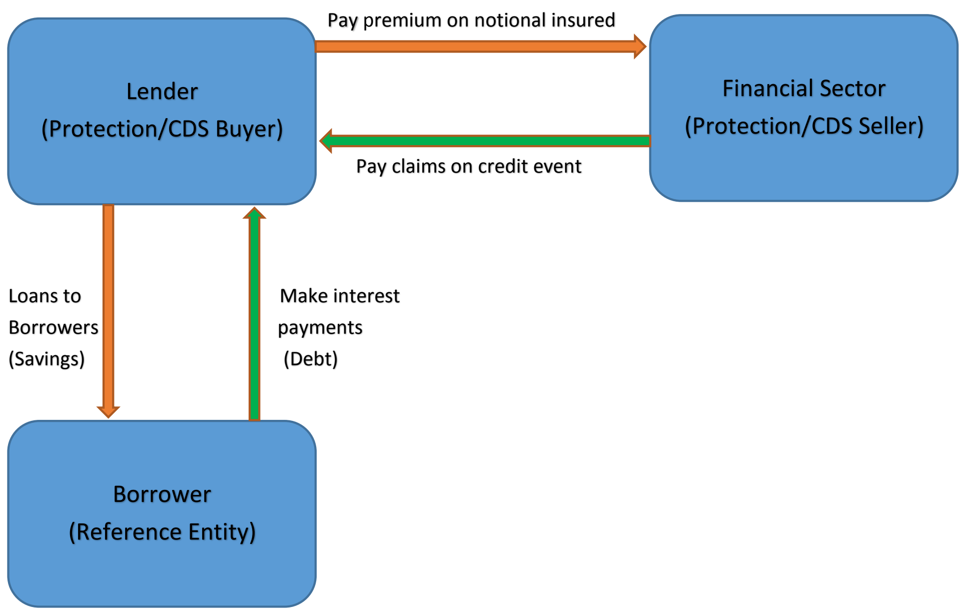


Figure 2.1: Visual Representation of the Model

2.1.1 Setup

There are three classes of agents. First, a continuum of heterogeneous lenders (saving households) of measure 1, indexed by l that are heterogeneous in their idiosyncratic income shocks. Second, a continuum of heterogeneous borrowers (borrowing households) of measure 1, indexed by b that are heterogeneous in their idiosyncratic income shocks. Third, a financial sector that sells credit risk protection to the lenders (hereafter, CDS contract). Thus, our model structure closely captures the CDS market structure which represents an important instance of financial innovation. This is visually represented in Figure (2.1). Also, as we will see later, our model generates various new testable predictions while closely replicating some features of the data.

Borrowers

Borrowers consume c_t^b, get stochastic income z_t^b and take on debt by borrowing from lenders. They have standard preferences:

$$V^b(a_t^b, z_t^b) = \max_{\{c_t^b, a_t^b\}} \quad \mathbb{E}_0 \int_0^\infty e^{-\rho^b t} u(c_t^b) dt \tag{2.1}$$

where u is strictly increasing , concave and ρ^b is the discount rate. Note $\rho^b > \rho^l$ i.e. borrowers are less patient than lenders, and thus borrow. Income follows a Lévy Jump-Diffusion type process given by:

$$dz_t^b = \mu(z_t)^b dt + \sigma(z_t)^b dW_t^b + J_t(z_{t-}^b) dN_t \tag{2.2}$$

where W_t^b is a standard Brownian motion, $\mu(z_t^b)$ is the drift, $\sigma(z_t^b)$ is the diffusion of income process, J_t is the idiosyncratic jump/credit risk where the jump magnitude which may be drawn from some common exogenous distribution with support on $\mathbb{R}_-$), i.e. negative jump. Jump captures credit risk in the economy. N_t is the Poisson process with intensity λ.

Note that jumps arrive at the same time for all borrowers, so that they all default. This can be thought of as being correlated defaults or, in other words, jump process is the same across borrowers but the continuous drift-diffusion process is different. This can also be interpreted as borrowers using a rule-of-thumb to default when their income jumps at a random time. As a penalty for default, borrowers' income is reduced by some fraction δ (the borrowing limit in the economy can also be reduced at the same time)[1]. So a jump of magnitude J_t takes the income from z_{t-} (value of the process *before the jump*) to a new and

[1]This is a mild assumption because considering endogenous default decision complicates our model and does not yield additional insights for the type of questions this chapter is after. Given the explicit reduction in future income after the surprise jump in the income, this exogenous default assumption is a simple way to get around the complications of endogenous default decision, which is an important topic in and of itself.

lower value z_t'. For theoretical tractability (and also for computational simplicity later on), we consider the generalized Ornstein-Uhlenbeck (G-OU) process. This is just a generalization of classical OU process, constructed by replacing the Brownian motion component with a Lévy process. So, for a $\beta > 0$ and some initial value $z_0^b = z$, the general form of such a G-OU process is given by:

$$dz_t^b = -\beta z_t^b dt + d\hat{z}_t \tag{2.3}$$

where $\hat{z}_t$ is a spectrally negative Lévy process[2]. Income z_t^b evolves stochastically and we restrict the domain to $\mathbb{R}_+$ as zero is the economically meaningful lower bound on income[3].

Let τ denote the random arrival time of jump, i.e. when borrowers' default. This is a stopping time generated by the point process $N_t = 1$ adapted to the filtration $\mathcal{F}_t$, with $P(dN_t = 1) = \lambda dt$ faced by all borrowers. This means in an infinitesimal time dt the probability of default by borrowers is λdt. λ can also be interpreted as the probability of default over a small time interval Δt. Such a time is totally *inaccessible*. This means default comes as a surprise, as opposed to the case with just diffusions where the default is predictable, which removes the surprise element. While a stochastic process for intensity λ could be specified in Lando (1998) (say a square-root process), we abstract away from such complications given our goals and fix the jump intensity. Nevertheless, this can

[2]We know that the limiting distribution of such a G-OU process exists i.e. $z_t^b \to \int_0^\infty e^{-\beta t} d\hat{Z}_s$ as $t \to \infty$, with the Laplace transform given by $exp(\int_0^\infty \psi(e^{-\beta s}\theta)ds)$ where $\psi(\theta)$ is the Laplace exponent of $\hat{z}_s$. z_t^b is also called a Feller process, which is also a semimartingale.

[3]We could also assume that z_t^b evolves stochastically in a bounded interval $[\underline{z}^b, \bar{z}^b]$, $\bar{z}^b \geq 0$, such that the process either stays in the interval or is reflected at the boundaries. Further, as noted in Achodou et. al. (2015), theoretically it is not necessary to restrict the process in a bounded interval. Unbounded processes present no additional challenge. We would just assume that the unbounded process is ergodic so as to get the existence of a stationary equilibrium. Rather, this assumption is for practical reasons of having to perform numerical computations in a bounded interval. Xing, Zhang and Wang (2009) also proved the existence of a stationary distribution for similar jump reflected OU processes, and also provided the Laplace transform of such a distribution.

add another layer of richness (and complication) to the model that could be considered for future research.

Payment of loan takes place every instant and the law of default arrival time is given by the exponential distribution: $P(\tau \leq t) = 1 - e^{-\lambda t}$. Debt $a_t^b \leq 0$ takes the form of loan taken from lenders at interest rate r_t, to be determined in equilibrium, and evolves as:

$$\dot{a}_t^b = \frac{da_t^b}{dt} = (z_t^b + r_t a_t^b - c_t^b)\mathbb{1}_{\{\tau > t\}} + (z_t^b - \delta z_t^b + r_t a_t^b - c_t^b)\mathbb{1}_{\{\tau \leq t\}} \tag{2.4}$$

or equivalently,

$$\dot{a}_t^b = (z_t^b + r_t a_t^b - c_t^b)(1 - dN_t) + (z_t^b - \delta z_t^b + r_t a_t^b - c_t^b)dN_t$$

where δz_t^b represents the reduction in income upon default[4]. Borrowers also face a borrowing constraint as:

$$a_t^b \geq \underline{a}^b \tag{2.5}$$

where $\underline{a}^b < 0$ is the borrowing limit. This can also be explicitly specified as $\underline{a}^b = -\phi \tilde{z}^b$ if we want borrowing to depend on some fraction of the income, capturing the notion of collateral, i.e. for some income $\tilde{z}^b$, ϕ is the leverage ratio. This implicitly defines a lower bound on the borrowing. Since our economy is driven by Markov processes (memoryless property), after defaults occur some adjustments take place, i.e. fraction δ reduction in income and possibly a reduced borrowing limit. This will be reflected in the new equilibrium interest rate for times $t \geq \tau$. Other than these instantaneous adjustments (due to continuous time and Markov nature), the economy continues as before. This also means after default borrowers can still participate in the debt market, i.e. no

[4]This reduction in income captures the notion of a bankruptcy cost associated with default, analogous to the motivation for capturing the features of U.S. consumer bankruptcy law in Chaterjee et al. (2007) where households smooth their consumption with a riskless asset and unsecured credit, with the option to default. In our model, default comes as a total surprise.

autarky, albeit with some adjustment. This is a type of *limited liability* feature of the model which is behind some results to be discussed later.

Lenders

Lenders consume c_t^l, get stochastic income z_t^l each period and save. Their savings take the form of loans/credit made out to the borrowers. Financial wealth results from the saving/lending decisions of the lenders. We will use the words loan, saving or credit interchangeably throughout depending on the context, as they all mean the same in this economy. Thus, lenders are the only credit suppliers in the economy. They have standard preference:

$$V^l(a_t^l, z_t^l) = \max_{\{c_t^l, a_t^l\}} \quad \mathbb{E}_0 \int_0^\infty e^{-\rho^l t} u(c_t^l) dt \tag{2.6}$$

Their income follows a simple diffusion process:

$$dz_t^l = \mu(z)^l dt + \sigma(z)^l dW_t^l \tag{2.7}$$

where W_t^l is a standard Brownian motion, $\mu(z_t^l)$ is the drift and $\sigma(z_t^l)$ is the diffusion of income process[5]. Financial wealth $a_t^l \geq 0$ takes the form of loans made out to borrowers at interest rate r_t, determined in equilibrium. Thus, a_t^l evolves as:

$$\dot{a}_t^l = \frac{da_t^l}{dt} = (z_t^l + r_t a_t^l - s_t(\alpha_t a_t^l) - c_t^l)\mathbb{1}_{\{\tau > t\}} + (z_t^l + r_t a_t^l + L_t - s_t(\alpha_t a_t^l) - c_t^l)\mathbb{1}_{\{\tau \leq t\}} \tag{2.8}$$

or equivalently,

$$\dot{a}_t^l = (z_t^l + r_t a_t^l - s_t(\alpha_t a_t^l) - c_t^l)(1 - dN_t) + (z_t^l + r_t a_t^l + L_t - s_t(\alpha_t a_t^l) - c_t^l) dN_t$$

[5]Similar to the borrowers' problem discussed above, for computational purpose we can specify lenders' income to evolve stochastically on a bounded interval $[\underline{z}^l, \bar{z}^l]$, $\underline{z}^l > 0$, according to a stationary diffusion process (such as the standard OU process) that either stays in the interval or is reflected at the boundaries.

where $\alpha \in [0,1]$ and αa_t^l is the fraction of total loan notional a_t^l insured by purchasing a CDS contract from the financial sector. The lender (the protection buyer) pays CDS premium rate s_t on the fraction of notional insured $\alpha_t a_t^l$. Upon credit event, the financial sector (the protection seller) pays the lenders net amount L_t. For simplicity, we assume zero recovery of the defaulted debt[6].

Note that we introduce financial innovation exogenously, i.e. α is not a choice variable, because we are interested in understanding the systemic risk implications *given* that financial innovation exists (and is used). So, the parameter α that captures the total fraction of loans insured, is used to model the usage of financial innovation. Higher fraction α means increasing use of financial innovation, i.e. α captures the degree of financial innovation. Thus, by varying α, we can trace out various degrees of systemic risk implications associated with increasing use of financial innovation. And since α is exogenous, rate s_t will also be exogenous (discussed in financial sector's problem below).

Making α a choice variable that is endogenously determined in equilibrium would be too restrictive for our purposes, while also unnecessarily complicating the model. This endogenous choice would transform the model from the present partial equilibrium framework to a general equilibrium framework. We would then have to first address the issue of the endogenous need for financial innovation itself, i.e. how lenders would choose to allocate between amounts to save/lend and amounts to insure. However, going this path distracts from our main goal of understanding systemic risk implications *conditional* on the fact that lenders have access to and use financial innovation. This path also does not

[6]Though including positive recovery rate/mechanisms is possible in our model (eg. Madan and Unal (1998) and Duffie and Singleton (1999) consider recovery of market value), we abstract away from these. Understanding recovery mechanisms in itself is an interesting issue, and we refer the readers to the review paper by Jarrow (2009) for details and references.

help us gain insight for the types of questions we want to address. Hence, we abstract away from the endogenous determination of α (and thus the CDS rate s_t as well). This means, as discussed above, we can freely vary α to trace out systemic risk implications. We can perform various comparative statics exercises to understand the relation between systemic risk and financial innovation via other equilibrium objects, such as the leverage and financial wealth distributions across agents.

Finally, we note that our partial equilibrium model can be extended to a general equilibrium framework by making α a choice variable. This extension would allow for studying *counterparty risk pricing in dynamic general equilibrium*, which is another open area of research in macro-finance. The main technical challenge is that a general equilibrium does not exist in our current model setup, (especially in light of financial sector's problem that we discuss below). Even though it is quite possible to address this issue with some minor modifications in the current model structure, this would naturally lead to a different set of implications that we do not study in this chapter.

Financial Sector

The financial sector sells credit risk protection to the lenders, insuring them against the default risk of the borrowers. It receives regular premium from lenders, which forms its capital position, and it agrees to pay the lenders if borrowers default. This financial innovation role of the financial sector is motivated by the CDS market structure. Defaults arrive according to intensity λ, as described above in borrowers' problem. Thus, the financial sector's evolution of surplus-capital X_t is given by:

$$dX_t = (\gamma X_t + \xi_t)dt - dY_t, \qquad X_0 = x \geq 0 \tag{2.9}$$

where $\xi_t = (s_t M_t)$ is the total premium received per-unit time. $X_0 = x$ is the initial capital. Y_t is the aggregate claims whose dynamics is given by a driftless subordinator defined on our filtered probability space that charges only the positive real line $\mathbb{R}_+$. Thus, X_t is a spectrally negative type of Lévy process (see Theory of Lévy Processes for a review of the basics of the theory of Lévy process). The financial sector stands ready at each time t to insure the total notional amount M_t chosen by lenders, at a premium s_t that we will now call the *CDS rate*.

As discussed above in lenders' problem, we take this CDS rate s_t as given (possibly priced via some reduced form methods as in Jarrow (2009, 2011)) given our goals. Requiring s_t to be determined in equilibrium is equivalent to transforming our model to a general equilibrium framework. However, explicitly considering the counterparty risk can be complicated, even in a simple reduced form setting. This complexity is magnified in a dynamic equilibrium model such as ours[7]. Further, this would also take away some of the analytical tractability of our model, especially given that a general equilibrium does not exist (only a partial equilibrium exists with equilibrium interest rate r_t) with the financial sector's surplus-capital evolution process X_t as specified above.

Thus, we abstract away from the endogenous role of financial innovation for reasons discussed above. Rather, the main concern in this chapter is whether the financial sector can remain solvent at all times because that matters for the financial stability and the ensuing systemic risk implications. Studying counter-

[7]Jarrow and Turbull (1995, 1997) and Duffie and Huang (1996) are the first notable papers to study counterparty risk in a reduced form model. Jarrow and Yu (2001) also analyze a reduced form model to include default intensities dependent on the default of a counterparty.

party risk in a dynamic general equilibrium model wherein the need for financial innovation arises endogenously is an interesting topic on its own for future research that can shed light on why financial innovation even exists.

A simple observation suggests that the problem of the financial sector is not entirely in a reduced form. Claims are drawn from the *equilibrium leverage distribution* across borrowers (reflected about the y-axis to the positive half line for convenience). So, the dY_t component captures the equilibrium borrowing decisions of the borrowers. The ξ_t component captures the equilibrium lending decisions of the lenders. Thus, even though the financial sector and borrowers are not directly related, both are indeed inextricably linked via the decision of the lenders. This highlights the subtle and rich structure of the economy in our model that allows us to study various properties of systemic risk. This would not have been possible with endogenously determined α or s_t.

We can view this CDS rate s_t as being *fair* in the sense that it is the expected total loss amount from claims – the failure probability times the loss rate. This is the expected claim amount per-unit time: $s_t = \int y\Pi(dy) = \lambda\bar{L}$ where $\Pi(dy) = \lambda K(dy)$ is the Lévy measure, for some claims distribution K on $\mathbb{R}^+$ and $\bar{L}$ is the total loss when default occurs. Thus, the financial sector in our model acts like a *risk-neutral counterparty* to all the risk in the economy. It stands ready to assume the risk lenders want to insure against at a risk-neutral rate[8].

The parameter $\gamma \geq 0$ helps us explore some general cases. When capital-surplus X_t is positive, the financial sector can earn credit interest at some rate.

[8]In a reduced form setting and in practice, pricing of a CDS contract is analogous to determination of premium for an insurance contract. But the actual pricing can get quite complicated even in a reduced form model. Nevertheless, the key variables involved are largely similar to those in our model. See Jarrow (2011, 2009) for a derivation of the market clearing CDS rate in a reduced form setting.

When the surplus falls below zero, the financial sector can cover the deficit by borrowing at a debit interest at some rate. In general, the credit rate is lower than the debit rate. Here we assume one single rate γ for analytical tractability, implying the same credit and debit rates[9]. We will later see γ's stabilizing role – it can allow the capital position to go into negative territory while still allowing the financial sector to remain solvent. This can also be interpreted as monetary policy intervention during times of economic distress by affecting the short term rate γ.

We note one subtle feature of our model. The risk-neutral, price-setting financial sector fully knows the leverage distribution and the intensity λ, i.e. its a full information economy. Nevertheless, we can add a *behavioral friction*, i.e. the financial sector underestimates the default risk λ, under this risk neutral pricing. This means the financial sector could use some other intensity $\bar{\lambda} < \lambda$ for pricing the premium, implying that it under-prices the premium, i.e. *underestimates risk*, while still being a risk-neutral, price-setting agent. As discussed before, this is similar to the notion of neglected risks as in Gennaioli et al. (2012) that leads to over-issuance of securities. In our model, this additional friction further increases the failure probability of the financial sector, naturally making the results stronger. But the results are already already stronger under a weaker set of conditions not involving such frictions. Thus, this ease of adding other frictions underscores the flexibility of our model.

Another subtle concern is this: in a full information economy, why won't

[9]We can easily motivate use of γ. When the surplus is positive, the financial sector can deposit those funds with the central bank and earn interest rate. When surplus is negative, it can draw or borrow funds from the central bank to cover its deficits. So, γ is analogous to the fed funds rate or the overnight LIBOR rates: financial institutions facing capital shortfall can borrow from other institutions that have excess capital, at the overnight rate γ to meet the regulatory requirements.

the risk-neutral financial sector adjust the price to account for the fully known risk profile in the economy and save itself from failure at all times? This is actually not a proper question in our context. In any full information asset pricing model, the market prices all risks properly. However, when the risk event occurs, there is always the chance of a market crash. This is fully known (i.e. the systematic risk component in asset pricing literature) and it cannot be hedged away. Thus, pricing the risk *does not* eliminate the possibility of a risk event; pricing just accounts for such risks in agents' state price density. In this chapter, we show how introducing financial innovation (like a CDS market) in risk/default settings does not necessarily remove all risks, but rather can add to the overall risk. This implies that there is still systemic risk.

2.1.2 Equilibrium:

Let $g^b(a^b, z^b)$ be joint distribution of the states a^b and z^b. Define aggregate debt/leverage (negative value here) of borrowers by:

$$\Gamma^b(r_t) = \int_{z^b} \int_{\underline{a}^b}^{0} a^b g^b(a_t^b, z_t^b) da^b dz^b$$

Let $g^l(a^l, z^l)$ be another joint distribution of the states a^l and z^l. Define aggregate savings (also financial wealth) of the lenders by:

$$\Gamma^l(r_t) = \int_{z^l} \int_{0}^{\infty} a^l g^l(a_t^l, z_t^l) da^l dz^l$$

Define aggregate loans/credit insured by the lenders via CDS as:

$$\Gamma^{cds}(s_t) = \int_{z^l} \int_{0}^{\alpha a^l} (\alpha a^l) g^l(a_t^l, z_t^l) da^l dz^l$$

Then, the equilibrium is defined as the price (i.e. the interest rate r_t in this economy), such that savings/loans are in zero net supply:

$$\Gamma^l(r_t) + \Gamma^b(r_t) = 0 \tag{2.10}$$

Further, the financial sector insures whatever total notional M_t is exogenously chosen by the lenders, at the CDS rate s_t. So, as the loan market clears, the CDS, i.e. the financial market clears as well:

$$\Gamma^{cds}(s_t) = M_t \tag{2.11}$$

2.1.3 Dynamic Programming Formulation

Solving the above optimization problem is very convenient and more flexible with a dynamic programming approach than with an optimal control approach. One particularly important advantage is the ease with which borrowing constraints can be handled in the continuous time dynamic programming formulation. Such constraints become state constraints (Soner, 1986a,b, Capuzzo-Dolcetta and Lions, 1990). In stark contrast to discrete time formulations, these constraints never bind in the interior of the state space and only show up in the boundary conditions, giving rise to state constraint boundary condition (Achdou et al., 2015). This considerably simplifies the analysis as the entire economic dynamics can be described by a system of two non-linear partial differential equations (PDE): a Hamilton-Jacobi-Bellman equation and a Kolmogorov Forward equation. We now re-cast the model in a very compact and efficient way using forward and backward operators. In appendix Operators, we briefly discuss some basics of operators used in this chapter. This operator-based approach can allow for a great deal of computational efficiency.

Dynamics of Lenders: The Hamilton-Jacobi-Bellman (HJB) and Kolmogorov Forward (KF) equations for lenders are, respectively:

$$\rho^l V^l(a_t^l, z_t^l) = \max u(c^l) + \mathcal{L}V^l(a_t, z_t) + \partial_t V^l(a^l, z^l, t) \tag{2.12}$$

$$\partial_t g^l(a_t^l, z_t^l) = \mathcal{L}^* g^l(a_t, z_t) \tag{2.13}$$

where $\mathcal{L}V^l(z_t, a_t)$ is the backward operator[10] defined as:

$$\begin{aligned}\mathcal{L}V^l(a_t, z_t) = {} & \partial_a V^h(a^l, z^l, t)(\eta^l(a_t^l, z_t^l)) + \partial_z V^l(a_t^l, z_t^l)\mu(z^l) \\ & + \frac{1}{2}\partial_{zz} V^l(a_t^l, z_t^l)\sigma^2(z^l)\end{aligned}$$

and $\mathcal{L}^* g^l(a_t, z_t)$ is the forward operator, defined as:

$$\begin{aligned}\mathcal{L}^* g^l(a_t, z_t) = {} & -\partial_a[\eta^l(a_t^l, z_t^l) g^l(a_t^l, z_t^l)] - \partial_z[\mu(z^l) g^l(a_t^l, z_t^l)] \\ & + \frac{1}{2}\partial_{zz}[g^l(a_t^l, z_t^l)\sigma^2(z^l)]\end{aligned}$$

Also,

$$\begin{aligned}\eta^l(a_t^l, z_t^l) = {} & (z_t^l + r_t a_t^l - s_t(\alpha_t a_t^l) - c_t^l)\mathbb{1}_{\{\tau > t\}} \\ & + (z_t^l + r_t a_t^l + L_t - s_t(\alpha_t a_t^l) - c_t^l)\mathbb{1}_{\{\tau \leq t\}}\end{aligned} \tag{2.14}$$

is the optimally chosen drift of the financial wealth (i.e. optimal savings policy function). $c^l(a_t^l, z_t^l)$ is given by the first order condition:

$$u'(c^l(a_t^l, z_t^l)) = \partial_a V^l(a_t^l, z_t^l) \tag{2.15}$$

[10]Note that $\partial_t V^l(a^l, z^l, t) + \mathcal{L}V^l(z_t, a_t) - \rho^l V^l(a^l, z^l, t)$ is the analogue of Black-Scholes-Merton (BSM) PDE, also known as the *Feynman-Kac formula*. Then $\partial_t V^l(a_t^l, z_t^l) + \mathcal{L}V^l(a_t, z_t) = 0$ is the usual Kolmogorov *backward* equation where $\mathcal{L}V^l(z_t, a_t)$ is the backward operator. The discount rate ρ^l is analogous to the risk-free rate in BSM framework. If we were in a finite horizon where the final random payoff function were defined, then the value function would be the expected discounted payoff that solved this PDE. Thus, we can see a clear connection between the finance literature, especially that involving stochastic calculus methods, and how that literature's knowledge can be naturally adapted and used in equilibrium settings commonly explored in macro literature.

The boundary and terminal conditions below are mostly useful for computational purposes as we need to work on a bounded domain for computing the equilibrium i.e $z^l \in [\underline{z}^l, \bar{z}^l]$ for $\underline{z}^l \geq 0$. If we assume that income z^l is reflected at boundaries, the value function V^l satisfies the boundary conditions:

$$\partial_z V^l(a_t^l, \underline{z}_t^l) = 0, \qquad \partial_z V^l(a_t^l, \bar{z}_t^l) = 0, \text{ for all } a^l \tag{2.16}$$

V^l also satisfies a terminal condition, for a *large enough* T:

$$V^l(a^l, z^l, T) = V_\infty(a^l, z^l) \qquad (\approx \text{stationary value function}) \tag{2.17}$$

Finally, the density g^l satisfies the initial condition:

$$g^h(a^l, z^l, 0) = g_0^l(a^l, z^l) \tag{2.18}$$

Dynamics of Borrowers: Again, the HJB and KF equations are, respectively:

$$\rho^b V^b(a_t^b, z_t^b) = \max u(c^b) + \mathcal{L}V^b(a_t, z_t) + \partial_t V^b(a^b, z^b, t) \tag{2.19}$$

$$\partial_t g^b(a_t^b, z_t^b) = \mathcal{L}^* g^b(a_t, z_t) \tag{2.20}$$

where $\mathcal{L}V^b(a_t, z_t)$ is the backward operator[11] defined as:

$$\begin{aligned}\mathcal{L}V^b(a_t, z_t) &= \partial_a V^b(a_t, z_t)(\eta^b(a_t^b, z_t^b)) + \partial_z V^b(a_t, z_t)\mu(z^l) \\ &\quad + \frac{1}{2}\partial_{zz} V^b(a_t, z_t)\sigma^2(z^l) + \lambda \mathbb{E}_J[V^b(a^b, z_J^b) - V(a^b, z^b)]\end{aligned}$$

[11]Note that $\partial_t V^b(a_t, z_t) + \mathcal{L}V^b(a_t, z_t) - \rho^b V^b(a_t^b, z_t^b)$ is sometimes called the *partial integro-differential equation* (PIDE). This includes the BSM PDE (the differential part) with an added jump component (the integral part). The added operator relative to lenders' problem, $\lambda \mathbb{E}_J[V^b(a^b, z_J^b) - V(a^b, z^b)]$, represents the effect of finite jumps generated by Poisson process N_t with mean arrival intensity λ where jumps are independent of income z^b.

and $\mathcal{L}^* g^b(a_t, z_t)$ is the forward operator, defined as:

$$\begin{aligned}\mathcal{L}^* g^b(a_t, z_t) = &-\partial_a[\eta^b(a^b, z^b, t) g^b(a_t^b, z_t^b)] - \partial_z[\mu(z^b) g^b(a_t^b, z_t^b)] \\ &+ \frac{1}{2}\partial_{zz}[g^b(a_t^b, z_t^b)\sigma^2(z^b)] + \lambda \mathbb{E}_J[g(a^b, z_j^b) - g(a^b, z^b)]\end{aligned}$$

Also,

$$\begin{aligned}\eta^b(a_t^b, z_t^b) = &(z_t^b + r_t a_t^b - c_t^b)\mathbb{1}_{\{\tau > t\}} \\ &+ (z_t^b - \delta z_t^b + r_t a_t^b - c_t^b)\mathbb{1}_{\{\tau \le t\}}\end{aligned} \tag{2.21}$$

is the optimally chosen drift of the debt (i.e. debt policy function).

$c^b(a_t^b, z_t^b)$ is given by the first order condition:

$$u'(c^b(a_t^b, z_t^b)) = \partial_a V^b(a_t^b, z_t^b) \tag{2.22}$$

Borrowing constraint never binds in the interior of the state space and only shows up in the HJB equation's boundary conditions. This ensures borrowing limit is never violated, and thus gives rise to state constraint boundary condition at $a^b = \underline{a}^b$ given by:

$$u'(z_t^b + r_t \underline{a}^b) \le \partial_a V^b(\underline{a}_t^b, z_t^b), \quad \text{for all } z^b \tag{2.23}$$

The value function V^b satisfies the terminal condition for a *large enough* T and g^b satisfies the initial condition:

$$V^b(a^b, z^b, T) = V_\infty(a^b, z^b) \tag{2.24}$$

$$g^b(a^b, z^b, 0) = g_0^b(a^b, z^b) \tag{2.25}$$

Thus, above equations (2.12) to (2.25) that include HJB and KF equations, optimality, boundary, terminal and initial conditions, along with conditions (2.10) and (2.11), fully describe this economy[12].

[12]To specify reflecting boundary we need boundary conditions. But defining reflecting barri-

2.1.4 Stationary Equilibrium

We can easily transform the above dynamic programming equations into their stationary counterparts as follows:

Lenders: For lenders, the stationary counterparts to their respective dynamic equations above are:

$$\rho^l V^l(a^l, z^l) = \max u(c^l) + \mathcal{L}V^l(a, z) \tag{2.26}$$

$$0 = \mathcal{L}^* g^l(a, z) \tag{2.27}$$

$$\eta^l(a^h, z^h) = \text{as above} = \dot{a}^l \tag{2.28}$$

$$u'(c^l(a^l, z^l)) = \partial_a V^l(a^l, z^l) \tag{2.29}$$

Borrowers: Similarly for borrowers:

$$\rho^b V^b(a^b, z^b) = \max u(c^b) + \mathcal{L}V^b(a, z) \tag{2.30}$$

$$0 = \mathcal{L}^* g^b(a, z) \tag{2.31}$$

$$\eta^b(a^b, z^b) = \text{as above} = \dot{a}^b \tag{2.32}$$

$$u'(c^b(a^b, z^b)) = \partial_a V^b(a^b, z^b) \tag{2.33}$$

$$u'(z^b + r\underline{a}^b) \leq \partial_a V^b(\underline{a}^b, z^b), \qquad \text{for all } z^b \tag{2.34}$$

---

ers with jumps can get complicated. This is a theoretical issue, but this issue might become important for computational purposes. Economically meaningful conditions would call for setting the derivative of value function zero at those boundaries, just as in lenders' boundary problem. One simple solution could be to pick a fixed jump value and use indicator functions to ensure that we do not overshoot or undershoot the boundary to make the income remain in the bounded interval. The method for solving BSM PDE with appropriate boundary conditions can be generalized to models involving jumps by considering PIDE featured in the borrowers' problem. One way is to use finite difference scheme commonly used in pricing barrier options. The value function has to be specified at the boundary, i.e. the barrier and beyond the boundary. This complication arises if we want to solve the fully general case of the PIDE. See Cont & Tankov (2004) and Cont & Voltchkova (2005) for more details on solving such problems. We show the generalized version in the model for completeness, but resort to simplified setting, by reducing the PIDE to PDE, for computational simplicity and to emphasize economic intuition.

Here $g^l(a,z)$ and $g^b(a,z)$ denote stationary distributions. Finally, the stationary equilibrium interest rate r must satisfy conditions analogous to equation (2.10).

$$\Gamma^l(r) + \Gamma^b(r) = 0 \tag{2.35}$$

For the financial sector, condition analogous to equation (2.11) holds:

$$\Gamma^{cds}(s) = M \tag{2.36}$$

Thus, equations (2.26) to (2.36) that include HJB and KF equations along with optimality, boundary, and equilibrium conditions fully characterize the stationary equilibrium of this economy.

2.2 Results

In this section, we will a systematic treatment of systemic risk in stationary equilibrium[13]. We first give a precise and rigorous definition of systemic risk. Then we will derive a series of results to understand systemic risk with respect to leverage and financial wealth distributions, risk aversion, use of financial innovation (CDS purchases) and default intensity. We will then combine these series of results to get to one of our main results, i.e. financial innovation can either increase or decrease systemic risk, *under some conditions*. Transition dynamics part will be analyzed in the next chapter where many of the insights developed in this stationary equilibrium chapter will continue to hold.

Proposition 1. ***(Existence of Stationary Equilibrium)****: There exists a stationary equilibrium in our economy.*

[13]In Numerical Computation we have outlined the details of the numerical computation, which largely relies on the finite-difference method as used in Achdou et al. (2015), based on the work by Achdou, Camilli and Capuzzo-Dolcetta (2012).

2.2.1 Defining Systemic Risk

Leverage distribution is more useful in studying systemic risk than just the average leverage, because average might not be representative of the actual risk profile of the economy. Stiglitz (2012) argues that economists who looked at just the average equity of homeowners before the 2007-2009 financial crisis, ignoring the distribution, were confident that the economy could easily weather a housing price decline of 20% or more, given that average indebtedness was less than 80% of the market price. However, such an approach was misguided because it ignored the fact that a *large fraction* of homeowners would owe more than their home value. This dire proposition ultimately materialized, leading to widespread foreclosures and the downturn of the U.S. economy[14].

In our model, the *endogenously* determined (marginal) distribution of leverage, especially the measure around the tail, gives the first clear notion of systemic risk. Leverage distribution can be easily motivated as an important risk factor as it is associated with the scenario where after some exogenous shock, households need to massively and rapidly deleverage. This can lead to a sharp decline in consumption and spending. Even in the absence of the financial sector, this would lead the economy into a downward spiral of demand-driven slump. Once the financial sector is explicitly incorporated, the insolvency considerations of the financial sector, the counterparty to all risks in the economy, gives another economically meaningful channel to characterize systemic risk. This motivates our general definition of systemic risk below:

Definition 1. ***(Systemic Risk):*** *Let* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ *denote the filtered probability space*

[14] This example also indicates that the representative agent framework may not be well suited to answering questions regarding systemic risk because all distributional considerations are assumed away. However, a heterogeneous agents framework easily allows us to consider the entire leverage distribution (as well as financial wealth distribution).

characterizing randomness in the economy where $\mathbb{P}$ is the statistical probability measure, $\mathbb{F} = \{\mathcal{F}\}_{t\geq 0}$ is the filtration of the σ- field $\mathcal{F}$.

Let $\underline{\mathcal{B}}$ denote a Borel set around $\underline{a}^b$ such that $\underline{\mathcal{B}} \subset \Omega$ and $\underline{a}^b \in \underline{\mathcal{B}}$. Define all the states $a^b \in \underline{\mathcal{B}}$ as systemic risk states. Let $\int_{\underline{\mathcal{B}}} d\mathbb{P}(\omega) = \bar{\mathbb{P}}^b$ denote the left-tail of (marginal) leverage distribution across borrowers. Then we get two closely related notions of systemic risk:

(1). We say that our economy exhibits systemic risk if $\bar{\mathbb{P}}^b \geq \hat{\mathbb{P}}$ for some positive $\hat{\mathbb{P}}$ (i.e. a heavy tail in the systemic risk states), possibly with Dirac measure or spike at or near $\underline{a}^b$. Increasing systemic risk implies increasing $\bar{\mathbb{P}}^b$.

(2). With the financial sector, systemic risk is characterized in term of the financial sector's failure probability $\pi(x)$, the distribution of first passage time $\mathcal{P}(\tau_{\bar{x}} \leq t)$ of capital below some solvency threshold level $\bar{x}$, and total capital available $\mathcal{X}_\tau$ at the random (stopping) jump time τ. Increasing $\pi(x)$ or $\mathcal{P}(\tau_{\bar{x}} \leq t)$ and decreasing $\mathcal{X}_\tau$ imply increasing systemic risk.

Note that systemic risk is defined in terms of probability, as it should be. Realization of systemic risk is crisis. Consequently, increased systemic risk in our model is first captured by an increased probability measure around the systemic risk states. Given the heterogeneous agents model, the measure around the systemic risk states is the total fraction of borrowers around the systemic risk states. Thus, increasing systemic risk would first imply an increasing measure of borrowers concentrated around the systemic risk states[15]. When a credit event arrives, i.e. negative jump in income, borrowers default.

[15]For practical purposes, the threshold level $\hat{\mathbb{P}}$ can be computed using historical events when the economy experienced financial crises. Assuming one can get access to such data or a good proxy (eg., through credit unions), one would look at the fraction of total borrowers at and around their borrowing/credit limit. If the sample is representative of the economy, one could construct the threshold fraction $\hat{\mathbb{P}}$.

Further, the jump size does not need to be large for the systemic risk effects as per our definition. Even with a small jump, i.e. small shock, there can be systemic risk implications. This is why systemic risk is defined via the tail of the equilibrium leverage distribution. Even with small jumps systemic risk implications can arise if there is a large fraction of borrowers around the high leverage systemic risk states, i.e. a heavy tail in the leverage distribution. Thus, we have reduced the problem of understanding a plethora of systemic risk issues to analyzing the conditions in the economy that affect the tail measure $\bar{\mathbb{P}}^b$ across borrowers. Interestingly, when the financial sector is explicitly considered, we get a second layer of systemic risk due to the failure possibility $\pi(x)$ of the financial sector. This systemic risk channel is related to $\bar{\mathbb{P}}^b$. Thus, it is not necessarily the size of negative shock (either jump size or the reduction in credit limit or both) that drive the results, but rather the size of the tail measure $\bar{\mathbb{P}}^b$ that serves to amplify the effects of small shocks, yielding systemic risk implications. Further, given our general model structure, we can also consider jump-size distribution across borrowers, which can yield other sets of insights. But given our goals, we abstract away from such generalization.

2.2.2 Baseline Systemic Risk Properties

Define $\mathcal{A}^b = -\lim_{a^b \to \underline{a}^b} \frac{u''(\underline{z}^b + ra^b)}{u'(\underline{z}^b + ra^b)}$. Here $\mathcal{A}^b$ represents the coefficient of absolute risk aversion $(-u''(c)/u'(c))$ when debt approaches the borrowing limit $\underline{a}^b$, capturing the risk aversion of households around the constraint[16]. Let $\underline{c}^b = c^b(\underline{a}^b, \underline{z}^b)$

[16]Note that some well-known features of Aiyagari-Huggett economies with uninsured idiosyncratic income risk easily generalize to our setting as well. For example, the stationary interest rate is smaller than the rate of time preferences of lenders who supply loans i.e. $r < \rho^l$. Also, theorems (1) and (2) in this chapter are closely related to similar results in Achdou et al.(2015) in terms of derivation methods. However, the model and the economic structure, ev-

and define another object that will be helpful for characterizing the speed and other related notions:

$$\mathcal{S}(z^b) = \sqrt{\left(2\frac{(r-\rho^b)u'(\underline{c}^b) + \partial_z u'(\underline{c}^b)\mu(z^b) + (1/2)\partial_{zz}u'(\underline{c}^b)\sigma^2(z^b) + \lambda(u'(\underline{c}(z')^b) - u'(\underline{c}^b))}{u''(\underline{c}^b)}\right)^+} > 0$$

Theorem 1. ***(Speed and Time towards Systemic Risk)*** *Consider the solution to the optimization problem of borrowers with $\mathcal{A}^b < \infty$. There is an z_c^b such that borrowers with income $z^b \leq z_c^b$ and debt $a^b = \underline{a}^b$ are constrained with $\eta(\underline{a}^b, z^b) = 0$ and borrowers with income $z^b > z_c^b$ are unconstrained.*

(a). The borrowing behavior and marginal propensity to consume near $\underline{a}^b$ is:

$$\eta(a^b, z^b) \approx -\mathcal{S}(z^b)\sqrt{a^b - \underline{a}^b},$$
$$\partial_a c(a^b, z^b) \approx r + \frac{\mathcal{S}(z^b)}{2\sqrt{a^b - \underline{a}^b}}$$

This implies, as $a^b \to \underline{a}^b$, $\partial_a \eta(a^b, z^b) \to -\infty$ and $\partial_a c(a^b, z^b) \to \infty$.

(b). Given $\eta(a^b, z^b)$ as above, borrowers with any initial debt $a_0^b > \underline{a}^b$ and continued low income realizations $z^b \leq z_c^b$ converge to $\underline{a}^b$- systemic risk state- in finite time $\mathcal{T}^b$, where $\mathcal{S}(z^b)$ determines the speed of convergence, which is proportional ($\propto$) to model elements as:

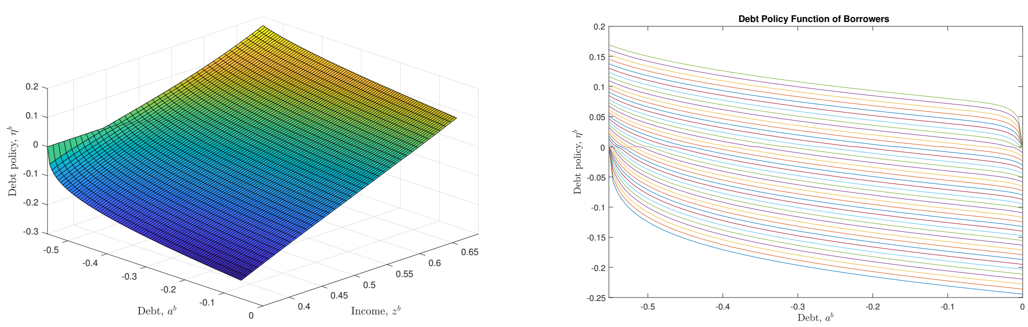
$$\mathcal{S}(z^b) \propto \lambda, \qquad \mathcal{S}(z^b) \propto \frac{1}{\mathcal{A}^b}, \qquad \mathcal{T}^b \propto \frac{1}{\mathcal{S}(z^b)}$$

Theorem (1) has interesting implications. The first part (a) implies that even an infinitesimal relaxation of the constraint results in an unbounded marginal

ident by now, and the consequent implications and interpretations are completely different in this chapter.

propensity to consume when agents are at the constraint, i.e. $\partial_a c(a^b, z^b) \to \infty$ as $a^b \to \underline{a}^l$. This replicates some empirical observations we discussed in Empirical Observations, particularly the credit card utilization study showing that agents who are close to their spending limit on their credit card spend a much higher portion of their increased limit. Specifically, agents who are at their limit spend about \$0.45 for every \$1 increase in their limit. But agents with larger breathing room in their cards spend only about \$0.07 for every \$1 increase in borrowing limit (Gross and Souleles, 2002).

Further, $\partial_a \eta(a^b, z^b) \to -\infty$ as $a^b \to \underline{a}^b$ implies that as borrowers get closer to the constraint $\underline{a}^l$, the rate of change in leverage increases rapidly, eventually becoming unbounded at $\underline{a}^l$. Figure (2.2), obtained from numerically computing the stationary equilibrium, visually illustrates the equilibrium behavior of borrowing households with regards to systemic risk. In particular, Figure (2.2b) clearly shows $\partial_a \eta(a^b, z^b)$ getting unbounded as $a^b \to \underline{a}^b$.



(a) (b)

Figure 2.2: Debt Policy Function in Stationary Equilibrium

The second part (b) of theorem (1) says that the higher the default intensity λ, the higher the speed and lower the time $\mathcal{T}^b$ taken to reach the systemic risk states. So, with increasing credit risk of borrowers captured by increasing λ, the systemic risk states are reached faster. This means as borrowers become riskier, they reach systemic risk states faster. Also, the lower the (coefficient of) risk aversion $\mathcal{A}^b$ around the constraint, the higher the speed $\mathcal{S}(z^b)$ and the smaller the time $\mathcal{T}^b$ taken to reach the systemic risk states. In the limiting case of infinite absolute risk aversion $\mathcal{A} = \infty$, the constraint may never be reached in finite time.

These results imply that even small shocks, say a one-time small shock that reduces the borrowing limit $\underline{a}^b$, can elicit large changes in consumption, rendering the system fragile. The situation can get critical if there is a large mass of borrowers at or around the constraint. The corollary (1.1) in fact shows that our model can incorporate such a situation.

Corollary 1.1. ***(Dirac measure at the constraint):*** *There is a Dirac measure of the borrowing households at the constraint $\underline{a}^b$.*

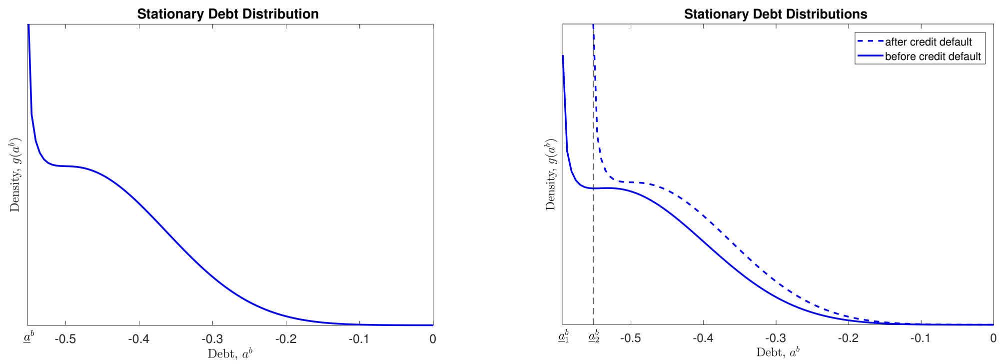
This corollary (1.1) implies that the left-tail of the leverage distribution is heavy. This result, in combination with theorem (1), reinforces our discussion regarding potentially dire consequences if the economy is hit by even small shocks. Note that this risk scenario is regardless of the presence of the financial sector. We want to further explore other factors that could be important for affecting the measure of the set around the constraint because increasing the left-tail of leverage distribution implies increasing first order systemic risk.

Corollary 1.2. ***(Heavy tail around the constraint):*** *Consider a set $\underline{\mathcal{B}}$ around the left-tail of the debt distribution. The measure of this set $\mathbb{P}(a^b \in \underline{\mathcal{B}})$ is increasing in intensity λ and lower risk aversion of the borrowers $\mathcal{A}^b$.*

Thus, the higher the default intensity λ and the lower the risk aversion $\mathcal{A}^b$, the heavier the left-tail of the debt/leverage distribution. This scenario translates into increasing systemic risk. Comparing this to the results of theorem (1), it appears that whatever factors increase the speed $\mathcal{S}(z^b)$ also seem to increase the left-tail of the debt/leverage distribution, leading to increasing systemic risk.

All the results explored until now can also be visually seen in Figure (2.3) below where the equilibrium debt/leverage distribution is computed before and after credit default, and the consequent reduction in borrowing limit and lower income for borrowers. This is just one *instance* of how systemic risk might manifest. Regardless of how risk materializes, the important aspect ultimately boils down to the measure of borrowers at and around the systemic risk states. If such a measure is large, the effects of even a small negative shock to the economy can have dire consequences. While not a perfect source of data for debt, for simple comparison we have included the debt distribution data in Figure (2.4) from the Census Bureau for 2005 and 2011 (periods before and after the recent financial crisis). We do not want to overemphasize that our model appears to replicate the shape of debt distribution in data. This survey data is certainly not representative enough to allow for complete comparisons. For proper comparison, one would need access to a much richer data set, possibly from credit unions or similar non-public/propriety sources. However, as a first approximation, this comparison seems encouraging.

Again, note that these results are independent of the financial sector. And it is exactly due to such initial risks in the economy that there arises a need for the financial sector's financial innovation function that can help mitigate such initial risks. We will visit this issue in detail in the financial sectors section below. We now turn to characterizing the lenders problem which is directly related to



(a) (b)

Figure 2.3: Instance of Systemic Risk

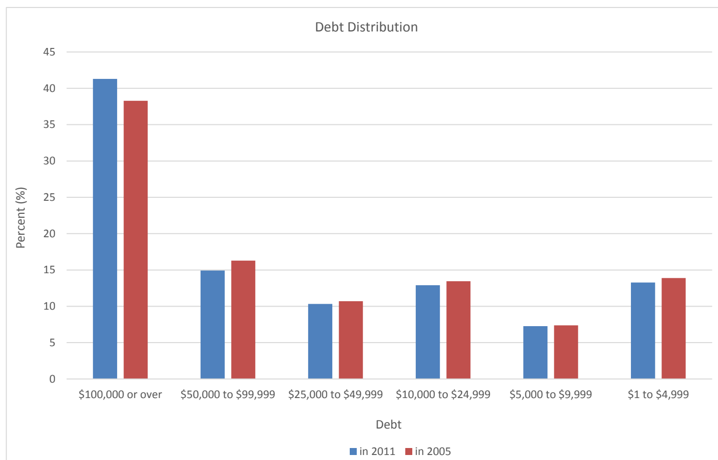
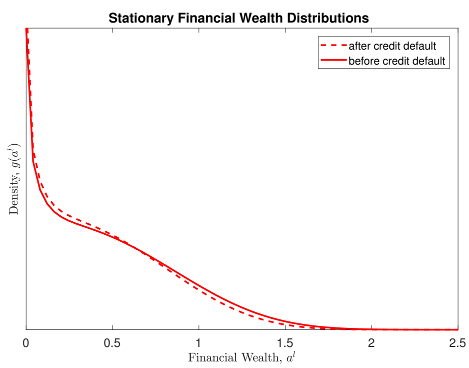


Figure 2.4: Debt Distribution in Data (Census Bureau)

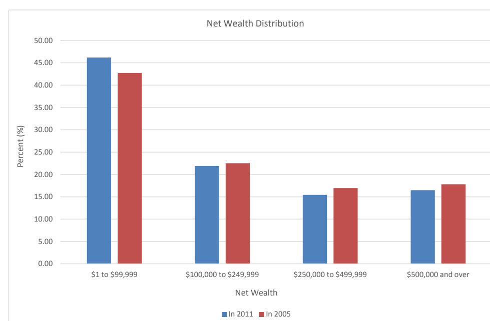
credit supply and outstanding leverage in the economy.

Theorem 2. ***(Credit Supply and Financial Innovation)****: Consider the solution to the optimization problem of lenders with initial wealth $a^l_0 > \bar{a}^l$ for some wealth level $\bar{a}^l$ and continued high income realizations. Let $E(a^l)$ be the total credit supplied and $\mathbb{P}(a^l \in \mathcal{B})$ be the measure of some set $\bar{\mathcal{B}}$ in the right tail. Then $E(a^l)$ and $\mathbb{P}(a^l \in \mathcal{B})$ are increasing in financial innovation, i.e. CDS purchases α by the lenders.*

From this theorem (2), we see that the more the CDS contracts α purchased (i.e. greater use of financial innovation), the higher the credit supplied. This is the moral hazard induced by the presence of financial innovation. Interestingly, this is also associated with an increase in the right tail of the financial wealth distribution. This means that with positive probability, the tail of the leverage distribution across borrowers also increases as lenders supply more credit. This result sharply highlights the equivocal nature of financial innovation. In studying the financial sector's problem below, we will add more dimensions to these results and be able to explicitly see different multiple risk mechanisms and channels through which financial innovation can either increase or decrease systemic risk. Note that this result is only for the stationary equilibrium. In the transition dynamics part, we will explore additional conditions under which greater use of financial innovation can lead to either higher or lower total credit supply.



(a) In Model



(b) In Data (Census Bureau)

Figure 2.5: Changes in Financial Wealth Distribution

Figure (2.5) is analogous to Figures (2.3) and (2.4) above in the borrowers' problem. For the same changes in leverage distribution as shown in Figure (2.3), the corresponding changes in *financial* wealth distribution can be seen in Figure (2.5a). After the credit event, the total financial wealth also declines as the mass of the distribution shifts to the left[17].

2.2.3 Systemic Risk with Financial Sector

Now we will explore various systemic risk properties by explicitly considering the financial sector. The equilibrium behavior of borrowers and lenders analyzed above constitute the core upon which we will build this section. The main concern of the financial sector is to remain solvent. This issue is particularly critical *at the random time* when a credit event occurs, i.e. when its services are needed most. This directly relates to our notion of systemic risk. The equilibrium leverage distribution across borrowers and the total financial innovation used, i.e. CDS purchased by lenders, naturally affect the financial sector's insolvency conditions, thereby affecting systemic risk.

Let us fix some notations and definitions first. Consider the first passage time (a stopping time) of capital position X_t below a certain threshold level $\bar{x} \le x$, defined as:

$$\tau_{\bar{x}} = \inf\{t > 0 : X_t < \bar{x}\}$$

If the capital decreases below this threshold level the premium rate cannot compensate for the decline in capital, and capital will decrease to minus infinity. The insolvency conditions of the financial sector can be characterized by the failure probability $\pi(x)$ and the distribution of first passage time $\mathcal{P}(\tau_{\bar{x}})$ of the

[17]We do not study the general income/wealth inequality in this chapter. Note that proper data for full (financial) wealth distribution is limited. For the purpose of being consistent, the wealth distribution data here is also obtained from the Census Bureau for 2005 and 2011. The Survey of Consumer Finances is another potential source, but it is also far from being a good source. Once again, we are confronted with data availability issues, especially the financial wealth distribution as required by the theory.

capital X_t below the threshold level $\bar{x}$, which are the same in our model:

$$\pi(x) = P(\tau_{\bar{x}} < \infty | X_0 = x) = \mathcal{P}(\tau_{\bar{x}})$$

Our main goal is to analytically characterize $\pi(x)$ and $\mathcal{P}(\tau_{\bar{x}})$ and then relate these to various equilibrium decisions of borrower and lenders. The capital position X_τ at the random time τ of the credit event is another important feature that defines whether the system remains stable or not. The approach of Lévy process (illustrated in Theory of Lévy Processes) allows us to jointly analyze the nature of the aggregate claims as it contains information about the default intensity λ and the *endogenously* determined leverage distribution across borrowers. Further, the premium collection rate per-unit time ξ_t contains information on the financial innovation used by lenders.

Recall that the dynamics of the financial sector is given by equation (2.9): $dX_t = (\gamma X_t + \xi_t)dt - dY_t$. First we will consider a simple case with $\gamma = 0$ and initial capital $x = 0$ to fix some baseline results because explicit closed-form expression is possible in this specification. After that we will analyze fully the general case.

Theorem 3. ***(Failure of Financial Sector):*** *Consider the dynamics of the financial sector as given in equation (2.9). Assume $\gamma = 0$, $x = 0$ and $\lambda\mathbb{E}(y) < \xi$. Then the failure probability $\pi(x)$ and the distribution of first passage time $\mathcal{P}(\tau_0)$ of the capital X_t below the threshold level zero is given by:*

$$\pi(x) = \mathcal{P}(\tau_0) = \frac{\lambda}{\xi} \int_{\mathbb{R}+} \bar{\mathbb{P}}(y)^b dy$$

This result essentially combines most of our previous results, and gives our first main result, compactly showing the conditions under which financial innovation can either increase or decrease systemic risk. First, as lenders purchase more credit risk protection (i.e. use more financial innovation α), this increases the total capital position of the financial sector. This makes the financial sector stable as the denominator ξ_t in theorem (3) increases, thereby decreasing failure probaility $\pi(x)$, all else being equal. However, since lenders can insure against the credit risk, they tend to supply even more credit as shown by theorem (2) above. This means that increased credit supply due to increasing use of financial innovation *endogenously creates risk*, because with positive probability, the tail of the equilibrium leverage distribution $\bar{\mathbb{P}}(y)^b$ is increased. Since, $\bar{\mathbb{P}}(y)^b$ is in the numerator, this increases the failure probability $\pi(x)$, making the financial sector unstable. Thus, while financial innovation improves risk sharing by helping insure against credit default risks in the model, it can also increase systemic risk as it encourages higher leverage due to the safety of insurance. This is the combination of counterparty risk and moral hazard story we discussed earlier, for the stationary equilibrium.

There are other interesting dimensions in this result. We see that the high default intensity λ is directly proportional to $\pi(x)$. If $\lambda = 0$, i.e. there is no credit risk in the economy, the financial sector never fails. For $\lambda > 0$ there are additional implications. From corollary (1.2) regarding borrowers' problem, we know that increasing λ increases the left-tail of the leverage distribution, captured by $\bar{\mathbb{P}}(y)^b$. Since $\pi(x)$ is directly proportional to $\bar{\mathbb{P}}(y)^b$, it is clear that if credit default risk λ increases, the failure probability $\pi(x)$ increases non-linearly, essentially getting magnified. Further, if the distribution of levered agents is concentrated at or around the constraint, the tail $\bar{\mathbb{P}}(y)^b$ is much heavier, which again leads to increased failure probability $\pi(x)$. To make things worse, if there is a Dirac measure of borrowers at the constraint, which is shown to exist in corollary (1.1), the effect on $\pi(x)$ could be enormous. All these translate into systemic risk in the economy. This situation clearly resonates with the situation before the recent financial crisis when a large fraction of households were at or near their borrowing limit. This ultimately resulted in the failure of Lehman Brothers and almost bankrupted AIG, which were some of the largest counterparty firms.

Further, if we add a *behavioral friction* by assuming that the financial sector underestimates the default risk by using ($\bar{\lambda} < \lambda$) in pricing the contracts (CDS rate s_t), the effects can be much higher. The λ in the numerator is the true default intensity value and it remains the same. But the perceived intensity $\bar{\lambda}$ used for pricing the total premium ξ_t in the denominator is lower than it would be otherwise. This underestimation of risk by the financial sector serves to amplify the systemic risk effects given that actual $\pi(x)$ above becomes much higher than it would be otherwise[18].

We now address the following questions: What is the total capital position of the financial sector *at the random (stopping) time* of the credit event? What economic factors affect the capital at such times? The answers to these questions will also serve to highlight the robustness of our results so far.

Theorem 4. ***(Financial sector's capital at credit event time*** τ***):*** *Let* $\mathcal{X}_\tau$ *be the total capital position of the financial sector at the random credit event time* τ *with intensity* λ*. Let* ξ*,* α *and* $\bar{\mathbb{P}}^b$ *be as defined above with* $\gamma = 0$*. Let* $\psi(\theta)$ *be the Laplace exponent associated with the Lévy- Khintchine exponent, characterizing the dynamics of capital* X_t *for* $t \le \tau$*. Let* ϕ *be the right-inverse function of* ψ*. Then we get the following relations:*

$$\mathcal{X}_\tau = \frac{1}{\phi(\lambda)}, \quad \textit{such that} \qquad \phi(\lambda) \propto \bar{\mathbb{P}}^b, \qquad \phi(\lambda) \propto \frac{1}{\xi}, \frac{1}{\alpha}, \qquad \phi'(\lambda) > 0$$

[18]This is analogous to the notion of neglected risks as in Gennaioli et al. (2012). They explore links between financial innovations and aggregate leverage in a three-period model in which agents neglect certain states of the world (unlikely risks), i.e. investors underestimate the risks (and actually bear more risks than they think they are holding). This leads to over-issuance of securities.

This result gives a clear picture of the financial sector's capital position not just at any time, but at the random time when a credit event in the economy materializes. It is exactly at such a random time that we need the financial sector to be well capitalized and be able to serve its function. We see that $\mathcal{X}_\tau$ is directly proportional to the tail of the leverage distribution across borrowers $\bar{\mathbb{P}}^b$ and to the default intensity λ. $\mathcal{X}_\tau$ is also inversely proportional to the premium ξ and financial innovation α. This reinforces the findings above, acting as a robustness check for our baseline result of theorem (3). Again, if the financial sector underestimates risk by using ($\bar{\lambda} < \lambda$), this can erroneously make it seem that there is enough capital to cover the losses at the random time of credit event. However, since the actual/true default intensity λ is higher than presumed intensity $\bar{\lambda}$, the capital position can be much more lower than anticipated when the credit event occurs. This *behavioral friction* again magnifies systemic risk effects.

Now we answer the following related questions: What threshold level of capital must the financial sector satisfy to remain solvent at all times and what determines this threshold level? What is the time that the financial sector's capital might fall below this level and what determines such an event of insolvency? Here, we relax the assumption of $\gamma = 0$ and $x = 0$, and consider the fully general case of the capital evolution equation (2.9).

Consider the first passage time, but now defined with respect to a different solvency threshold level $\bar{x} = -(\xi/\gamma)$ for $X_0 = x$:

$$\tau_{-(\xi/\gamma)} = \inf\{t > 0 : X_t < -(\xi/\gamma)\}$$

From equation (2.9), we can see that when the surplus gets below $\bar{x}$, the premium collected will not be able to bring the capital position above the threshold level. This means, the surplus X_t will drift to $-\infty$ whenever $X_t < -(\xi/\gamma)$ i.e probability $\mathbb{P}[X_t \geq -(\xi/\gamma) | X_t < -(\xi/\gamma)] = 0$ almost surely.

First notice the generality of this formulation. Considering $\gamma > 0$ makes the new solvency threshold level quite generous in the sense that now the capital position of the financial sector can go negative, up to $-(\xi/\gamma)$, while still remaining solvent. Thus, γ acts like a stabilizing force. This general formulation allows us to further discuss financial stability as well, where the stability is due to the fact that the system can withstand a larger capital shortfall. In the model section above, we discussed the motivations and some situations in financial markets where this type of scenario can be observed.

Before stating theorem (5), we will prove a simple lemma that will serve to make our proof and the results of theorem (5) more intuitive.

Lemma 1. ***(Equivalent Processes):*** *Let $dX_t = (\gamma X_t + \xi)dt - dY_t$ be the stochastic differential equation of a process X_t with $X_0 = x$, and denote its law by $\mathbb{P}_{X_0=x}$. Let $dX_t = \gamma X_t dt - dY_t$ be the stochastic differential equation of another process X_t with $X_0 = x + (\xi/\gamma)$, and denote its law by $\mathbb{P}_{X_0=x+(\xi/\gamma)}$. Then the law of two such processes, $\mathbb{P}_{X_0=x}$ and $\mathbb{P}_{X_0=x+(\xi/\gamma)}$, are equivalent.*

Theorem 5. ***(Failure of Financial sector in finite time):*** *Consider the dynamics of the financial sector $dX_t = (\gamma X_t + \xi)dt - dY_t$ with $\gamma > 0$ and $x > 0$. Then the solvency threshold $\bar{x}$ is $\bar{x} = -(\xi/\gamma)$. Given the driftless subordinator Y_t , $\mathcal{V}_t = \int_0^t e^{-\gamma s} dY_s$ is the total flow of present discounted value of claims until time t. Then the insolvency condition is given by:*

$$\pi_t(x) = \mathcal{P}(\tau_{-(\xi/\gamma)} \leq t) = \mathbb{P}(\mathcal{V}_t > \mathcal{V}_0)$$

where $\mathcal{V}_0 = (x + \xi/\gamma)$ denotes the initial capital plus the discounted premium collection rate. Then:

(a). For increasing initial capital requirement $x_2 > x_1$, $\mathcal{P}$ is decreasing, i.e. $\mathcal{P}_2 < \mathcal{P}_1$.

(b). For increasing tail distribution of leverage across borrowers $\bar{\mathbb{P}}^b_2 > \bar{\mathbb{P}}^b_1$, $\mathcal{P}$ is increasing, i.e. $\mathcal{P}_2 > \mathcal{P}_1$.

(c). For an increasing purchase of CDS contracts by lenders $\alpha_2 > \alpha_1$, $\mathcal{P}$ is decreasing, i.e. $\mathcal{P}_2 < \mathcal{P}_1$.

We now discuss the implication of this theorem in conjunction with other results, and highlight some policy implications. Theorem (5) provides a broad way to think about financial stabilization policies. Policies such as the Dodd-Frank Act and the Fed's annual stress test that heavily focus on capital buffer requirements (x in our model) are at best *partially informed*. Higher capital buffer certainly reduces systemic risk, but it does not account for other intricate risk flow mechanisms through the channel of leverage distribution *across borrowing households* in the overall economy. This is seen from the $\mathcal{V}_t$ component in the expression $\mathbb{P}(\mathcal{V}_t > \mathcal{V}_0)$ of theorem (5) which can be interpreted as representing the equilibrium demand side. It is affected by time and speed towards systemic risk states and the leverage distribution across borrowing households in the economy. Thus, mainly focusing just on the capital buffers (and its variants) clearly obscures other important channels, casting doubts over the efficacy of such policies[19].

Further, as our results show, the effect of making lenders purchase more credit risk protection is not clear. The $\mathcal{V}_0$ component of theorem (5) can be interpreted as representing the credit/loan supply side which embeds various

[19]From a modeling standpoint, this again underscores the importance of considering full fledged heterogeneous agents models in understanding systemic risk as this allows us to study the entire leverage and the financial wealth distributions across agents. Formulating policies based on just the aggregates can be misleading as we ignore the dynamics coming from the distributional and tail properties.

determinants of credit supply in addition to the capital requirement x. As seen, the total premium collected ξ relates to the loan notional insured. So, policies that require more purchase of credit risk protections (higher α in the model) against the risky loans seem to improves financial stability, evident from decreasing failure probability $\mathcal{P}$ in theorem (5). This constitutes the first order benefit of financial innovation– providing an efficient risk management tool for improving financial stability. However, this is again an incomplete story. As seen from theorem (2), greater use of credit risk protections can also increase the total credit supply and consequently the total outstanding debt/leverage in the economy. And with positive probability the tail of the leverage distribution can also increase. This can again feeds back into the $\mathcal{V}_t$ component which is directly related to the leverage distribution. This serves to increase the failure probability $\mathcal{P}$, thereby undoing the first-order stabilizing effect of financial innovation.

Also note the stabilizing role of the γ component in theorem (5). During economic distress, the financial sector might face a capital shortfall, rendering its capital position negative. However, the financial sector can borrow funds at the rate γ to cover capital shortfall. Thus, if γ is made lower, say via monetary policy, such a response effectively allows the capital position to go further into the negative territory. Since γ is in the denominator of the threshold $\bar{x} = -(\xi/\gamma)$, this allows the financial sector to absorb the effects of negative shocks during times of economic downturns. Thus, the zero lower bound interest rate policy implemented by the U.S. Federal Reserve Bank in the aftermath of the recent crisis seems to have been in the right direction as seen through the lens of our model.

Stabilizing policies also need to address the issue of capital position, not just at any time but *at the random time* (τ in our model) when the economy is hit with credit events. Our theorem (4) provides some policy insights. The aggregate capital position of the financial sector at the random time of a credit event depends on a variety of similar factors discussed above, and not just on the capital buffer x that the financial sector is required to hold. The equations in theorem (4) make this point clear and our above discussions apply.

So, how should policy address the issue of systemic risk? The answer is not as simple as just imposing different variants of capital buffer regulations. This approach is incomplete. Given that all objects of interest in our model are measurable in data (the only issue is access to data), our policy suggestions can be easily applied in practice. Fully incorporating a variety of factors aforementioned, especially that relate to the distribution of leverage and financial wealth across agents, would require richer data sets, many of which might be confidential or publicly unavailable. For example, data from credit unions or detailed over-the-counter derivatives trading data. It should not be difficult for policy makers to access such data sets. And if such comprehensive and proper data sets are not available, it would be a great long-term investment to create such data sets for scholarly use. This will serve to enhance our understanding of a wide variety of macro-finance issues related to economic and financial stability.

2.3 Conclusion

In this chapter we studied systemic risk and financial innovation in context of an stationary equilibrium model. In the following chapter we will continue to build on our analysis above by exploring different evolutionary properties of systemic risk in a transition dynamics setting. We will also highlight some new policy implications.

CHAPTER 3

SYSTEMIC RISK AND FINANCIAL INNOVATION: A TRANSITION DYNAMICS ANALYSIS

In this chapter we *analytically* characterize the transition dynamics of our economy which translates into the evolutionary properties of systemic risk over time. This is a difficult problem in the context of a heterogeneous agent model because we are not just looking at the evolution of a variable over time, but rather the evolution of the measure itself, i.e. the entire distribution across agents. Once an equilibrium condition (partial or general) is imposed, the complexity magnifies and it becomes challenging to get analytical traction on the model. Hence the analysis of transition dynamics in the literature is purely numerical, with the notable exception of Gabaix et al. (2016), albeit in a reduced form model with focus on the transition speed. More discussion on this can be found in in Gabaix et al. (2016), Guvenen (2012) and Heathcote, Storesletten and Violante (2009).

Nevertheless, the basic idea is simple: given that the economy starts to evolve some initial distribution/conditions, we want to characterize the evolutionary process itself, i.e. the transition process, towards some long run stationary equilibrium. For example, imagine some shock hits the economy, moving it off its initial (stationary) distribution. This shock could come from a variety of sources. Say there is a shock to the overall borrowing limit that reduces $\underline{a}^b$ as in Figure (2.3). Or say, after the shock, the volatility of the income σ^b or default intensity λ changes. Such economic changes set in motion the transition of the economy to a new long-run stationary equilibrium distribution (assuming it exists). In other words, as $t \to \infty$, the economy eventually converges to a new stationary distribution[1].

Consider Figure (2.3) in our stationary equilibrium section above. This shows before and after effects of the credit shocks, and both of those represent stationary equilibrium distributions. The transition dynamics study seeks to characterizes how the economy moves from the some initial distribution/conditions to some other (stationary) distribution/conditions in the future. Further, there may be many interesting economic properties *during* such an evolutionary process. With respect to systemic risk, there could be important policy implications regarding evolution of the stability features of the economy. This section analyzes such systemic risk properties *during* the evolutionary/transition phase of the economy.

Our treatment of transition dynamics is quite unique in the literature as we are able to obtain a variety of analytical characterizations of transition dynamics in context of an equilibrium model by introducing the Feynman Path Integral (FPI hereafter) method of quantum mechanics. As we will see, these methods are completely general, which means these can be used to give analytical answers to many more questions that relate to transition dynamics. In the literature review, we have already discussed various reasons for choosing this method over other alternatives that rely on the spectral theory of operators as in Gabaix et al. (2016). We will give more details below. In short, to address transition dynamics issues, the mathematical sophistication (and complexities) of the spectral theory of operators can be substituted (not necessarily always though) with the equally powerful FPI method because both methods are mathemati-

[1]Gabaix et al. (2016) also present a similar thought experiment: initially the economy is in a stationary distribution with some initial parameters and say a parameter changes (for example, the variance in income). So asymptotically as $t \to \infty$, the distribution converges to its new stationary distribution, and they explore the speed of this transition.

cally equivalent. In our model, FPI appears most suitable for gaining analytical traction[2] [3].

3.1 Notes on Feynman Path Integral

Before moving to our results, we discuss some key aspects of FPI formalism. In Feynman Path Integral appendix we provide further details. A fundamental concern of quantum mechanics is to analyze how the state of particles evolve over time – given some initial conditions governing the state, how does the system propagate over time to some other set of conditions? More specifically, one needs to determine the time evolution of the wave-function of a particle from some initial conditions. In a standard quantum mechanics approach, this can be answered by working with the Schrödinger equation, which propagates the particle in time. The fundamental question addressed by FPI formulation of quantum mechanics is this: if a particle is at position x_0 at time $t = 0$, what is the *transition probability amplitude* that it will be at some other position x_t at a later time $t > 0$? The key idea, simple but seemingly counterintuitive, is that the particle takes all possible paths, and the total probability amplitude can be calculated as the sum of the amplitude over all possible paths. This ultimately

[2] Feynman (1948) gave a new formulation of quantum mechanics using path integrals, which was different from but mathematically equivalent to the standard approach via Schrödinger equation. The standard approach (via Hamiltonian) heavily relies on the spectral theory of operators on Hilbert Space, with focus on the Hermitian, i.e. self-adjoint operators. Thus, the methods in Gabaix et al. (2016) reflect those used in the standard formulation of quantum mechanics. The well-known Schrödinger equation is analogous to their Kolmogorov equations associated with the reduced-form income process.

[3] Developing path integral allowed Feynman (1948) to completely side step the standard formalism of quantum mechanics, as noted in this seminal paper: *Non-relativistic quantum mechanics is formulated here in a different way. It is, however, mathematically equivalent to the familiar formulation. In quantum mechanics the probability of an event which can happen in several different ways is the absolute square of a sum of complex contributions, one from each alternative way.*

yields the path integral expression that we are about to explore.

The gist of the method can be summarized by comparison as follows. Generally, we can associate a Kolmogorov Forward (KF) equation to describe the time evolution of the distribution associated with a well-defined stochastic process. The solution of KF equation gives the expression for such a distribution. In FPI formulation, we associate a Lagrangian to a given stochastic process. This Lagrangian is a function of time, position x_t and time derivative $\dot{x}_t$, i.e. the velocity describing the dynamic trajectory of the process[4]. The process can take all possible paths between the initial position x_i and final position x_f where each path contributes to the total probability. The Lagrangian implicitly associates certain probability to each possible path. Then a time integral of the Lagrangian is evaluated, called the *action* functional, on each possible path satisfying something called the *principle of least action*, which is a variational principle, i.e. the *action* is evaluated at the optimal path. The optimal path is obtained by solving the variational problem in which the objective is defined via the *action* functional. This is equivalent to but slightly different than the solution via optimal control technique. This eventually gives the expression for the transition density K of the process:

$$K(t_i, x_i; t_f, x_f) = \int_{x(0)=x_i}^{x(t)=x_f} \mathcal{D}x(t) e^{-S[x_t]}$$

where $\mathcal{D}x(t)$ is the notation that denotes integral over all paths (details in appendix) and the exponential of the negative of *action* $S[x_t]$ gives weight to each path. So the optimization solution via variational technique gives less weight to paths that are far from the optimal path, and more weight to those paths near

[4]Note that this Lagrangian is different from the usual notion of Lagrangian that economists use in optimization problems. The Lagrangian involved in optimization methods, involving Lagrange multipliers, are generally referred to as Lagrange equations of the *first kind*. In our context, the Lagrangian representation with FPI is the *second kind*. Both are related concepts but used in different settings. See Hand and Finch (1998) for additional details.

the optimal path.

The fundamental concern for us is to describe the evolution of the entire distribution of economic agents over time. In general, this can be done efficiently in a continuous time setting via KF equations, as discussed in the dynamic programming solution method above. Thus, the Schrödinger equation and Kolmogorov Forward equations are mathematically analogous. The former describes the quantum dynamics, and the latter describes the diffusion dynamics. Since we are characterizing equilibrium transition dynamics, our problem is already quite involved right from the start. For this reason, following the standard quantum mechanics method, or analogously the spectral theory methods, do not offer a clear resolution to this complexity. However, FPI appears to be the more suitable method given our model[5].

3.2 Results

We now start our transition dynamics analysis. First we derive a general expression for the transition density function K^b of the leverage distribution, a key object in the study of transition dynamics (similar derivation applies for financial wealth distribution). This is directly related to the stochastic process governing the evolution of the optimal leverage/debt in equilibrium i.e. $da_t^b = \eta_t^b(a_t, z_t)dt$. The first main difficulty is that compared to the models with reduced form dynamics, we cannot freely make assumptions on η_t^b as it is endogenously determined and highly non-linear. However, our model structure somewhat simpli-

[5] In fact, the power of FPI comes from the fact that it is very efficient while dealing with particularly complicated problems (such as those appearing in quantum field theory) that are otherwise very cumbersome and difficult to solve, if not impossible, via standard spectral theory methods.

fies the problem, given that our primary expression characterizing the equilibrium evolution of debt/leverage is quite simple: $da_t^b = \eta_t^b(a_t, z_t)dt$. This allows us to use FPI to easily obtain the analytic expression for the transition density function, paving the way for analytically characterizing various transition dynamics properties that the literature has not explored before[6].

We denote $\eta_t = \eta_t(a_t, z_t)$ for notational simplicity. Note that we have changed the notation of (marginal) density functions from g^b and g^l to K^b and K^l, respectively (and to $\mathbb{K}^b$ and $\mathbb{K}^l$ respectively for the associated probability). This is done so as to be consistent with notation in FPI literature and for readers to be able to easily reference relevant articles.

Theorem 6. ***(Transition Probability/Trajectory):*** *Consider the evolution of the debt $da_t^b = \eta_t^b dt$ where η_t^b is the optimally chosen debt policy function. Then the transition density function of going from initial debt a_i at time t_i to some final debt a_f at time $t_f > t_i$ is given by:*

$$K^b(a_i, t_i; a_f, t_f) = \int_{a(t_i)=a_i}^{a(t_f)=a_f} \mathcal{D}a(s) e^{-\int_{t_i}^{t_f} (1/2)(\eta_s^b)^2 ds}$$

Similar expression for K^l holds for the evolution of the financial wealth $da_t^l = \eta_t^l dt$.

This result succinctly characterizes the transition probability between any two points in space-time, thereby allowing us to analytically understand the properties of the otherwise complicated transition trajectory of the economy. In other words, this describes the evolution of the entire distribution of leverage (and financial wealth), and thereby the evolution of systemic risk[7]. We can

[6]Note that we could have also associated an operator acting on leverage distribution, giving us Kolmogorov Forward equation. But this approach does not give the clear analytical expression we are after. Nevertheless, the KF equation can still be helpful in deriving additional properties of transition dynamics, in conjunction with the transition probability obtained from FPI. We will revisit this issue when we discuss transition speed.

[7]It is worth noting that the proof of this expression, provided in the appendix, is quite sim-

also directly interpret this theorem in our model's context as follows: transition while the economy is at extremes (characterized by either high or low drift η_t^b) is relatively low. When the drift is at extremes, transition probability away from any given state is low, relative to when the drift is not at either extremes, i.e. around the middle of the spectrum. This is analogous to the stationary equilibrium spectrum of the drift shown in Figure (2.2) for η_t^b. This shows that extremes are *sticky* in the sense that once the system enters such states it tends to stay there for a relatively longer period of time during the transition phase of the economy. This idea motivates a very intuitive corollary below.

Corollary 6.1. ***(Leverage Trap):*** *Let $a_i^b \in \underline{\mathcal{B}}$ and let $\mathcal{B}_f = (a_f^b - a_i^b)$ denote an increasing Borel set indexed by f such that whenever $a_{f'} \geq a_f$ we have $\mathcal{B}_f \subseteq \mathcal{B}_{f'}$. Let $\mathbb{K}(a_f, t_f; a_i, t_i)$ denote the transition probability. Then $\mathbb{K}((a_f, t_f) \in \mathcal{B}_{f'}; (a_i, t_i)) \leq \mathbb{K}((a_f, t_f) \in \mathcal{B}_f; (a_i, t_i))$, and the rate of decline is exponential.*

This corollary (6.1) says that if agents are around high leverage systemic risk states, then the probability of transitioning away to the lower leverage states gets increasingly difficult. Stated differently, the deeper the economy moves into high leverage systemic risk states, the lower the chances of moving out of such states, essentially trapping the economy in the high leverage states, i.e. once the economy moves to high leverage systemic risk states, it tends to stay there. In particular, the transition probability may be the lowest if agents are at the constraint $\underline{a}^b$, i.e. agents might get stuck in those states for a long period of time (we will make this notion much more precise in our next result). Further,

ple. This makes the formulation of transition dynamics via quantum mechanics methods very intuitive and immediately applicable to a wide variety of economic settings. This helps, to some extent, bypass some of the subtle/complex aspects of functional analysis, and specifically the spectral theory. For example, the transition probability for the baseline model in Gabaix et al. (2016) without jumps can be obtained by replacing the exponential in the above theorem (6) with $exp(-\int_{t_i}^{t_f} \frac{1}{2\sigma^2}(\dot{x}_s - \mu)^2 ds)$.

the rate of decline in transition probability is exponential as agents move deeper into systemic risk states, eventually becoming lowest at the constraint.

The intuition here is a bit subtle: in the model, when borrowing households take on more leverage they consume more. And highly leveraged households want to take on even more leverage. This is due to the incentive structure in the model. Given their budget constraints, borrows can borrow up to their borrowing limit, with the possibility of default and limited liability (as a certain fraction of income is reduced but they can continue to participate in debt market). This means borrowers have the incentive to keep on borrowing as much as they can to maximize their lifetime value from consumption. And when there are many highly leveraged borrowing households, it makes it difficult for the economy as a whole to deleverage. Hence, the *leverage trap*.

The implications of this corollary (6.1) are interesting. Say some shock hits the economy and the borrowing conditions get tighter. This change in economic environment makes the system start its transition towards a new stationary distribution. If a huge mass of agents are at or near the constraint, then they must either default or deleverage rapidly to meet the new constraint. Thus, if a large fraction of agents are affected, the economy faces massive debt-deleveraging, defaults, and consequent sharp reduction in consumption. During such a transition phase, those agents at or around the constraint are affected the most because they have very low probabilities of escaping the high leverage systemic risk states. Hence the term *leverage trap*, as such agents might get trapped in those high leverage states for a long time as the economy continues to transition. This means their consumption might be quite low for a sustained period of time. So, larger the mass of agents around the systemic risk states, greater the effects of economic shocks. This scenario is exemplified in Figure 2.3) above via shift in stationary equilibrium. Further, *at a given time*, the results that we discussed regarding the financial sector's failure probability also hold here.

Following the discussion above, two related questions arise naturally. How long can we expect the economy to remain in systemic risk states before it exits such states? How fast is the transition and what economic factors determine this transition speed? Answers to these two broad questions are independent of the leverage trap interpretation we discussed above. This is because the transition dynamics we explore is quite general, allowing for the flexibility of adapting to various situations to get particular insights. For the results below, it is useful to think of the economy starting from some initial conditions, not necessarily after a negative economic shock event as discussed above, and evolving towards a long run equilibrium distribution.

Corollary 6.2. ***(Exit-time distribution from systemic risk states):*** *Let $\tau_e = inf\{t > 0 : a_t^b \notin \underline{\mathcal{B}}; a_0^b \in \underline{\mathcal{B}}\}$ be the first exit time of a_t^b from the set $\underline{\mathcal{B}}$. Let $\mathbb{P}_{\tau_e}(t)$ be the probability distribution of τ_e. Then:*

$$\mathbb{P}_{\tau_e}(t) = 1 - \int_{\underline{\mathcal{B}}} K^b(a_0 \in \underline{\mathcal{B}}; a_t \in \underline{\mathcal{B}}) da^b$$

Thus, quite intuitively, we see that exit-time probability is the complement of probability of staying in systemic risk states $\underline{\mathcal{B}}$. This result and corollary (6.1) show that if the economy is around the systemic risk states, the chances of exiting such states in a given time frame are low. Thus, the economy might spend a lot of time around those systemic risk states during the transition process. This reinforces and complements the idea of the *leverage trap* by adding a time dimension to the intuition – if agents are trapped in high leverage states, the chances of exiting such states in any given time are very low[8].

Another important aspect of transition dynamics is transition speed. We discussed various comparative statics exercises in our stationary equilibrium analysis – we considered changes to different parameters and analyzed how that would change the distribution. Here, the main idea is to understand *how fast* such changes occur in the economy. We also want to understand different economic factors that are important in characterizing such a speed. In the context where the economy is hit with some negative shocks (like the recent financial crisis), transition speed captures the *time to recovery* aspect. After shocks to the initial stationary distribution, the economy starts transitioning to a new stationary distribution in the long run, where reaching that new distribution is analogous to having recovered. Consequently, the time taken to teach the new distribution can be naturally interpreted as capturing the time to recovery.

Before stating the transition speed result, we first need to discuss some technical details, comparing FPI approach to the standard operator theory (spectral analysis) method for obtaining transition speed, as in Gabaix et al. (2016). We can write the transition speed in terms of the KF equation associated with the equilibrium debt/leverage dynamics $da_t^b = \eta_t^b dt$ as:

$$\partial_t \mathbb{K}^b = \mathcal{H}^* \mathbb{K}^b = \kappa \mathbb{K}$$

where $\mathcal{H}^*$ is a linear differential operator and κ is an eigenvalue. We know that the optimal debt policy function η_t^b is highly non-linear, endogenous, and

[8]Understanding the distribution of first exit time from systemic risk states is crucial for addressing different types of policy-relevant situations where the time frame of the policy implementation is important. This result helps take into account the notion of *how long* the system might remain in systemic risk states during the transition phase.

does not have an analytic solution (only a first order approximation around $\underline{a}^b$ was explored in theorem (1)). From previous discussions, we know that there is a stationary equilibrium and the system converges to a stationary distribution, i.e. $\partial_t \mathbb{K}^b = 0$ has a solution. Then the standard operator theory suggests that the distribution $\mathbb{K}^b$ converges exponentially, where the rate of convergence (transition speed) is given by the non-trivial eigenvalue κ of the operator $\mathcal{H}^*$. Also, directly solving for $\partial_t \mathbb{K}^b$ for a given time t would yield a solution in terms of a unitary/evolution operator which forms the continuous one-parameter group generated by $\mathcal{H}^*$ operator.

Thus, analyzing the transition speed (and other dynamics) via spectral analysis methods discussed above offers no resolution to our problem given the highly non-linear and endogenous nature of η_t^b. However, specifying the debt evolution process in a simple way as $da_t^b = \eta_t^b dt$, allows for considerable simplification to our problem as we leverage the analytical tractability offered by FPI method and side step complications arising in standard operator theory route.

Corollary 6.3. ***(Transition Speed):*** *Let $\partial_t \mathbb{K}^b$ denote the transition speed characterizing the transition from some initial condition $\mathbb{K}_0^b$. Then,*
(a). $\partial_t \mathbb{K}^b$ is increasing whenever $\eta_t^b < 0$ i.e. when the economy is leveraging up.
(b). Assume $\frac{\eta e^{\int \eta_s ds}}{2e^{S[x_t]}} < 1$. Then $\partial_t \mathbb{K}^b$ is decreasing whenever $\eta_t^b > 0$ i.e. when the economy is deleveraging.

Thus, the transition speed (which captures the time to recovery) of moving from some initial stationary equilibrium to another new one after some shocks is faster if the economy can leverage up, and slower if the economy deleverages. If the economy is leveraging up, which occurs whenever $\eta_t^b < 0$, we can expect the economy to transition quickly. By contrast, when the economy is deleveraging, the transition speed decreases (under a condition).

A different interpretation makes this result seem somewhat counterintuitive. It implies that faster adjustment to shocks goes hand in hand with increasing systemic risk (as leverage increases), and slower adjustment with decreasing systemic risk. This highlights the trade-off between faster recovery time and accepting increased systemic risk in the process. In other words, the possibility of systemic risk materialization is relatively higher during good economic times[9]. The main intuition behind this results is this: the economy tends to go back to some stationary equilibrium distribution after some shocks. In the model, stationary equilibrium distribution features many highly leveraged borrowing households. So, if the borrowing households can leverage up, this makes it easier for the economy as a whole to adjust and recover faster after some shocks, i.e. reach the stationary equilibrium distribution faster.

The FPI $K(a_i, t_i; a_f, t_f)$ takes into account all possible paths between two points in space-time, and those paths around the classical/modal, i.e. the optimal path provide important contributions to the probability amplitude. Clearly, this is the path of least action that corresponds to the optimally chosen path in our economy. In our model, the evolution of the optimally chosen drift of the debt η_t^b describes the optimal path of a_t^b. Then using calculus of variations technique, i.e. applying the variational principle to the action $S[a_t^b]$, we obtain the

[9]This corollary (6.3) replicates a defining characteristic of many economies after some negative shocks when the deleveraging process makes the transition process to new equilibrium painfully slow. This was evident in the aftermath of the recent financial crisis when U.S. households had massive debt overhang, and had to deleverage on a large scale, making the recovery process painfully slow. This result appears to imply that the transition may be faster (slower) with higher intensity λ, because when the economy is leveraging up (deleveraging), it could also signal increasing (decreasing) credit/default risk of borrowers captured by λ.

well-known *Euler-Lagrange equation*:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{a}} = \frac{\partial L}{\partial a}$$

The immediate implication of this differential equation is the following result characterizing the endogenous acceleration mechanisms of the leverage.

Corollary 6.4. ***(Acceleration of Leverage):*** *Let $\ddot{a}^b$ denote the rate of change of η_t^b i.e. the acceleration of the optimal leverage. Then,*

$$\ddot{a}^b = \eta_t^b(a_t, z_t)(r_t - \partial_a C_t^b(a_t, z_t)) = \eta_t^b(a_t, z_t)\partial_a \eta_t^b(a_t, z_t)$$

Now we turn to the dynamics associated with lenders. We want to understand how aggregate credit supply changes with time as lenders use more/less of financial innovation (CDS purchases α). In the stationary equilibrium section above (theorem (2)), we showed that greater use of financial innovation leads to higher credit supply in the economy (the moral hazard aspect). We now expand on that result, as we explore conditions during the transition dynamics when greater use of financial innovation may lead to either higher or lower total credit supplied in the economy. In other words, the moral hazard aspect of the financial innovation may or may not hold during the transition dynamics phase.

Theorem 7. ***(Financial Innovation Amplifies Credit Cycles):*** *Consider the evolution of loan/credit supply $da_t^l = \eta_t^l dt$ where η_t^l is the optimally chosen drift. Let $E(a_t^l)$ be the aggregate credit supply at a given time t. Then,*

(a). If $\eta_t^l > 0$, then $E(a_t^l)$ is increasing in CDS purchases α.

(b). If $\eta_t^l < 0$, then $E(a_t^l)$ is decreasing in CDS purchases α.

Further,

(c). Let $\bar{\mathcal{B}} \in \mathbb{R}_{++}$ denote the Borel set to the right rail of the financial wealth distribution, such that for for all $a^l \in \bar{\mathcal{B}}$ we have $\eta_t^l > 0$. Then the tail measure $\mathbb{P}(a^l \in \bar{\mathcal{B}})$ also increases with CDS purchase α.

Thus, if the drift of the optimal savings η_t^l is positive, the total credit supply increases as lenders use more financial innovation. Interestingly, the right tail of the financial wealth distribution also gets heavier. On the other hand, if the drift η_t^l is negative, the total credit supply decreases as lenders use more financial innovation. The implication is that greater use of financial innovations during credit contraction periods after some shocks can enhance such contractions. Consequently, greater use of financial innovation during credit expansion/boom periods can also enhance such expansions. Thus, during good economic times (when credit supply is presumably already booming) increased use of financial innovation leads to increased total credit supply. And with positive probability the tail of the leverage distribution across borrowers also increases. This increases the systemic risk in the economy. This has similar undertones of the moral hazard story in stationary equilibrium (theorem (2)). However, this story gets reversed when economic times are not as good (or even bad) characterized by negative drift η_t^l.

The intuition here is straightforward: when credit and the economy are expanding, having the safety provided by financial innovation makes the lending households supply even more credit in the economy. But when credit and the economy are contracting, increased use of financial innovation reduces the total amount of credit supply, which also reduces the overall leverage. Then, a natural question is, what is the impact of such a reduction in overall leverage in the economy during the transition phase? This leads to another new insight.

Corollary 7.1. ***(Financial Innovation and Recovery):*** *Greater use of financial innovation, i.e. increasing purchases α, after some negative shocks when credit and the economy are contracting, can slow the recovery process.*

The proof of this result immediately follows by combining theorem (7) with corollary (6.3) on transition speed. Theorem (7) shows that financial innovation amplifies credit cycles. This means greater use of financial innovations during credit contraction periods can enhance such contractions, leading to the reduction in overall leverage. Corollary (6.3) shows that the transition speed (which captures the time to recovery) of moving from some initial stationary equilibrium to another new one after some shocks (like the recent financial crisis), is faster if the economy can leverage up, and slower if the economy deleverages. So, theorem (7) and corollary (6.3) together imply that greater use of financial innovations after some shocks first reduces the overall leverage in the economy; and, second, this deleveraging process makes the transition speed slower, i.e. the recovery process is slower.

We now briefly discuss how our transition dynamics results explain some of the empirical motivations and observations (in Empirical Observations) discussed earlier and also generate some new policy implications. These results offer another perspective on the massive debt overhang just before the recent financial crisis, and the slow recovery process after the crisis was. First, we know that the U.S. economy and credit were expanding since the late 1990s until the crisis. In the same time period, we also saw an exponential growth in the use of various financial innovations like the CDS. So, it is entirely plausible that greater use of financial innovations, during the period from late 1990s to just before the start of financial crisis of 2007-2009, helped amplify the already expanding credit supply in the economy. This is a new testable prediction generated by our model. Second, after the 2007-2009 financial crisis, U.S. households had to massively deleverage (as seen in Figure (A.1)), and the overall recovery process was painfully slow. Our model also predicts that deleveraging makes the transition speed slower which captures the recovery process.

We now highlight some new policy insights. Our results suggest that policy should not require greater use of credit risk protections when credit and economy are already contracting. Such a policy will, first, induce even lower credit supply in the economy, as above results predict. Second, such a policy can also make the recovery process slower as lower credit supply generally means the economy is not able to leverage up, which makes the recovery process slower. So, during economic downturns when credit is already contracting, policies that require the lenders to purchase more credit risk protections against existing risky loans (that is, purchase more risk protections like CDS to hedge against the existing risk, which seems to be a normal response in downturn when risk is perceived to be high), can be counterproductive. Such policies will, first, result in even lower credit supply in the economy; and, second, due to lower leverage, the time to recover from the crisis periods also becomes longer, thereby prolonging the effects of crisis.

3.3 Conclusion

Chapters 2 and 3 study systemic risk, especially as it relates to financial innovation. We build a heterogeneous agents model in continuous time and study full equilibrium dynamics of an economy with financial frictions and three classes of agents: heterogeneous borrowers, heterogeneous lenders, and a representative financial sector. We analytically characterize some intricate relation between financial innovation and systemic risk, both in stationary equilibrium and transition dynamics settings, and we generate a variety of new testable predictions. Three results summarize key highlights of these two chapters. First, we show that financial innovation can either increase or decrease systemic risk under some conditions. Second, we introduce a new notion of a *leverage trap*, showing that as the economy moves to high leverage systemic risk states, it tends to stay there. Third, we find that greater use of financial innovations after some shocks, when credit and the economy are already contracting, can slow the recovery process.

These chapters make two main contributions. First, we consider a different modeling framework for exploring the financial innovation role of the financial sector. This is a special case of the financial sector that is still largely unexplored in the literature, as opposed to the usual intermediary role. For this, we introduce some powerful methods from the theory of Lévy Processes. We also make additional methodological contributions, in terms of showing how to *analytically characterize transition dynamics* in the class of equilibrium heterogeneous agent models, which represents the evolutionary dynamics of systemic risk. For this, we introduce Feynman Path Integral method from quantum mechanics.

The generality and flexibility of the methods introduced in these chapters open the door to answering many newer types of questions for future research. In particular, our approach to analyzing systemic risk jointly through the leverage distribution across economic agents and insolvency considerations of the financial sector is unique to the literature and also provides new policy insights.

Further, the literature has largely relied on numerical computations to study equilibrium transition dynamics. This limitation can be largely overcome as shown here. Thus, the methods explored here should be easily applicable to a variety of economic settings. Further, the tools considered here complement the methods of spectral theory explored in Gabaix et al. (2016). For example, one application could be to analytical study the transition dynamics properties of recent changes to the U.S. tax laws that have broad implications for individuals and national/multinational firms. Another application could be to study the long run (systemic risk) effects of growing income/wealth inequality. Our model/methods can also be extended to study the dynamic general equilibrium pricing of counterparty risk and to understand how the need for financial innovation arises endogenously.

CHAPTER 4

ASSET PRICE BUBBLES, MARKET LIQUIDITY AND SYSTEMIC RISK

4.1 The Model

This section presents the details of the model.

4.1.1 The Market

We consider a discrete time economy with a finite horizon $t \in \{0, 1, ..., T\}$. The randomness in the economy is characterized by a given complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ where Ω is state space, $\mathbb{P}$ is the (statistical) probability measure, and $\mathbb{F} = \{\mathcal{F}\}_{t\in\{0,1,...,T\}}$ is the filtration with $\mathcal{F} = \mathcal{F}_T$. The economy is populated with heterogeneous agents of total mass one. These agents are partitioned into a finite number of types, each type with a strictly positive mass, indexed by $i = 1, \ldots, I$. The discrete mass of a type i agent is denoted $\mathbb{I}(i)$ for $i = 1, \ldots, I$ with $\sum_{i=1}^{I} \mathbb{I}(i) = 1$. Agent types are heterogeneous in their beliefs $\mathbb{P}^i$, wealth W^i (to be defined later), and their preferences $U_i : \mathbb{R} \times \Omega \to \mathbb{R}$ which are defined over terminal consumption. We also assume that $\mathbb{P}_i$ are equivalent to $\mathbb{P}$ for all i. This means that agents' beliefs and the statistical probability measure agree on zero probability events. The filtration $\mathbb{F}$ corresponds to the information set of each agent implying that there is symmetric information.

Agents can trade two assets: a riskless bond, represented by a money market account (mma) and a risky asset. The risky asset is arbitrary, so it can represent equity (stock), physical capital, or even real estate depending upon the market analyzed. Here, for convenience, we will use the term risky asset or stocks interchangeably. Borrowing and lending takes place through a bond market that is in zero net supply. Agents also face a borrowing constraint (discussed below). The risky asset is in positive supply, which is constant across time. Short sales are allowed, but margin requirements are imposed to insure that the short position can be covered at a future date. Without loss of generality, prices are normalized by the value of the mma, i.e. the mma is the numeraire. Alternatively, we can think of this numeraire as the units of consumption good redeemable at the last date in the model.

We assume that the stock pays no dividends over times $\{0, \ldots, T-1\}$, but it pays a liquidating dividend at time T.

Assumption. *(Liquidating Dividends)*

The risky asset S_t is assumed to have no cash flows (dividends) over times $t \in \{0, 1, ..., T\}$ and has an exogenous liquidating dividend at time T, i.e. there exists a $\xi : \Omega \to \mathbb{R}$ at time T such that $S_T = \xi > 0$.

We assume that the stock is in positive supply with $N > 0$ shares outstanding and that the mma is in zero net supply, for all times. At the terminal time T, all debts, if any, must be paid, all the remaining positions liquidated, and the proceeds used for consumption. Since all positions must be liquidated at time T, liquidity costs are necessarily incurred by agents prior to consumption. Time $T + 1$ is a non-trading date when consumption occurs.

4.1.2 The Liquidity Cost Process

Market liquidity/illiquidity is captured by a stochastic quantity impact on the price from trading, where the trade size affects the magnitude of the impact. In this regard, we use the liquidity cost process of Cetin, Jarrow, Protter (2004) as modified by Cetin and Rogers (2007). Our approach is closely related to market liquidity as defined in Brunnermeier and Pedersen (2008)[1].

Define $s_t(x, \omega)$ to be the *per share* market price of the stock for a trade of size $x \in \mathbb{R}$. Both purchases and short sales are allowed. This is sometimes called the supply curve. We assume that: (i) $s_t(x, \omega)$ is $\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}_t$ measurable for each t where $\mathcal{B}(\mathbb{R})$ is the Borel sigma-algebra on $\mathbb{R}$, (ii) $s_t(x, \omega) > 0$ in x for for all t a.e. $\mathbb{P}$, and (iii) $s_t(x, \omega)$ is strictly increasing in x for for all t a.e. $\mathbb{P}$.

Define $s_t(0, \omega) \equiv S_t$ to be the market price for zero trades. This corresponds to the market price in a world with no quantity impact from trading. S_t is the *marked-to-market* price of the stock. We add the following assumption to characterize the liquidity costs of trading.

Assumption. *(Liquidity Cost Process)*

Define $\varphi_t(x, \omega)$ to be the liquidity cost for selling/buying x shares at time t given ω, i.e.

$$xs_t(x, \omega) \equiv \varphi_t(x, \omega)s_t(0, \omega) = \varphi_t(x, \omega)S_t.$$

[1]See Nikolau (2009) for a discussion of various liquidity notions. Our liquidity cost can also be interpreted as an endogenous transaction cost. In equity markets this cost is incurred by having to sell below the fair/mid-price or buying above the fair/mid-price. In housing or real estate markets this reflects the costs associated with either buying or selling the real estate property. This implies that standard transaction costs are a special case of our market liquidity formulation.

We assume that[2]

1. *$\varphi_t(x, \omega)$ is $\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}_t$ measurable for each t, and*

2. *for a fixed $t \in \{0, 1, \ldots, T\}$ and $\omega \in \Omega$, $\varphi_t(x, \omega) : \mathbb{R} \to (-\infty, \infty]$ is strictly convex, strictly increasing where $\varphi_t(0) = 0$, and φ_t is differentiable for all x.*[3]

This is a very general liquidity cost process. The functional form of $\varphi_t(x)$ is given exogenously, but validated endogenously in equilibrium. By condition (2), the larger the quantity purchased, the larger the price paid per share. Conversely, the larger the quantity sold, the less the price received per share. This captures the inelastic nature of the supply curve for shares. The increasing condition, in conjunction with $\varphi_t(0) = 0$, implies that $\varphi_t(x) > 0$ for $x > 0$ and $\varphi_t(x) < 0$ for $x < 0$.

The convexity condition is needed to incorporate nonlinearities in liquidity costs. Indeed, when there is no quantity impact on the price from trading (the traditional model), the quantity impact function is linear in the trade size, i.e.

$$\varphi_t(x) = x.$$

The strict convexity assumption insures that the liquidity cost is larger than proportional as the trade size increases.

The marginal cost from trading dx additional shares in a trade of size x is

$$\frac{d\left(\varphi_t(x)S_t\right)}{dx} = \varphi_t'(x)S_t > 0.$$

[2]For simplicity of notation, we will often drop the dependence of φ_t on ω.

[3]This implies that $\tilde{S}_t(x, \omega) = xs_t(x, \omega)$ is convex, lower semicontinuous with $\tilde{S}_t(0, \omega) = 0$ for every ω. Hence, by Pennannen (2014), p. 747, $\tilde{S}_t$ is a $\mathcal{F}_t$ - measurable normal integrand.

This represents the *quoted* or *transaction price* when trading x shares. It is the price paid/received for the last share traded. In this representation there is no distinction between a quoted and transaction price, both are the same. Since $\varphi_t'(0) = 1$ and $\varphi_t(x)$ is convex, $\varphi_t'(x) > 1$ for $x > 0$ and $\varphi_t'(x) < 1$ for $x < 0$.

Given this interpretation, we see that $\varphi_t'(0)S_t = S_t$ corresponds to the transaction price when trading zero shares, or equivalently, the marked-to-market price. This implies that $\varphi_t'(0+)S_t$ corresponds to the ask price paid for buying 0+ (an infinitesimal quantity of) shares and $\varphi_t'(0-)S_t$ corresponds to the bid price received for selling 0− (an infinitesimal quantity of) shares. The condition $\varphi_t'(0) = 1$ along with the differentiability of $\varphi_t(x)$ for all x implies that $\varphi_t'(x)$ is continuous[4], hence $\varphi_t'(0+)S_t = S_t = \varphi_t'(0-)S_t$, i.e. the ask price equals the marked-to-market price which equals the bid price. Finally, note that if there is no quantity impact on the price from trading, then $\varphi'(x) = 1$ for all x, and the transaction and marked-to-market prices are always equal. An example helps to clarify this assumption.

Example. (Liquidity Cost Process) *Consider the following stochastic liquidity cost process:*

$$\varphi_t(x, \omega) = \frac{e^{\alpha_t(\omega)x} - 1}{\alpha_t(\omega)}$$

where $\alpha_t(\omega) > 0$ is $\mathcal{F}_t$ - measurable. It can be easily checked that this process satisfies all the previous assumptions. We see that

$$\varphi_t'(x, \omega) = e^{\alpha_t(\omega)x}$$

with $\varphi_t'(0) = 1$. This liquidity cost process is graphed in Figure (4.1). The 45 degree line corresponds to the no liquidity cost case where $\varphi_t(x, \omega) = x$ for all (t, ω). As we will

[4]See Rockafellar (1970), p. 246.

discuss in subsequent sections, the liquidity cost curves associated with higher $\alpha_t(\omega)$ are relatively more illiquid compared to the curves with lower $\alpha_t(\omega)$.

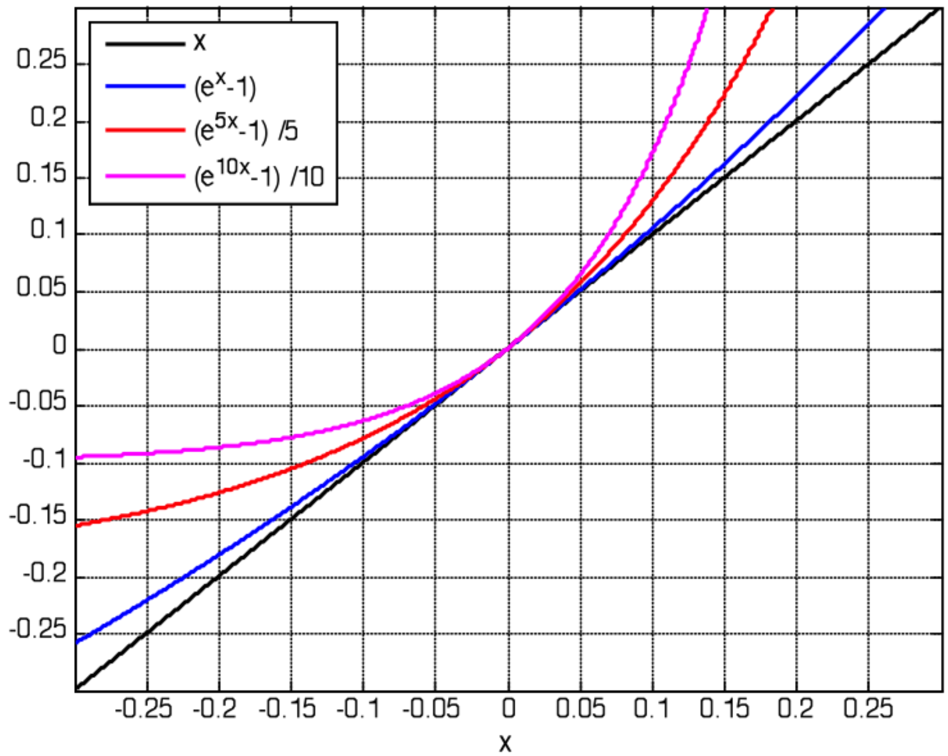


Figure 4.1: Liquidity Cost Functions

This liquidity cost function is analogous to the non-linear adjustment cost of investment or capital in the macroeconomics literature. Firms or households that own the capital face such costs when replacing old capital with new (physical) capital. Adjustment cost functions are often represented by $\varphi(I_t/k_t)$ where $\varphi(\delta) = 0$, $\varphi'(.) > 0$, $\varphi''(.) > 0$, and δ is the rate of capital depreciation. Here each unit of investment is transformed into less than one unit of capital, reflecting adjustment costs. $\varphi_t(.)$ usually enters the capital evolution process as $k_{t+1} = (1-\delta)k_t + I_t - k_t\varphi(I_t/k_t)$. In macro models, when this approach is used to analyze a firm's investment decision, we get Tobin's Q-theory, as in Tobin (1969) and Hayashi (1982). Further, such adjustment costs can also arise due to investment changes between periods as in Christiano, Eichenbaum and Evans (2005) where the investment adjustment cost function is $\varphi(I_t/I_{t-1})$ and in steady state $\varphi(1) = 0$, $\varphi'(1) = 0$, and $\varphi''(1) > 0$. Here, the capital accumulation is given as $k_{t+1} = (1 - \varphi(I_t/I_{t-1}))I_t + (1 - \delta)k_t$. The adjustment cost function is often assumed to be quadratic[5].

An advantage of using this liquidity cost process to characterize market liquidity is that it allows us to consider various asset market types with different liquidities. For example, equities markets are more liquid than housing markets. This difference is represented by different functional forms for the liquidity cost function $\varphi_t(x)$. We will exploit this benefit below. Additionally, this liquidity cost function can be used to generate liquidity policies in an analogous function to how the Taylor rule is used in monetary policy to set nominal interest rates depending on output/inflation gaps and the unemployment rate (see Eggertsson and Krugman (2012), Guerrieri and Iacoviello (2017); also see Christiano et al. (2010) for a review of DSGE models with monetary policy).

4.1.3 The Budget Set

In an asset pricing model, the budget set is characterized by the notion of a *self-financing trading strategy*. With a liquidity cost function, defining an agent's budget is subtle. This is because a trading strategy's wealth is not well defined.

[5] Bernanke, Gertler and Gilchrist (1999) use a similar non-linear adjustment cost to capital, allowing for dynamic amplification of a negative shock with $k_{t+1} = K_t\varphi(I_t/k_t) + (1-\delta)k_t$. Here, $\varphi(.)$ there is increasing and concave. Brunnermeier and Sannikov (2014) also use a similar approach interpreting the concavity of such an adjustment function as capturing a technological illiquidity when converting output to new capital and vice versa.

A trading strategy's wealth depends on the stock price, which depends on a trade quantity x. To avoid this ambiguity, we focus on the shares in the budget set.

A *trading strategy* is defined to be a $\mathcal{F}_{t-1}$ - measurable stochastic process (X_t, Y_t) representing the aggregate shares in stock and mma, given an initial position (x, y). The budget set, a *self-financing* trading strategy, for a generic (X_t, Y_t) can be written as

$$Y_{t+1} = Y_t - \Delta X_{t+1} s_t(\Delta X_{t+1}) \tag{4.1}$$

where $\Delta X_{t+1} = X_{t+1} - X_t$. Using the liquidity cost function we get

$$Y_{t+1} = Y_t - \varphi_t(\Delta X_{t+1}) S_t \tag{4.2}$$

Starting from (x, y), this is a difference equation whose solution is

$$Y_{t+1} = y - \sum_{j=0}^{t} \varphi_j(\Delta X_{j+1}) S_j. \tag{4.3}$$

By construction, the last trade occurs at time T, when the stock holdings are liquidated, i.e.

$$\Delta X_{T+1} = -x - \sum_{t=0}^{T-1} \Delta X_{t+1} = -X_T.$$

This implies that

$$Y_{T+1} = y - \varphi_T\left(-x - \sum_{t=0}^{T-1} \Delta X_{t+1}\right) S_T - \sum_{t=0}^{T-1} \varphi_t(\Delta X_{t+1}) S_t. \tag{4.4}$$

Given the self financing condition, we have that

$$\{(x, y), X_{t+1} \text{ for } t \in \{0, 1, \ldots, T-1\}\}$$

uniquely determines Y_{T+1}[6].

[6]Note that we do not need to include X_{T+1} in this expression because by construction it equals zero.

As seen above, the change in the budget set is determined by the change in the mma and risky asset as agents borrow and trade amongst each other. Later on we will see that this is related to the changes in the wealth of agents.

4.1.4 The Trading Constraints

We now characterize the trading constraints in the economy. This closely follows the trading constraints in Jarrow (2017b). We restrict the *self-financing* trading strategy (X_t, Y_t) to satisfy the following condition:

$$\{X_{t+1} \in K_t(\omega) \text{ for } t \in \{0, 1, \ldots, T-1\} \text{ a.e. } \mathbb{P}\}$$

where $K_t(\omega) \subset \mathbb{R}$ is a $\mathcal{F}_t$ - measurable[7], nonempty, closed, convex set with $0 \in K_t$ and $\{X_{t+1} = X_t\} \in int\,(K_t)$ for $t \in \{0, 1, \ldots, T-1\}$ a.e. $\mathbb{P}$ where $int(\cdot)$ denotes the interior in the usual topology on $\mathbb{R}$.

In this chapter, we represent the trading constraints in the economy by defining the set $K_t(\omega)$ as follows.[8].

Assumption. *(Trading Constraints)*

$$K_t(\omega) = \{X_{t+1} \in \mathbb{R} : \; Y_{t+1} \geq -S_t X_{t+1} \left(\gamma 1_{X_{t+1} \geq 0} + (1+\gamma) 1_{X_{t+1} \leq 0}\right)\}$$

where $0 \leq \gamma \leq 1$, and $K_t(\omega) \subset \mathbb{R}$ is a $\mathcal{F}_t$ measurable, nonempty, closed, convex set.

Proof. It is nonempty because $0 \in K_t(\omega)$. We can rewrite the constraint set as

[7] K_t being $\mathcal{F}_t$ - measurable means that $\{\omega \in \Omega : K_t(\omega) \cap A \neq \emptyset\} \in \mathcal{F}_t$ for every open set $A \subset \mathbb{R}^2$.

[8] See Jarrow (2017b) for some examples of different types of trading constraint sets that frequently arise in financial applications.

$$K_t(\omega) = \{X_{t+1} \geq 0 : \ Y_{t+1} \geq -\gamma S_t X_{t+1}\} + \{X_{t+1} \leq 0 : \ Y_{t+1} \geq -(1+\gamma) S_t X_{t+1}\}.$$

This set is closed and since each subset in the sum is convex, the sum is convex (see Ruszczynski (2006), p. 18). ■

Marked-to-market prices S_t are used to define the trading constraint because they reflect prices readily observable in the market. This is contrasted with transaction prices $\varphi_t'(x) S_t$ which depend on knowledge of the trade size and the liquidity cost paid via the liquidity cost function $\varphi_t(x)$. This constraint limits both borrowing and short selling. To see this, note that if the trading strategy shorts the risky asset ($X_{t+1} < 0$), then it simplifies to $Y_{t+1} \geq -(1+\gamma) S_t X_{t+1} > 0$. This implies that to cover the short position, a margin account must be held in the riskless asset consisting of the marked-to-market value of the short position plus a fraction $\gamma > 0$ more. Next, if the trading strategy is to buy the risky asset ($X_{t+1} > 0$), then it simplifies to $Y_{t+1} \geq -\gamma S_t X_{t+1}$ where $Y_{t+1} < 0$ is the borrowing. This implies that to finance the purchase, a borrowing of no more than a fraction $\gamma > 0$ of the market-to-market value of the long position is allowed.

Note that the trading constraint only applies for times $t \in \{0, 1, ..., T-1\}$ because at time T all the positions are liquidated. The assumption $\{0\} \in K_t$ implies that zero holdings in the stock always satisfies the borrowing constraint. $X_{t+1} = X_t \in int(K_t)$ means that no trading at any time t is always feasible. Here, the restrictions on the mma position are not explicit. However, the holdings in the mma are implicitly restricted because Y_{t+1} is uniquely determined by the position in the stock and the budget constraint.

Let us now define the normal cone to the set K_t for a given $\omega \in \Omega$ as

$$N_{K_t}(X) = \{\kappa \in \mathbb{R} : \kappa(Z - X) \geq 0 \quad \text{for all} \quad Z \in K_t\}$$

where $X \in K_t$ and $t \in \{0, 1, ..., T-1\}$. This set is $\mathcal{F}_t$ - measurable and by construction $N_{K_T}(X) = N_{\mathbb{R}}(X) = \{0\}$.

The following lemma will be needed to understand the shadow costs of the trading constraints in subsequent sections.

Lemma 2. *(Signs of the Elements $\kappa \in N_{K_t}(X_{t+1})$)*

Let X_{t+1} be a trading strategy and $\kappa \in N_{K_t}(X_{t+1})$.

(1) (Non-binding Constraint) If $X_{t+1} \in int(K_t)$, then $\kappa = 0$.

(2) (Binding Constraint) If $X_{t+1} \in bd(K_t)$ where $bd(\cdot)$ denotes the boundary in the usual topology on $\mathbb{R}$.

Given $\kappa \in N_{K_t}(X)$ with $\kappa \neq 0$. Then,

- *(a) if $X_{t+1} > 0$, then $\kappa > 0$.*
- *(b) if $X_{t+1} = 0$ and $[-\epsilon, 0] \subset K_t$ for $\varepsilon > 0$, then $\kappa > 0$.*
- *(c) if $X_{t+1} < 0$, then $\kappa < 0$.*
- *(d) if $X_{t+1} = 0$ and $[0, \varepsilon] \subset K_t$ for $\varepsilon > 0$, then $\kappa < 0$.*

Proof. If $X_{t+1} \in int(K_t)$, then $N_{K_t}(X_{t+1}) = \{0\}$, see Tuy (1998), p. 22. Hence, if $\kappa \in N_{K_t}(X_{t+1})$, then $\kappa = 0$.

(Case a) If $X_{t+1} \in bd(K_t)$, and $X_{t+1} > 0$, then by the convexity of K_t, $[0, X_{t+1}] \subset K_t$. This implies $N_{K_t}(X_{t+1}) \neq \{0\}$. Hence, if $\kappa \neq 0$, then $\kappa(Z - X_{t+1}) \leq 0$ for $0 \leq Z \leq X_{t+1}$, implying $\kappa > 0$.

(Case b) follows similarly.

(Case c) If $X_{t+1} \in bd(K_t)$, and $X_{t+1} < 0$, then by the convexity of K_t, $[X_{t+1}, 0] \subset K_t$. This implies $N_{K_t}(X_{t+1}) \neq \{0\}$. Hence, if $\kappa \neq 0$, then $\kappa(Z - X_{t+1}) \leq 0$ for $0 \leq X_{t+1} \leq Z$, implying $\kappa < 0$.

(Case d) follows similarly. This completes the proof. ■

This lemma shows that if X_{t+1} is on the boundary and cannot increase (cases (a) and (b)), then $\kappa \neq 0 \in N_{K_t}(X_{t+1})$ is strictly positive. Conversely, if X_{t+1} is on the boundary and cannot decrease (cases (c) and (d)), then $\kappa \neq 0 \in N_{K_t}(X_{t+1})$ is strictly negative. This abstraction will be explicitly used below to characterize bubble. A concrete version of this abstraction is the usual Lagrange multipliers in the optimization problems, where the multiplier $\kappa = 0$ means the constraint is not binding, and $\kappa \neq 0$ means the constraint is binding.

4.1.5 The Optimization Problem

As noted earlier, the economy is populated by a finite number of heterogeneous agents types indexed by i, with beliefs $\mathbb{P}_i$ and preferences $U_i : \mathbb{R} \times \Omega \to \mathbb{R}$ defined over terminal consumption. The agents of type i are initially endowed with shares in the stock and mma (x_i, y_i). We assume that for all $\omega \in \Omega$, $U_i(z, \omega)$ is strictly increasing, strictly concave, and satisfies the Inada conditions

$$\lim_{x \to -\infty} U_i'(z, \omega) = \infty, \quad \lim_{x \to 0} U_i'(z, \omega) = 0.$$

In addition, we assume that $U_i(0, \omega) = 0$ and that $U_i(z, \omega), U_i'(z, \omega)$ are bounded above by $\mathbb{P}_i$ integrable random variables (independent of z) for all i.[9]

Agents choose shares in the stock and mma, subject to borrowing constraints, to maximize their expected utility of terminal consumption:

$$u_i(x, y) = \sup_{\{X_{t+1} \in K_t \colon t \in \{0,1,\ldots,T-1\}\}} E^{\mathbb{P}_i}[U_i(Y_{T+1})] \tag{4.5}$$

where expectation is taken under $\mathbb{P}_i$.

4.1.6 Equilibrium

In an equilibrium, agents maximize their expected utility subject to their constraints, such that the price of risky asset is endogenously determined in the equilibrium. The aggregate holdings in stocks and mma are represented by the expectation with respect to the distribution $\mathbb{I}(i)$ across agents. To formalize this description, we need some definitions.

Definition 2. (An Asset Market)

An asset market *is a collection* $\left((\varphi, K, N), (\mathbb{P}_i, U_i, (x_i, y_i))_{i=1}^{I}\right)$

where $\mathbb{E}^{\mathbb{I}}[y^i] = 0$ *and* $\mathbb{E}^{\mathbb{I}}[x^i] = N$.

As defined, an asset market is a liquidity cost process, a trading constraint, a supply of shares outstanding i.e. (φ, K, N) and a set of economic agents $(\mathbb{P}_i, U_i, (x_i, y_i))$ of total mass one. The mma is in zero net supply and the supply

[9] These assumption imply that $U_i(z, \omega)$ is a normal integrand on $\mathbb{R} \times \Omega$ and when taking the derivative of $E^i[U_i(z)]$ with respect to z, one can exchange the expectation and derivative operators.

of the risky asset is strictly positive and equal to N. By construction, the total supply of shares and mma endowed at time 0 must equal the total supply of shares and mma outstanding. Finally, the underlying filtration and probability measure $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\{0,\dots,T\}}, \mathbb{P})$ are implicit in the specification of a market.

Definition 3. (Equilibrium)

Given an asset market $\left((\varphi, K, N), (\mathbb{P}_i, U_i, (x_i, y_i))_{i=1}^{I}\right)$ *with* $\mathbb{E}^{\mathbb{I}}[y^i] = 0$ *and* $\mathbb{E}^{\mathbb{I}}[x^i] = N$.

An equilibrium *is a price process S and risky asset demands* $(X^i_{t+1} : t \in \{0, \dots, T-1\})_{i=1}^{I}$ *a.e.* $\mathbb{P}$ *such that*

(i). $(X^i_{t+1} \in K_t : t \in \{0, \dots, T-1\})$ *are optimal for all i, and*

(ii) $\mathbb{E}^{\mathbb{I}}[\Delta X^i_{t+1}] = 0$ *for* $t \in \{0, \dots, T-1\}$ *a.e.* $\mathbb{P}$ *such that the market clears i.e. supply equals demand.*

There is no trading at time $T+1$.

This is standard definition of a dynamic competitive Radner equilibrium (Radner, 1982). Because the supply of shares is constant N for all times, in conjunction with $\mathbb{E}^{\mathbb{I}}[x^i] = \mathbb{E}^{\mathbb{I}}[X^i_0] = N$ this implies that $\mathbb{E}^{\mathbb{I}}[X^i_{t+1}] = N$ for $t \in \{0, \dots, T-1\}$. Note that by the budget/self-financing condition (4.1), this implies that mma market decreases in aggregate value if $\Delta X^i_{t+1} \neq 0$ for some i, i.e.

$$\mathbb{E}^{\mathbb{I}}[\Delta Y^i_{t+1}] = -\mathbb{E}^{\mathbb{I}}[\varphi_t(\Delta X^i_{t+1}) S_t] < 0 \quad \text{for } t \in \{0, \dots, T-1\} \text{ a.e. } \mathbb{P}.$$

This strict decrease in value follows because if $\Delta X^i_{t+1} \neq 0$ for some i, then by the strict convexity of $\varphi_t(\Delta X^i_{t+1})$, $\varphi_t(\Delta X^i_{t+1}) > \varphi_t(0) + \varphi_t'(0)\Delta X^i_{t+1}$. The fact that $\varphi_t(0) = 0$ in conjunction with $\mathbb{E}^{\mathbb{I}}[\Delta X^i_{t+1}] = 0$ yields

$$-\mathbb{E}^{\mathbb{I}}[\varphi_t(\Delta X^i_{t+1}) S_t] < -\varphi_t'(0) S_t \mathbb{E}^{\mathbb{I}}[\Delta X^i_{t+1}] = 0 \quad \text{for } t \in \{0, \dots, T-1\} \text{ a.e. } \mathbb{P}.$$

Thus, the decrease in the value of the aggregate mma market captures the the liquidity costs incurred by the aggregate trading in the economy. The parties to whom the liquidity costs are paid are the market makers in the exchange market. The equilibrium in the exchange market endogenously determines the equilibrium liquidity cost function φ. For the details of the exchange market equilibrium see Jarrow (2017b). In this chapter we take the liquidity cost function as given, but the functional form is itself determined in equilibrium.

4.2 Bubbles and Market Liquidity

This section characterizes equilibrium asset price bubbles and their relation to market liquidity. First, it can be shown, under a set of sufficient conditions, that there exists an equilibrium for the economy previously described (see Jarrow (2017b)). For the subsequent analysis, therefore, we assume the existence of such an equilibrium and we characterize this equilibrium.

4.2.1 Characterizing Asset Price Bubbles

We start with a theorem characterizing an agent's first order conditions. This characterization is given in the following theorem.

Theorem 8. *(Individual Agent's Optimal allocation)*

Let S be an equilibrium price process.

Then, there exists a unique optimal trading strategy $\{X_{t+1} \in K_t : \ t \in \{0, \ldots, T-1\}\}$

characterized by the following equations

$$0 \in \partial E_t^{\mathbb{P}_i}\left[U_i\left(Y_{T+1}^i\right)\right] + N_{K_t}(X_{t+1}^i) \tag{4.6}$$

for $t \in \{0, \ldots, T-1\}$ *a.e.* $\mathbb{P}_i$

where[10] $Y_{T+1}^i = y_i - \varphi_T\left(-x_i - \sum_{t=0}^{T-1} \Delta X_{t+1}^i\right) S_T - \sum_{t=0}^{T-1} \varphi_t(\Delta X_{t+1}^i) S_t$, $\Delta X_{T+1}^i = -X_T^i$, $\partial(\cdot)$ *is the subdifferential, and* $E_t^{\mathbb{P}_i}[\cdot] = E^{\mathbb{P}_i}[\cdot|\mathcal{F}_t]$.

Or equivalently,

$\varphi_t'(\Delta X_{t+1}^i)(S_t - \nu_t^i)$ *is a martingale under*

$$\frac{d\mathbb{Q}_i}{d\mathbb{P}_i} = \frac{U_i'(Y_{T+1}^i)}{E^{\mathbb{P}_i}\left[U_i'(Y_{T+1}^i)\right]} > 0$$

for $t \in \{0, \ldots, T\}$ *where* $\nu_t^i = \frac{\kappa_t^i}{E_t^{\mathbb{P}_i}\left[U_i'(Y_{T+1}^i)\right]\varphi_t'(\Delta X_{t+1}^i)}$ *and* $\kappa_t^i \in N_{K_t}(X_{t+1}^i)$.

Proof. (Step 1) *Existence and Uniqueness*

The above assumptions in conjunction with Examples 2.1 and 5.2 of Pennanen (2014) imply that the hypothesis of Theorem 5.1 in Pennanen (2014) hold. This theorem states that an optimal trading strategy exists. It is unique by the concavity of the utility function.

(Step 2) *Characterization of the Solution*

We use backward induction.

At time T, with share holdings (X_T^i, Y_T^i), the optimal trading strategy is $\Delta X_{T+1} = -X_T$ since the portfolio must be liquidated.

[10]Note that when $t = T$, expression (4.6) is identically zero.

At time $t < T$ with share holdings (X_t^i, Y_t^i), having determined the optimal $\{X_{j+1}^i : j \in \{t+1, \ldots, T-1\}\}$, the optimal $Z \in K_t$ must maximize

$$E_t^{\mathbb{P}_i}\left[U_i\left(Y_{T+1}^i(Z)\right)\right] = E_t^{\mathbb{P}_i}\Big[U_i\Big(y_i - \varphi_T(-\textstyle\sum_{j=t+1}^{T-1} \Delta X_{j+1}^i - Z)S_T$$
$$- \textstyle\sum_{j=t+1}^{T-1} \varphi_j(\Delta X_{j+1}^i)S_j - \varphi_t(Z - X_t)\,S_t + (Y_t^i - y_i)\Big)\Big].$$

The first order condition, which is necessary and sufficient (see Tuy (1998), p. 75), is that

$$0 \in \partial E_t^{\mathbb{P}_i}\left[U_i\left(Y_{T+1}^i(X_{t+1})\right)\right] + N_{K_t}(X_{t+1}^i).$$

But,

$$\partial E_t^{\mathbb{P}_i}\left[U_i\left(Y_{T+1}^i(X_{t+1})\right)\right] = \frac{dE_t^{\mathbb{P}_i}\left[U_i\left(Y_{T+1}^i(X_{t+1})\right)\right]}{dX_{t+1}^i}$$

$$= E_t^{\mathbb{P}_i}\left[U_i'(Y_{T+1}^i)\left(\varphi_T'(\Delta X_{T+1}^i)S_T - \varphi_t'(\Delta X_{t+1}^i)S_t\right)\right].$$

Hence, there exists a $\kappa_t^i \in N_{K_t}(X_{t+1}^i)$ such that

$$E_t^{\mathbb{P}_i}\left[U_i'\left(Y_{T+1}^i\right)\left(\varphi_T'(\Delta X_{T+1}^i)S_T - \varphi_t'(\Delta X_{t+1}^i)S_t\right)\right] + \kappa_t = 0. \tag{4.7}$$

$$\frac{E_t^{\mathbb{P}_i}\left[U_i'\left(Y_{T+1}^i\right)\left(\varphi_T'(\Delta X_{T+1}^i)S_T - \varphi_t'(\Delta X_{t+1}^i)S_t\right)\right]}{E_t^{\mathbb{P}_i}\left[U_i'(Y_{T+1}^i)\right]} + \frac{\kappa_t}{E^{\mathbb{P}_i}\left[U_i'(Y_{T+1}^i)\right]} = 0.$$

$$E_t^{\mathbb{Q}_i}\left[\varphi_T'(\Delta X_{T+1}^i)S_T - \varphi_t'(\Delta X_{t+1}^i)S_t\right] + \frac{\kappa_t}{E_t^{\mathbb{P}_i}\left[U_i'(Y_{T+1}^i)\right]}\frac{\varphi_t'(\Delta X_{t+1}^i)}{\varphi_t'(\Delta X_{t+1}^i)} = 0$$

where $E_t^{\mathbb{Q}_i}[\cdot] = E^{\mathbb{Q}_i}[\cdot|\mathcal{F}_t]$ is conditional expectation under $\mathbb{Q}_i$.

$$E_t^{\mathbb{Q}_i}\left[\varphi_T'(\Delta X_{T+1}^i)S_T - \varphi_t'(\Delta X_{t+1}^i)S_t + \varphi_t'(\Delta X_{t+1}^i)\frac{\kappa_t}{E_t^{\mathbb{P}_i}\left[U_i'(Y_{T+1}^i)\right]\varphi_t'(\Delta X_{t+1}^i)}\right] = 0.$$

Define $v_t^i = \frac{\kappa_t^i}{E_t^{\mathbb{P}_i}\left[U_i'(Y_{T+1}^i)\right]\varphi_t'(\Delta X_{t+1}^i)}$. Note that $\kappa_T^i = 0$ since $N_{K_T}(X) = \{0\}$ for all $X \in K_T$. Then,

$$E_t^{\mathbb{Q}_i}\left[\varphi_T'(\Delta X_{T+1}^i)(S_T - v_T^i) - \varphi_t'(\Delta X_{t+1}^i)(S_t - v_t^i)\right] = 0.$$

This implies that $\varphi_t'(\Delta X_{t+1}^i)(S_t - v_t^i)$ is a martingale for $t \in \{0, \ldots, T\}$ under $\frac{d\mathbb{Q}_i}{d\mathbb{P}_i} = \frac{U_i'(Y_{T+1}^i)}{E^{\mathbb{P}_i}\left[U_i'(Y_{T+1}^i)\right]} > 0$. This completes the proof. ■

Given a characterization of the agent's first order conditions in equilibrium, we now turn to a characterization of asset price bubbles. We first need to define an asset's fundamental value.

Definition 4. (Fundamental Value)

The fundamental value *of an asset to a type i agent is*

$$E_t^{\mathbb{Q}_i}[\varphi_T'(\Delta X_{T+1}^i)S_T]$$

where $E_t^{\mathbb{Q}_i}[.] = E^{\mathbb{Q}_i}[.|\mathcal{F}_t]$ *is conditional expectation taken under* $\mathbb{Q}_i$.

The fundamental value is the time t expected liquidation value of the agent's portfolio at time T, given their optimal trading strategy. This liquidation value reflects *the transaction or market price for the last shares traded* by the agent, i.e. $\varphi_T'(\Delta X_{T+1}^i)S_T$. The personalized state price density $\frac{d\mathbb{Q}_i}{d\mathbb{P}_i}$ is used to value the future transaction because it includes an adjustment for risk.[11] The standard definition of a price bubble now follows.

Definition 5. (Asset Price Bubble)

The asset price bubble for a type i agent is

$$\beta_t^i = \varphi_t'(\Delta X_{t+1}^i)S_t - E_t^{\mathbb{Q}_i}[\varphi_T'(\Delta X_{T+1}^i)S_T].$$

From this definition it is clear that a price bubble is related to the beliefs of an individual agent which translates into their personalized fundamental value for the risky asset. Consequently, some agents may see a price bubble based on their beliefs, while others may not. Note that if there were no liquidity costs to

[11]The state price density is also called the stochastic discount factor.

trading shares, i.e. $\varphi_t'(\Delta X_{t+1}^i) = 1$, then the market and transaction prices would equal S_t, and this definition collapses to the standard definition of an asset price bubble appearing in the literature (see Jarrow (2015)). We can now characterize an agent's price bubble.

Corollary 8.1. *(Bubbles in Equilibrium)*

For all $i = 1, \ldots, n$ and $t = 1, \ldots, T$,

$$\beta_t^i = \frac{-\kappa_t^i}{E_t^{\mathbb{P}_i}\left[U_i'\left(Y_{T+1}^i\right)\right]\varphi_t'(\Delta X_{t+1}^i)} \tag{4.8}$$

for $\kappa_t^i \in N_{K_t}(X_{t+1}^i)$.

If the constraint is non-binding, then $\beta_t^i = 0$. No asset price bubble.

If the constraint is binding, then

(a) if $X_{t+1} > 0$, then $\beta_t^i < 0$.

(b) if $X_{t+1} = 0$ and $[-\epsilon, 0] \subset K_t$ for $\varepsilon > 0$, then $\beta_t^i < 0$.

In cases (a) and (b), the transaction price for the last share traded is a $\mathbb{Q}_i$ submartingale.

(c) if $X_{t+1} < 0$, then $\beta_t^i > 0$.

(d) if $X_{t+1} = 0$ and $[0, \varepsilon] \subset K_t$ for $\varepsilon > 0$, then $\beta_t^i > 0$.

In cases (c) and (d) the transaction price for the last share traded is a $\mathbb{Q}_i$ supermartingale.

Proof. Given $E_t^{\mathbb{P}_i}\left[\varphi_T'(\Delta X_{T+1}^i)(S_T - v_T^i) - \varphi_t'(\Delta X_{t+1}^i)(S_t - v_t^i)\right] = 0.$

$$E_t^{\mathbb{P}_i}\left[\varphi_T'(\Delta X_{T+1}^i)S_T\right] - \varphi_t'(\Delta X_{t+1}^i)S_t + \varphi_t'(\Delta X_{t+1}^i)v_t^i = 0.$$

Algebra yields

$$\varphi_t'(\Delta X_{t+1}^i)S_t - E_t^{\mathbb{P}_i}\left[\varphi_T'(\Delta X_{T+1}^i)S_T\right] = -\varphi_t'(\Delta X_{t+1}^i)v_t^i, \text{ i.e.}$$

$\beta_t^i = -\varphi_t'(\Delta X_{t+1}^i)v_t^i$. But, $v_t^i = \frac{\kappa_t^i}{E_t^{\mathbb{P}_i}\left[U_i'(Y_{T+1}^i)\right]\varphi_t'(\Delta X_{t+1}^i)}$.

Hence $\beta_t^i = \frac{-\kappa_t^i}{E_t^{\mathbb{P}_i}\left[U_i'(Y_{T+1}^i)\right]\varphi_t'(\Delta X_{t+1}^i)}$.

Using Lemma 2 completes the proof. ∎

This corollary characterizes an individual agent's asset price bubble. If trading constraints are non-binding and $\kappa_t^i = 0$, then there are no asset price bubbles. However, if trading constraints are binding with $\kappa_t^i \gtrless 0$, then there are four cases. With respect to cases (a) and (b), the trader desires to buy more of the stock, but is constrained. Hence, the stock price is too low relative to the trader's valuation given their state price density. Here the asset price bubble is negative (the stock is undervalued) and the stock price process is a $\mathbb{Q}_i$ submartingale. With respect to cases (c) and (d), the trader desires to short more of the stock, but is constrained. Hence, the stock price is too high relative to the trader's valuation given their state price density. Here the asset price bubble is positive (the stock is overvalued) and the stock price process is a $\mathbb{Q}_i$ supermartingale. This is analogous to the characterization of asset price bubbles in continuous time, competitive, and frictionless markets (see Jarrow (2015)).

This corollary can also be used to understand the birth and death of asset price bubbles in this market. In an asset market with no price bubbles, a bubble starts if trading constraints suddenly become binding. This might occur because of exogenous shocks to the constraints imposed by regulators, or because beliefs and or preferences change so that a previously non-binding constraint becomes binding. Conversely, analogous random shocks to the market can also cause bubbles to burst.

4.2.2 Characterizing Market Liquidity

Now we move on to characterize the relation between bubbles and market liquidity. This characterization requires another definition.

Definition 6. (Liquid versus Illiquid Markets)

For a given t, let φ_t^l and φ_t^{nl} be two liquidity cost processes associated with different market types that satisfy the assumptions given above, and such that

$$\begin{array}{lll} \frac{d\varphi_t^l(x)}{dx} < \frac{d\varphi_t^{nl}(x)}{dx} & if & x > 0 \\ \frac{d\varphi_t^{nl}(x)}{dx} < \frac{d\varphi_t^l(x)}{dx} & if & x < 0 \end{array} \tag{4.9}$$

for all $x \neq 0$ a.e. $\mathbb{P}$. Then, we say that the time t asset market with liquidity cost $\varphi_t^{nl}(x)$ is illiquid (not liquid) relative to the market with liquidity cost $\varphi_t^l(x)$.

As defined above, a market is more liquid the smaller the transaction price paid $\frac{d(\varphi_t(x)S_t)}{dx} = \varphi_t'(x)S_t$ for a purchase $x > 0$, and the larger the transaction price received for a sale $x < 0$. Given the marked-to-market value of the stock, this is measured by $\varphi_t'(x) \neq 1$, hence the definition. It is easy to show that the more liquid the market, the smaller is the liquidity cost function for a trade of size $x \neq 0$.

Lemma 3. *(Price Impact Costs)*

Let the liquidity cost function φ_t^l be more liquid than φ_t^{nl}. Then

$$\varphi_t^l(x) < \varphi_t^{nl}(x) \tag{4.10}$$

for all $x \neq 0$ a.e. $\mathbb{P}$.

Proof. Using the fundamental theorem of calculus and the assumption that $\varphi_t^k(0) = 0$, we have $\varphi_t^k(x) = \int_0^x \frac{d\varphi_t^k(u)}{du} du$ for $k \in \{l, nl\}$ for $x > 0$ and $\varphi_t^k(x) = -\int_0^x \frac{d\varphi_t^k(u)}{du} du$ for $k \in \{l, nl\}$ for $x < 0$. The result follows because the integral is a positive linear operator. ∎

We have already discussed asset markets that can have different liquidities (say housing versus equity markets). The implication of the above lemma is that, if an agent is buying or selling (shorting) a risky asset in a relatively illiquid market, the liquidity cost is higher, i.e. an illiquid market is unfavorable to buyers or sellers in terms of overall cost. This observation is central to many of the results below. This lemma is also analogous to the liquidity cost function example in Figure (4.1). Larger values of $\alpha_t(\omega)$ represent more illiquid markets. With above insights, we now analyze the nature of bubbles in liquid versus illiquid markets.

Theorem 9. *(Bubbles in Liquid vs. Illiquid Markets)*

(Borrowing Constrained Bubbles, $\beta_t^i < 0$)

If $X_{t+1}^i > 0$, then $\beta_t^i(\varphi_t^l) < \beta_t^i(\varphi_t^{nl}) < 0$, i.e. bubbles in liquid markets are larger in absolute value, all else constant.

(Short Selling Constrained Bubbles, $\beta_t^i > 0$)

If $X_{t+1}^i < 0$, then $0 < \beta_t^i(\varphi_t^l) < \beta_t^i(\varphi_t^{nl})$, i.e. bubbles in illiquid markets are larger, all else constant.

Proof. From expression (4.8),

$$\beta_t^i = \frac{-\kappa_t^i}{E_t^{\mathbb{P}_i}\left[U_i'\left(Y_{T+1}^i\right)\right]\varphi_t'(\Delta X_{t+1}^i)}.$$

Assume that Y_{T+1}^i, κ_t^i, X_{t+1}, and ΔX_{t+1}^i are fixed. There are two cases to consider.

(Case 1) Assume that $X_{t+1}^i > 0$, then we know from above $\kappa_t^i > 0$, and $\Delta X_{t+1}^i > 0$. Thus, $\beta_t^i < 0$ since $\varphi_t'(\Delta X_{t+1}^i) > 0$ for all ΔX_{t+1}^i.

By the definition (6), $1 > \frac{d\varphi_t^{nl}(x)}{dx} > \frac{d\varphi_t^{l}(x)}{dx} > 0$ for $x > 0$. This means the denominator is larger for φ_t^{nl} than for φ_t^l.

Thus, the absolute value of the bubble $\left|\beta_t^i\right|$ is larger for φ_t^l, implying $\beta_t^i(\varphi_t^l) < 0$ is a larger negative number.

(Case 2) Assume that $X_{t+1}^i < 0$, then we know from above $\kappa_t^i < 0$, and $\Delta X_{t+1}^i < 0$. Thus, $\beta_t^i > 0$ since $\varphi_t'(\Delta X_{t+1}^i) > 0$ for all ΔX_{t+1}^i.

By the definition (6), $1 < \frac{d\varphi_t^{nl}(x)}{dx} < \frac{d\varphi_t^{l}(x)}{dx}$ for $x < 0$. This means that the denominator is smaller for φ_t^{nl} than for φ_t^l.

Thus, the bubble $\beta_t^i > 0$ is larger for φ_t^{nl}. ■

This theorem shows that market liquidity and the magnitude of asset price bubbles have an asymmetric relation depending on the sign of the price bubble. When the risky asset is undervalued due to borrowing constraints and the price bubble is negative, bubbles are smaller in illiquid markets. In contrast, when the risky asset is overvalued due to short selling constraints and the price bubble is positive, bubbles are larger in illiquid markets. This asymmetry is due to the difference in the agents' actions to exploit the perceived mispricing and the asymmetry in the liquidity cost function's transaction price $\varphi_t'(x)S_t$ across the two actions.

When agents view the risky asset as undervalued, they borrow to *purchase* the asset. The more risky assets they purchase, the *larger* the transaction price paid per share in an illiquid market. Due to the convexity of $\varphi_t(x)$, this increase in the transaction price occurs at an increasing rate in an illiquid market relative to a liquid market. This increasing marginal cost of a purchase at an increasing rate, relative to the marginal benefit attained, reduces the optimal constrained purchase quantity in an illiquid market relative to a liquid market, consequently reducing the size of the price bubble in a more illiquid market.

In contrast, when agents' view the risky asset as overvalued, they want to *short sell* the risky asset and provide margin to cover the short position. The more assets they short sell, the *smaller* the transaction price proceeds received per share in an illiquid market. Due to the convexity of $\varphi_t(x)$, this decrease in the transaction price is at an decreasing rate in an illiquid market relative to a liquid market. This decreasing marginal cost of a sale at a decreasing rate, relative to the marginal benefit attained, increases the optimal constrained sale quantity in an illiquid market relative to a liquid market, consequently increasing the size of the price bubble in a more illiquid market.

We now move on to characterize the percentage of agents in an economy that see bubbles at any given time. A direct application of theorem 9 yields the following result.

Theorem 10. *(Liquidity and Size of the Price Bubble Distribution)*

Let $\mathbb{I}^-_{\beta_t} \equiv \mathbb{I}(i : \beta^i_t < 0)$ *denote the percentage of agents in the economy that see negative* *bubbles at time t, and* $\mathbb{I}^+_{\beta_t} \equiv \mathbb{I}(i : \beta^i_t > 0)$ *denote the percentage of agents in the economy that see positive bubbles at time t. Then,*

(i) $\mathbb{I}^-_{\beta_t}(\varphi^l_t) > \mathbb{I}^-_{\beta_t}(\varphi^{nl}_t)$*, i.e. the percentage of agents viewing negative bubbles increase in more liquid markets, and*

(ii) $\mathbb{I}^+_{\beta_t}(\varphi^l_t) < \mathbb{I}^+_{\beta_t}(\varphi^{nl}_t)$*, i.e. the percentage of agents viewing positive bubbles increase in more illiquid markets.*

Proof. This follows from the following facts. For case (i), if $X^i_{t+1} > 0$, then $\beta^i_t(\varphi^l_t) < \beta^i_t(\varphi^{nl}_t) < 0$. Thus, more agents will see bubbles in liquid markets, everything else constant, because there will be some agents with $\beta^i_t(\varphi^{nl}_t) = 0$ but $\beta^i_t(\varphi^l_t) < 0$. The same argument holds for case (ii). ∎

This theorem states that when market liquidity declines in a given asset market, the percentage of agents seeing price bubbles declines as well. Consequently, theorems (9) and (10) have policy implications with respect to the existence of asset price bubbles. This is especially true for monetary policy, which can immediately affect market liquidity. In our model, in equilibrium, bubbles depend on agents' beliefs and preferences. This means that what is relevant for macro-policy is the *percentage of agents in the economy that see bubbles*, and not whether a uniform price bubble exists. This frees policy makers from determining the existence and magnitude of a uniform price bubble. Instead, policy makers can simply monitor the financial markets and statistically infer traders' beliefs concerning bubbles and react accordingly.

While a separate model is needed to fully explore the policy implications we have just discussed, our results provide a framework for analyzing these issues. In the next section of this chapter we show a direct relation between the percentage of agents in the economy viewing price bubbles and systemic risk. We now explore the relation between the magnitude of the trading constraints and the size of asset price bubbles.

4.2.3 Characterizing the Shadow Price of Trading Constraints

This subsection characterizes the shadow price of the trading constraints in equilibrium.

Theorem 11. *(Trading Constraint and Shadow Price)*

Let the trading constraint be binding at time t.

Then, as γ increases,

(i) $\kappa_t^i > 0$ decreases when $X_{t+1}^i > 0$, and

(ii) $\kappa_t^i < 0$ increases in absolute value when $X_{t+1}^i < 0$.

Proof. At time t, the optimal share holdings are either $X_{t+1}^i > 0$ or $X_{t+1}^i < 0$. Since the constraint is binding at the optimum, when $X_{t+1}^i > 0$, $Y_{t+1}^i < 0$ and when $X_{t+1}^i < 0$, $Y_{t+1}^i > 0$. This follows from our trading constraint set $K_t(\omega)$ as discussed before. Consequently, at time t, only one side of the constraint

$$K_t(\omega) = \{X_{t+1} \in \mathbb{R} : \ Y_{t+1} \geq -S_t X_{t+1} \left(\gamma 1_{X_{t+1} \geq 0} + (1+\gamma) 1_{X_{t+1} \leq 0}\right)\}$$

is binding for each agent i. Thus, in the time t optimization problem, the original constraint can be replaced by the constraint

$$K_t(\omega) = \{X_{t+1} \in \mathbb{R}_{++} : -\gamma S_t X_{t+1} \leq Y_{t+1}\} = \left\{X_{t+1} \in \mathbb{R}_{++} : -\frac{Y_{t+1}}{S_t X_{t+1}} \leq \gamma\right\}$$

for $X_{t+1} > 0$, and

$$K_t(\omega) = \{X_{t+1} \in \mathbb{R}_{--} : -(1+\gamma) S_t X_{t+1} \leq Y_{t+1}\} = \left\{X_{t+1} \in \mathbb{R}_{--} : \frac{Y_{t+1}}{S_t X_{t+1}} \leq -(1+\gamma)\right\}$$

for $X_{t+1} < 0$ without changing the optimal solution.

Note that $Y_{t+1} = Y_t - \varphi_t(X_{t+1} - X_t)S_t$ from expression (4.2).

Define the function $f(X_{t+1}) = -\left(\frac{Y_t - \varphi_t(X_{t+1} - X_t)S_t}{S_t X_{t+1}}\right)$ on $X_{t+1} \in \mathbb{R}$.

We note that

$f'(X_{t+1}) = \frac{Y_t - \varphi_t(X_{t+1} - X_t)S_t}{S_t X_{t+1}^2} + \frac{\varphi_t'(X_{t+1} - X_t)S_t}{S_t X_{t+1}} = \frac{Y_{t+1} + \varphi_t'(X_{t+1} - X_t)S_t X_{t+1}}{S_t X_{t+1}^2}$.

We have, $f'(X_{t+1}) > 0$ when $X_{t+1} > 0$.

Indeed, $\varphi_t'(X_{t+1} - X_t)S_t X_{t+1} > \gamma S_t X_{t+1}$ when $X_{t+1} > 0$ because $\varphi_t'(x) > 1 > \gamma$ when $x > 0$.

Thus, $f'(X_{t+1}) = \frac{Y_{t+1} + \varphi_t'(X_{t+1} - X_t)S_t X_{t+1}}{S_t X_{t+1}^2} > \frac{Y_{t+1} + \gamma S_t X_{t+1}}{S_t X_{t+1}^2} \geq 0$ because $Y_{t+1} + \gamma S_t X_{t+1} \geq 0$ when $X_{t+1} > 0$.

We have, $f'(X_{t+1}) > 0$ when $X_{t+1} < 0$.

Indeed, $(1+\gamma) S_t X_{t+1} < \varphi_t'(X_{t+1} - X_t)S_t X_{t+1}$ when $X_{t+1} < 0$ because $\varphi_t'(x) < 1 < 1 + \gamma$ when $x < 0$.

Thus, $f'(X_{t+1}) = \frac{Y_{t+1} + \varphi_t'(X_{t+1} - X_t)S_t X_{t+1}}{S_t X_{t+1}^2} > \frac{Y_{t+1} + (1+\gamma)S_t X_{t+1}}{S_t X_{t+1}^2} \geq 0$ because $Y_{t+1} + (1+\gamma) S_t X_{t+1} \geq 0$ when $X_{t+1} < 0$.

We can rewrite the time t constraint as:

(i) $f(X_{t+1}) \leq \gamma$ when the optimum is $X_{t+1} > 0$, and

(ii) $g(X_{t+1}) \leq \delta$ when the optimum is $X_{t+1} < 0$ with $g(X_{t+1}) = -f(X_{t+1})$ and $\delta = -(1+\gamma)$.

(Case 1) At time t the optimum is $X^i_{t+1} > 0$ with a binding constraint and $f'(X_{t+1}) > 0$.

Looking at the proof of Theorem 8, we see that at time t the optimization problem is

$$v(\gamma) = \sup_{\{X_{t+1} \in \mathbb{R}\}} E^{\mathbb{P}_i}[U_i(Y_{T+1}(X_{t+1})] \quad \text{subject to} \quad f(X_{t+1}) = \gamma$$

where the constraint is binding. As written, the Lagrangian is

$$\mathscr{L} = E^{\mathbb{P}_i}\left[U_i\left(Y^i_{T+1}(X_{t+1})\right)\right] + \lambda_t\left(f(X_{t+1}) - \gamma\right).$$

The first order (necessary and sufficient) condition is

$$\frac{dE^{\mathbb{P}_i}\left[U_i\left(Y^i_{T+1}(X_{t+1})\right)\right]}{dX_{t+1}} + \lambda_t f'(X_{t+1}) = 0.$$

From the proof of Theorem 8, expression (4.7), we have that $\kappa_t = \lambda_t f'(X_{t+1})$, i.e. $\lambda_t = \frac{\kappa_t}{f'(X_{t+1})}$.

But, standard results yield $v'(\gamma) = \lambda_t$, see Holmes (1970), p. 39. Hence, $v'(\gamma) = \lambda_t = \frac{\kappa_t}{f'(X_{t+1})}$. By Lemma 2, $\kappa_t > 0$ for $X_{t+1} > 0$, and $f'(X_{t+1}) > 0$ implies that $\lambda_t > 0$.

Also, it is well known that the value function is concave (in the parameter γ), see Holmes (1970), p. 37. Hence, $v''(\gamma) < 0$, which implies that as γ increases the slope $v'(\gamma) = \lambda_t = \frac{\kappa_t}{f'(X_{t+1})}$ decreases. Since $f'(X_{t+1}) > 0$, all else equal, this implies that $\kappa_t > 0$ decreases.

(Case 2) At time t the optimum is $X^i_{t+1} < 0$ with a binding constraint and $g'(X_{t+1}) = -f'(X_{t+1}) < 0$.

A similar argument to Case 1 yields that $\lambda_t = \frac{\kappa_t}{g'(X_{t+1})}$. By Lemma 2, $\kappa_t < 0$ for $X_{t+1} < 0$. So, $g'(X_{t+1}) < 0$ implies that $\lambda_t > 0$.

By standard results $\upsilon'(\delta) = \lambda_t$, hence $\upsilon'(\delta) = \frac{\kappa_t}{g'(X_{t+1})} > 0$. By the concavity of $\nu(\delta)$, $\upsilon''(\delta) < 0$, which implies that as δ increases, the slope $\upsilon'(\delta) = \lambda_t = \frac{\kappa_t}{g'(X_{t+1})}$ decreases. Since $g'(X_{t+1}) < 0$, this implies that $\kappa_t < 0$ becomes less negative, i.e. decreases in absolute value. But, $\delta = -(1+\gamma)$ increasing implies that γ decreases. Hence, as γ increases, $\kappa_t < 0$ becomes more negative, i.e. increases in absolute value.

This completes the proof. ∎

Using expression (4.8), we get an immediate corollary relating trading constraints to the size of asset price bubbles.

Corollary 11.1. *(Trading Constraints and Bubble Magnitudes)*

Suppose that the trading constraint is made more restrictive. Then, the absolute value of a bubble's magnitude increases, i.e. $\left|\beta_t^i\right|$ *increases.*

Proof. Given $\beta_t^i = \frac{-\kappa_t^i}{E_t^{\mathbb{P}^i}[U_i'(Y_{T+1}^i)]\varphi_t'(\Delta X_{t+1}^i)}$, we have that γ increasing (decreasing) implies: (i) $\kappa_t^i > 0$ decreases (increases) when $X_{t+1}^i > 0$, which implies $\beta_t^i < 0$ decreases (increases) in absolute value, and (ii) $\kappa_t^i < 0$ increases in absolute value when $X_{t+1}^i < 0$, which means $\beta_t^i > 0$ increases as there is a negative sign in the above expression.

We note the asymmetry in the increase of γ. In case (i) when $X_{t+1}^i > 0$, γ *decreasing* means the constraint is becoming more binding/restrictive (less borrowing possible). In case (ii) when $X_{t+1}^i < 0$, γ *increasing* means the constraint is becoming more binding/restrictive (less short selling possible). Thus, as γ becomes more binding, in both cases the bubble increases in absolute value. ■

This result is subtle and might appear to be counter-intuitive. One might argue that relaxing the trading constraints should lead to more speculative behavior and larger bubbles. However, this argument ignores the reason bubbles exist in our economy. As we have shown, there are no bubbles in an economy when trading constraints are non-binding i.e. $\kappa_t^i = 0$. They only appear when the constraints are binding. Hence, given their beliefs, agents would like to either buy or short sell more stocks to reduce their personalized under- and over-valuation. But, they cannot. The more binding the trading constraints, the larger the personalized under- and over-valuation. This is consistent with the fact that bubbles appear when either the short selling or borrowing constraints are binding.

4.3 Systemic Risk

This section studies the relation between asset price bubbles, market illiquidity, trading constraints, and systemic risk. We will show that systemic risk is the risk of market failure resulting from funding risk, which is the conjunction of increased market illiquidity (liquidity risk) due to widespread selling pressure (like a fire-sale), leading to binding trading constraints that generate massive defaults across agents. As implied, funding risk is closely related to the notions of default, an accelerator effect, and the percentage of agents in an economy that see bubbles. To see all of these relationships, we first need to define an agent's wealth.

Definition 7. (Wealth)

Let W_t^i denote the wealth (or net-worth of a self-financing *trading strategy) of an individual agent of type i, defined as*

$$W_t^i = X_{t+1}^i S_t + Y_{t+1}^i$$

for all times $t \in \{0, 1, ..., T-1\}$.

Analogous to the definition of the trading constraint, wealth is defined using the marked-to-market value of the risky assets. This makes sense because the transaction price requires keeping track of the share purchases at any time t and the liquidity cost paid $\varphi_t'(x)$. In addition, defining wealth using market-to-market price implies that an agent's wealth is always positive, as the following lemma shows.

Lemma 4. *(Wealth is Always Nonnegative)*

If $X_{t+1}^i > 0$ and $\beta_t^i < 0$, then $Y_{t+1}^i = -\gamma S_t X_{t+1}^i$ and *$W_t^i = (1-\gamma) X_{t+1}^i S_t > 0$.*

If $X_{t+1}^i < 0$ and $\beta_t^i > 0$, then $Y_{t+1}^i = -(1+\gamma) S_t X_{t+1}^i$ and $W_t^i = -\gamma X_{t+1}^i S_t > 0$.

Intuitively, systemic risk is the risk of market failure in an economy due to an exogenous and unanticipated shock that causes massive defaults. In our setting, we define a market failure as the nonexistence of an equilibrium. The shock could be due to catastrophic events or changes in monetary policy, fiscal policy, regulatory policy (trading constraints), or changes to agents' beliefs and preferences. Thus, in our setting market failure results from an exogenous and unanticipated shock that causes the non-existence of an economic equilibrium due to the inability of some agents to satisfy their trading constraints, thereby defaulting with a corresponding loss in wealth. As characterized, systemic risk is the result of a funding illiquidity (funding risk), which is the conjunction of market illiquidity and binding trading constraints that lead to defaults across agents.

It is important to note that the exogenous shock must be unanticipated and not included in the original specification of the economy. Indeed, if the shock is anticipated and included into the economy's structure, then because our equilibrium is a Radner equilibrium, it holds for all times t a.e. $\mathbb{P}$. Hence, the economy is always in equilibrium and there is never any market failure. Anticipated but random shocks only increase the volatility/risk of consumption and equilibrium prices. In equilibrium, agents' trading constraints are never violated, and there is no default (wealth stays positive). This is a direct implication of the existence of an agent's optimal portfolio subject to the trading constraints. We show below that for systemic risk effects, the price shocks do not have to be large. The aggregate effect of small price shocks can also lead to the realization of funding illiquidity and market failure, leading to a significant wealth loss in the economy.

We use a reduced form approach to represent an exogenous shock to the economy. Rather than explore an exogenous shock to the fundamentals of the economy (beliefs, preferences, endowments, trading constraints) and explore its implications on economic equilibrium prices and consumption, we analyze the implications of an exogenous shock on the price process itself. This reduced form approach simplifies the analytics because it enables us to determine, in a competitive economy where all agents are price takers, whether an optimal portfolio still exists after the shock, given the trading constraints. Alternatively stated, it enables us to determine if such an exogenous shock to prices causes any agents to default, i.e. violate the trading constraints, which corresponds to too much debt or an inability to meet short sale margin calls. Default for an agent type, which represents a positive mass in the economy $\mathbb{I}^-_{\beta_t}$ or $\mathbb{I}^+_{\beta_t}$, implies market failure because it results in the non-existence of an economic equilibrium.

Given this discussion, we analyze an unanticipated random shock to the price process equal to ΔS at time t, which could be a large or small, positive or negative change from S_t to $S_t + \Delta S$. The next theorem characterizes the conditions under which such a price shock results in the non-existence of an economic equilibrium and market failure, which is an *instance* of systemic risk materializing into systemic crisis.

Theorem 12. *(Margin Calls, Market Failure, Wealth Loss)*

(Borrowing Constrained)

If $X^i_{t+1} > 0$ *and* $\beta^i_t < 0$.

Let the price shock be $\Delta S < 0$ *at time t.*

Then, the minimum sale $\Delta X^i < 0$ *necessary to satisfy the borrowing constraint satisfies*

$$\left(\varphi_t(\Delta X^i) - \gamma \Delta X^i\right) = \gamma \frac{\Delta S}{S_t + \Delta S} X^i_{t+1}.$$

A feasible $\Delta X^i < 0$ *exists (staying in equilibrium) if and only if* $\varphi_t(\Delta X^i) < \gamma \Delta X^i$.

Market failure occurs otherwise with a loss of wealth.

(Short Sale Constrained)

If $X^i_{t+1} < 0$ and $\beta^i_t > 0$.

Let the price shock be $\Delta S > 0$ at time t.

Then, the minimum purchase $\Delta X^i > 0$ necessary to satisfy the margin constraint satisfies

$$\left(\varphi_t(\Delta X^i) - (1+\gamma)\,\Delta X^i\right) = (1+\gamma)\,\frac{\Delta S}{S_t + \Delta S} X^i_{t+1}.$$

A feasible $\Delta X^i > 0$ exists (staying in equilibrium) if and only if $\varphi_t(\Delta X^i) < (1+\gamma)\,\Delta X^i$.

Market failure occurs otherwise with a loss of wealth.

Proof. Two cases to consider.

(Case 1) $X^i_{t+1} > 0$ and $\beta^i_t < 0$ with $\Delta S < 0$ at time t.

Before the shock borrowings are $Y^i_{t+1} = -\gamma S_t X^i_{t+1} < 0$.

After the shock, the maximum borrowings are $-\gamma\,(S_t + \Delta S)\,X^i_{t+1} > -\gamma S_t X^i_{t+1}$. Hence, the constraint is violated and the change in wealth is negative.

Indeed,

$$\left(X^i_{t+1} S_t - \gamma S_t X^i_{t+1}\right) - \left(X^i_{t+1} S_t - \gamma\,(S_t + \Delta S)\,X^i_{t+1}\right) = \gamma X^i_{t+1} \Delta S < 0.$$

This implies, borrowers must sell shares $\Delta X^i < 0$ to obtain cash to reduce borrowings so that the constraint is not violated. Next, we determine the minimum shares to sell to stay on the constraint.

After selling shares, the constraint is $-\gamma\,(S_t + \Delta S)\left(X^i_{t+1} + \Delta X^i\right) < 0$.

The cash needed is $-\gamma\,(S_t + \Delta S)\left(X^i_{t+1} + \Delta X^i\right) + \gamma S_t X^i_{t+1} > 0$.

From the liquidity cost of trading, the cash obtained from selling shares is

$-\varphi_t(\Delta X^i)\,(S_t + \Delta S) > 0.$

From above cash needed/obtained conditions, the solution is ΔX^i such that

$\gamma\,(S_t + \Delta S)\left(X^i_{t+1} + \Delta X^i\right) - \gamma S_t X^i_{t+1} = \varphi_t(\Delta X^i)\,(S_t + \Delta S).$

$\gamma\,(S_t + \Delta S)\,X^i_{t+1} + \gamma\,(S_t + \Delta S)\,\Delta X^i - \gamma S_t X^i_{t+1} = \varphi_t(\Delta X^i)\,(S_t + \Delta S).$

$\gamma \Delta S\,X^i_{t+1} + \gamma\,(S_t + \Delta S)\,\Delta X^i = \varphi_t(\Delta X^i)\,(S_t + \Delta S).$

$\gamma \frac{\Delta S}{S_t + \Delta S} X^i_{t+1} + \gamma \Delta X^i = \varphi_t(\Delta X^i).$

$\gamma \frac{\Delta S}{S_t + \Delta S} X^i_{t+1} = \left(\varphi_t(\Delta X^i) - \gamma \Delta X^i\right).$

The left side is negative as $\Delta S < 0$ and $X^i_{t+1} > 0$.

A solution exists with $\Delta X^i < 0$ if and only if $\varphi_t(\Delta X^i) - \gamma \Delta X^i < 0$.

(Case 2) $X^i_{t+1} < 0$ and $\beta^i_t > 0$ with $\Delta S > 0$ at time t.

Before the shock the margin is $Y^i_{t+1} = -(1 + \gamma)\,S_t X^i_{t+1} > 0$.

After the shock, the margin is $-(1 + \gamma)\,(S_t + \Delta S)\,X^i_{t+1} > -(1 + \gamma)\,S_t X^i_{t+1}$. Hence, the constraint is violated and the change in wealth is negative as in case 1 above.

This implies, short sellers must buy shares $\Delta X^i > 0$ to reduce the short position so that the margin constraint is not violated. Next, we determine the minimum shares to buy to stay on the constraint.

After buying shares, the constraint is $-(1 + \gamma)\,(S_t + \Delta S)\left(X^i_{t+1} + \Delta X^i\right) > 0$.

The cash reduction in the margin from purchase is

$$(1+\gamma)(S_t+\Delta S)\left(X^i_{t+1}+\Delta X^i\right)-(1+\gamma)S_tX^i_{t+1}>0.$$

The cash needed to buy shares is $\varphi_t(\Delta X^i)(S_t+\Delta S)>0$.

From above cash reduction/needed conditions, the solution is ΔX^i such that

$$(1+\gamma)(S_t+\Delta S)\left(X^i_{t+1}+\Delta X^i\right)-(1+\gamma)S_tX^i_{t+1}=\varphi_t(\Delta X^i)(S_t+\Delta S).$$

$$(1+\gamma)(S_t+\Delta S)X^i_{t+1}+(1+\gamma)(S_t+\Delta S)\Delta X^i-(1+\gamma)S_tX^i_{t+1}$$

$$=\varphi_t(\Delta X^i)(S_t+\Delta S).$$

$$(1+\gamma)\Delta S\,X^i_{t+1}+(1+\gamma)(S_t+\Delta S)\Delta X^i=\varphi_t(\Delta X^i)(S_t+\Delta S).$$

$$(1+\gamma)\tfrac{\Delta S}{S_t+\Delta S}X^i_{t+1}+(1+\gamma)\Delta X^i=\varphi_t(\Delta X^i).$$

$$(1+\gamma)\tfrac{\Delta S}{S_t+\Delta S}X^i_{t+1}=\left(\varphi_t(\Delta X^i)-(1+\gamma)\Delta X^i\right).$$

The left side is negative as $\Delta S>0$ and $X^i_{t+1}<0$.

A solution exists with $\Delta X^i>0$ if and only if $\varphi_t(\Delta X^i)-(1+\gamma)\Delta X^i<0$. ■

Before discussing the theorem, we note that the price shock ΔS considered for both the borrowing and short sale constrained agents results in a loss of wealth, independent of whether or not market failure occurs. As shown below, the magnitude of the wealth loss is larger when market failure occurs.

This theorem characterizes the conditions under which a price shock of size ΔS results in the inability of an agent to satisfy her trading constraint, resulting in default and market failure. There are two such possibilities. The first situation occurs when the i^{th} agent is borrowing to buy the stock, is borrowing constrained, and therefore believes the stock is undervalued ($X^i_{t+1} > 0$ and $\beta^i_t < 0$). Default and market failure occur if the liquidity cost of selling the stocks $\Delta X^i < 0$ to satisfy the constraint does not generate enough cash. This occurs when $0 > \varphi_t(\Delta X^i) > \gamma \Delta X^i$. This is the standard mechanism often discussed intuitively in the literature (see Brunnermeier and Pedersen (2008) and Brunnermeier et al. (2013)). The second situation is largely unexplored in the macroeconomics literature. It occurs when the i^{th} agent is short selling the stock, is margin constrained, and therefore believes the stock is overvalued ($X^i_{t+1} < 0$ and $\beta^i_t > 0$). Default and market failure occur if the liquidity cost of buying back the shorted stocks $\Delta X^i > 0$ to satisfy the constraint is too costly. This occurs when $\varphi_t(\Delta X^i) > (1 + \gamma)\,\Delta X^i > 0$. In both cases, when funding illiquidity (market illiquidity and binding trading constraints) causes an agent's default, market failure occurs and there is a significant loss of wealth in the economy.

Market failure occurs in both of these situations because of funding illiquidity. Indeed, the i^{th} agent cannot sell/buy enough stock to satisfy the trading constraints due to the liquidity costs of executing trades. In practice, both the stock and borrowing positions would be in the same brokerage account. If after the price shock an insufficient quantity of stock/mma exists in the portfolio to satisfy the trading constraint, then the entire brokerage account is necessarily liquidated by the agent's broker. Under this assumption, it is easy to show that under some mild conditions the i^{th} agent's wealth becomes negative, which implies bankruptcy.

Corollary 12.1. *(Market Failure Implies Bankruptcy)*

Let market failure occur.

(Borrowing Constrained)

Let $X^i_{t+1} > 0$ *,* $\beta^i_t < 0$*, and* $\Delta S < 0$ *at time t.*

If $\varphi_t(-X^i_{t+1})\,[S_t + \Delta S] > \gamma(-X^i_{t+1})S_t$*, then the* i^{th} *agent is bankrupt.*

(Short Sale Constrained)

Let $X^i_{t+1} < 0$ *,* $\beta^i_t > 0$*, and* $\Delta S > 0$ *at time t.*

If $Z < -X^i_{t+1}$ *where* $Z > 0$ *is the solution to* $\varphi_t(Z)\,[S_t + \Delta S] = -(1+\gamma)\,S_t X^i_{t+1}$*, then the* i^{th} *agent is bankrupt.*

Proof. (Step 1)

Consider the first case where the agent is borrowing to buy the stock. Here, before the shock, the agent's wealth is

$$W^i_t = S_t X^i_{t+1} + Y^i_{t+1} = (1-\gamma)\,S_t X^i_{t+1} > 0$$

because $Y^i_{t+1} = -\gamma S_t X^i_{t+1} < 0$. After the price shock of $\Delta S < 0$, the agent cannot satisfy the trading constraint. Given that the entire stock position is liquidated ($-X^i_{t+1} < 0$), the wealth after liquidation is

$$W^i_t(after) = -\varphi_t(-X^i_{t+1})\,[S_t + \Delta S] + Y^i_{t+1}$$

$$= -\varphi_t(-X^i_{t+1})\,[S_t + \Delta S] - \gamma S_t X^i_{t+1}.$$

Note here that the position in the mma is fixed. This new wealth is negative if $\varphi_t(-X^i_{t+1})\,[S_t + \Delta S] > -\gamma S_t X^i_{t+1}$.

(Step 2)

Consider the second case where the agent has a margin account to short the stock. Here, before the shock, the agent's wealth is

$$W^i_t = S_t X^i_{t+1} + Y^i_{t+1} = -\gamma S_t X^i_{t+1} > 0$$

because $Y^i_{t+1} = -(1+\gamma)\, S_t X^i_{t+1} > 0$. After the price shock of $\Delta S > 0$, the agent cannot satisfy the trading constraint. Given the entire mma position is liquidated to buy stock to cover the short position, the wealth after buying back stock is

$$W^i_t(after) = [S_t + \Delta S]\,(X^i_{t+1} + Z)$$

where $Z > 0$ is the solution to $\varphi_t(Z)\,[S_t + \Delta S] = -(1+\gamma)\, S_t X^i_{t+1}$. Note in this case the position in the shorted shocks is fixed. This new wealth is negative if $Z < -X^i_{t+1}$. This completes the proof. ■

The two sufficient conditions for bankruptcy are quite mild and similar to the market failure conditions $\varphi_t(\Delta X^i) > \gamma \Delta X^i$ for $\Delta X^i < 0$ when the agent is borrowing constrained and $\varphi_t(\Delta X^i) > (1+\gamma)\,\Delta X^i$ for $\Delta X^i > 0$ when the agent is short sale constrained where ΔX^i correspond to the solutions to the respective equations in theorem 12.

Besides characterizing the conditions under which default and market failure occur, the previous theorem also provides some insights into the impact of the shock, if equilibrium still exists and a market failure does not occur. This implication is highlighted in the following corollary.

Corollary 12.2. *(An Accelerator Effect and Augmented Loss of Wealth)*

Assuming the minimum purchase necessary to satisfy the trading constraints is feasible in theorem (12) i.e. staying in equilibrium, then the minimum sale or purchase ΔX^i necessary to satisfy the trading i.e. borrowing or short sale constraint exceeds that in an economy with no liquidity costs. Consequently, the loss of wealth is larger in an economy with liquidity costs.

(Borrowing Constrained)

Let $X^i_{t+1} > 0$ *,* $\beta^i_t < 0$*, and* $\Delta S < 0$ *at time t.*

The wealth loss is for $\Delta X^i > 0$*,*

$$(1-\gamma)\, S_t X^i_{t+1} - (1-\gamma)\,[S_t + \Delta S_t]\left[X^i_{t+1} + \Delta X^i\right] > 0.$$

(Short Sale Constrained)

Let $X^i_{t+1} < 0$ *,* $\beta^i_t > 0$*, and*$\Delta S > 0$ *at time t.*

The wealth loss is for $\Delta X^i < 0$*,*

$$-\gamma S_t X^i_{t+1} + \gamma\,[S_t + \Delta S_t]\left[X^i_{t+1} + \Delta X^i\right] > 0.$$

Proof. (Step 1)

If $X^i_{t+1} > 0$ and $\beta^i_t < 0$. Let the price shock be $\Delta S < 0$ at time t. Then, by the previous theorem (12) the minimum sale $\Delta X^i < 0$, necessary to satisfy the borrowing constraint satisfies $\Delta X^i = \frac{1}{\left(\frac{\varphi_t(\Delta X^i)}{\Delta X^i} - \gamma\right)} \gamma \frac{\Delta S}{S_t + \Delta S} X^i_{t+1} > \frac{1}{(1-\gamma)} \gamma \frac{\Delta S}{S_t + \Delta S} X^i_{t+1}$. The right side of this inequality is the trade size in a market with no illiquidities where $\varphi_t(\Delta X^i) = \Delta X^i$.

(Step 2)

If $X^i_{t+1} < 0$ and $\beta^i_t > 0$. Let the price shock be $\Delta S > 0$. Then, by the previous theorem the minimum purchase $\Delta X^i > 0$ necessary to satisfy the margin constraint satisfies $\Delta X^i = \frac{1}{\left(\frac{\varphi_t(\Delta X^i)}{\Delta X^i} - (1+\gamma)\right)} (1+\gamma) \frac{\Delta S}{S_t + \Delta S} X^i_{t+1} > \frac{1}{\gamma} (1+\gamma) \frac{\Delta S}{S_t + \Delta S} X^i_{t+1}$. The right side of this inequality is the trade size in a market with no illiquidities where $\varphi_t(\Delta X^i) = \Delta X^i$.

In both steps above, the loss in wealth is larger than in the economy with liquidity costs.

(Step 3)

The change is wealth is computed by recognizing that the constraint is satisfied both before and after the trades. ■

This corollary (12.2) can be explained as follows. After a price shock, a sale/purchase is needed to satisfy the trading constraints. If such a sale/purchase exists that is feasible and an equilibrium still exists, then the size of the sale/purchase exceeds that which would occur in a market with no illiquidities. This is intuitive because if there are liquidity costs to trading, then these liquidity costs must be subtracted from the proceeds, and the size of the trade must be larger to account for these costs. As such, this implies that the aggregate trades (aggregated by the percentage of agents that trade) will affect equilibrium prices more in an economy with liquidity costs than in an economy without liquidity costs. This increased trading will result in a secondary price change (of the same sign) that augments the original price shock of ΔS. This in turn will lead to another necessary sale/purchase to satisfy the binding trading constraints, leading to another subsequent price change to the second price shock, and so forth until the sequence either converges to a stable and feasible price change or default and market failure occurs. Consequently, the loss of wealth is larger in an economy with liquidity costs. This is the accelerator effect of funding illiquidity due to an exogenous price shock in the economy. This accelerator effect is analogous to that in Brunnermeier and Pedersen (2008) and Brunnermeier and Sannikov (2014).

We can quantify the probability of massive agent defaults and market failure, or systemic risk, in our economy due to an exogenous price shock of size ΔS.

Definition 8. (Systemic Risk)

The probability of market failure at time t for a random shock of size ΔS *is*

$$\mathbb{P}_{fail} = \mathbb{P}\{\omega \in \Omega : \exists\, i \text{ where } \left(\beta_t^i < 0, Y_{t+1}^i = -\gamma S_t X_{t+1}^i < 0, \Delta X^i < 0, \varphi_t(\Delta X^i) > \gamma \Delta X^i\right)$$
$$\text{or } \left(\beta_t^i > 0, Y_{t+1}^i = -(1+\gamma) S_t X_{t+1}^i > 0, \Delta X^i > 0, \varphi_t(\Delta X^i) > (1+\gamma)\Delta X^i\right)\} \tag{4.11}$$

where ΔX^i *is the solution to*

$\varphi_t(\Delta X^i) - \gamma \Delta X^i = \gamma \frac{\Delta S}{S_t + \Delta S} X_{t+1}^i$ *for* $\Delta S < 0$ *and*

$\varphi_t(\Delta X^i) - (1+\gamma)\Delta X^i = (1+\gamma)\frac{\Delta S}{S_t + \Delta S} X_{t+1}^i$ *for* $\Delta S > 0$.

Note that this definition implicitly takes into account the probability of wealth loss as disequilibrium necessarily entails some agents' inability to satisfy the trading constraint, resulting in default and a loss of wealth (see theorem 12). Thus, this definition is seen to be equivalent to the probability of a significant loss of wealth in the economy due to massive agent defaults and consequent market failure. By direct inspection of expression (4.11), the following theorem follows.

Theorem 13. *(Bubbles, Liquidity, Constraints, and Systemic Risk)*

(i) As $\mathbb{I}_{\beta_t}^{+}$ *or* $\mathbb{I}_{\beta_t}^{-}$ *increase, all else constant,* $\mathbb{P}_{fail}$ *increases.*

(ii) As the market becomes more illiquid, all else constant, $\mathbb{P}_{fail}$ *increases.*

(iii) As the constraints become more restrictive, all else constant, $\mathbb{P}_{fail}$ *decreases.*

Proof. By expression (4.11), we have three cases.

(i) As $\mathbb{I}_{\beta_t}$ increases, $\left|\beta_t^i\right|$ is increasing, which makes $\beta_t^i < 0$ more likely for $\Delta X^i < 0$ and it makes $\beta_t^i > 0$ more likely for $\Delta X^i > 0$. This gives the result.

(ii) As markets become more illiquid, $\varphi_t(\Delta X^i)$ increases for all $\Delta X^i \neq 0$. See Lemma (3). This makes $\varphi_t(\Delta X^i) > \gamma \Delta X^i$ more likely for $\Delta X^i < 0$ and it makes $\varphi_t(\Delta X^i) > (1+\gamma)\Delta X^i$ more likely for $\Delta X^i > 0$, due to monotonicity of probability measure. This gives the desired result.

(iii) As the constraints become more binding, this means that γ decreases for $X^i_{t+1} > 0$ and that $1+\gamma$ increases for $X^i_{t+1} < 0$. This makes $0 > \varphi_t(\Delta X^i) > \gamma \Delta X^i$ less likely for $\Delta X^i < 0$ and it makes $\varphi_t(\Delta X^i) > (1+\gamma)\Delta X^i > 0$ less likely for $\Delta X^i > 0$, due to monotonicity of probability measure. This gives the desired result. ■

This theorem shows the effects that changing fundamentals have on systemic risk. It states that if the percentage of agents that believe bubbles exist increases or if the market becomes more illiquid or if trading constraints are relaxed, then systemic risk increases. We note that using corollary (11.1), a secondary effect of making the trading constraints more restrictive is that for any agent, the absolute value of a bubble's magnitude increases, i.e. $\left|\beta^i_t\right|$ increases. This means that the percentage of agents seeing bubbles, $\mathbb{I}^+_{\beta_t}$ and $\mathbb{I}^-_{\beta_t}$, increases as well. Then, by implication (i) above, $\mathbb{P}_{fail}$ increases. This secondary effect, which involves staying in equilibrium, may or may not dominate the primary implication in (iii) above, generated under the market failure condition described in theorem (12) where equilibrium does not exist, assuming all else constant, including the size of the asset price bubble.

All of these implications are intuitive and consistent with recent market experience during the financial crisis of 2007-2009. U.S. households were highly levered during the pre-crisis periods and entered the crisis with huge amounts of debt. Indeed, U.S. household gross debt as a percent of personal income was 96% in 2000 and 128% in 2008. Then the shock emanating from the financial sector eventually lead to the collapse of housing bubbles, stock market crash and consequently reduced borrowing limits in the economy drastically. It further lead to massive and rapid de-leveraging in the economy accompanied by a significant loss of wealth. See Eggertsson and Krugman (2012) for further discussion on this debt and deleveraging mechanism. In our analysis of systemic risk, we have shown multiple effects of similar deleveraging mechanisms in a much more general set up through market failure and accelerator effects due to borrowing and short sale constraints.

4.4 Conclusion

This chapter provides an equilibrium model with heterogeneous agents and trading constraints to study asset price bubbles, market illiquidities, and systemic risk. Systemic risk is defined as the risk of market failure due to an exogenous shock to the economy. This results in funding illiquidity, which is the conjunction of market illiquidity (i.e. liquidity risk) and binding trading constraints. To do this, we introduce a different framework for analyzing asset price bubbles and how they affect the macroeconomy. In our framework, asset price bubbles arise endogenously in a rational equilibrium due to the heterogeneous beliefs, heterogeneous preferences, and binding trading constraints. We show that: (i) positive price bubbles are larger in more illiquid markets, (ii) the percentage of agents in the economy who believe a positive price bubble exists decreases as a market becomes more liquid, (iii) a bubble's magnitude increases when trading constraints are more restrictive, and (iv) systemic risk increases as the percentage of agents seeing bubbles increases or as the market becomes more illiquid. The realization of systemic risk results in a large fraction of agents violating their trading constraints thereby defaulting, the non-existence of an equilibrium, and a large loss of wealth in the economy.

Our results also have policy implications because market liquidity, trading constraints and asset price bubbles affect systemic risk. We show that improved market liquidity decreases systemic risk, and it is well known that monetary and other regulatory policies can affect market liquidity. We also show that as the percentage of agents viewing price bubbles increase, systemic risk increases. Regulators can directly influence agents' beliefs with appropriate policy actions. To influence the impact of bubbles on systemic risk, policy makers should focus on determining the *percentage of agents that see bubble*s, and *not* whether a uniform price bubble exists, which is a difficult if not an impossible task. Alternatively stated, the percentage of agents that see price bubbles matters because this is the most relevant channel through which policy with respect to bubbles works, and not via the overall level of market prices. This is important because the realization of systemic risk results in a significant fraction of agents defaulting with a corresponding loss of wealth in the economy. Lastly, we show that as trading constraints become more binding, systemic risk declines. Regulators can certainly modify trading constraints. More work still needs to be done to fully understanding how asset price bubbles, trading constraints, and market liquidities are influenced by regulatory policies. In particular, it would be useful to introduce monetary policy directly into the model and explore its equilibrium implications. This is an open and fruitful future research area.

APPENDIX A

APPENDIX

A.1 Numerical Computation

As seen from the dynamic programming formulation in the model section, the optimization problem is highly non-linear. The analytical solutions for Kolmogorov-Forward equation- a partial differential equation- is generally elusive, and only known for the linear oscillator cases. Thus, no analytical solution is known for the non-linear dynamics of our savings and debt policy functions, $\eta^l(a^l, z^l)$ and $\eta^b(a^b, z^b)$ respectively. Thus, we need to rely on numerical approximation to the stationary densities $g^b(.)$ and $g^l(.)$. A popular and efficient way to do so is finite difference method (finite element method is also widely used in physics/mathematics literature). Since the dynamic problem is already formulated in terms of operators, this considerably simplifies the algorithm design. Once the operator associated with HJB is solved, the operator associated with KE is just the adjoint of the operator in HJB equation. And the adjoint in finite dimension is just the matrix transpose. Since we need to appropriately discretize the problem for computations, we have to reduce the original infinite dimensional problem to the finite dimensional case. That means one needs to be careful on defining and grouping terms while solving the HJB equations first.

Building on an earlier work by Lasry and Lions (2007), and Achdou, Camilli and Capuzzo-Dolcetta (2012), Achdou et al. (2015) have made progress in the study of heterogeneous agents by taking advantage of the rich theories and methodology available in continuous time. They highlight efficient algorithms for numerical computations. This section describes the numerical techniques of finite difference method to compute stationary equilibrium, closely following Achdou et al. (2015). This method is very simple to implement while being robust, flexible and efficient. In addition to simplifying the model and making it amenable to various analytical techniques, the continuous time modeling approach also offers practical computational advantage compared to the discrete-time case. For overview of the computational advantages and for detailed numerical outline/algorithm we refer the readers to Achdou et. al. (2015). Here we only provide a brief sketch of the numerical techniques involved.

While Achdou et. al. (2015) do not incorporate jump-diffusion, similar discretization method can be used for processes with jumps, as we have in our model. For example, see Kaplan, Moll and Violante (2016) for processes involving jumps with two assets and non-convex adjustment costs. Our HJB and KF equation involving jumps are also discretized similarly. As discussed in the model section, for solving a fully general model with partial integro-differential equation (PIDE), finite difference scheme developed for pricing barrier options can also be considered. See Cont & Tankov (2004) and Cont & Voltchkova (2005) for more details on solving such PIDEs. Here, we use a simplified setting.

Conditions required for the convergence of the finite difference scheme is guided by the theory developed in Barles and Souganidis (1991). They have shown that under some conditions the solution to a finite difference scheme converges to the (unique viscosity) solution of the HJB equation. Roughly stating, one of their key theorems states that if the finite difference scheme satisfies some conditions of monotonicity, consistency and stability, then the solution of the optimization problem converges locally uniformly to the unique viscosity solution. The main method is called *upwind scheme*, and more specifically the *implicit* upwind scheme. This method can easily handle the boundary conditions and the borrowing constraint of our model. Here we briefly discuss the stationary equilibrium computations. We do not compute transition dynamics because we explore it analytically. For detailed steps/theory behind computations we again refer the readers to Achdou et al. (2015) as those computational steps easily apply to our model, and here we will try to be consistent with their notations.

We use a bisection algorithm to get the stationary interest rate r. We start the iteration with an initial guess r_0 and for successive iterations $k = 0, 1, 2, ...$ we follow the steps below:

1. Given r_k, solve the stationary version of HJB equations above for both households using a finite difference method. Calculate the saving and debt policy functions η_k^l and η_k^b.
2. Given η_k^b and η_k^l from (1) above, solve for the stationary KF equations for both households.
3. Given g_k^b and g_k^l, compute the net supply of loans equation (2.35) i.e. $\Gamma^l(r) + \Gamma^b(r) = 0$. If net supply greater than zero, decrease the interest rate $r_{k+1} < r_k$ and vice versa.

When r_{k+1} is close to r_k within some tolerance level, we call the resulting interest rate along with the HJB and KF equations a stationary equilibrium.

We now briefly describe the discretization of the stationary HJB and KF equations of the borrowers as in equations (2.30) and (2.31), reproduced below. Similar steps apply to the lenders when default intensity is set to $\lambda = 0$. For simplicity we consider fixed jump size as discussed in the model section. So we remove the expectation operator across the jump size J. The computations after the credit event is similar. Given that we have a Markov economy, we just reduce the income of borrowers by some fraction after they default. Further, we also reduce the borrowing limit as well. Then the economy evolves just as before (i.e. memoryless property), but now with lower income and tighter borrowing constraint. This is essentially capturing the bankruptcy cost feature of default. For lenders, they would receive some fixed amount after borrowers default i.e. the insurance payment. Note that lump-sum payments only scale the budget sets and do not affect the optimal decision. Thus, the economy evolves just as before after these adjustments to the budget sets.

$$
\begin{aligned}
\rho^b V^b(a^b, z^b) &= \max_c u(c^b) + \mathcal{L}V^b(a, z) \\
0 &= \mathcal{L}^* g^b(z, a) \\
\mathcal{L}V^b(a, z) &= \partial_a V^b(a^b, z^b)(\eta^b(a^b, z^b)) + \partial_z V^b(a^b, z^b)\mu(z^l) \\
&\quad + \frac{1}{2}\partial_{zz} V^l(a^l, z^l)\sigma^2(z^l) + \lambda[V^b(a^b, z'^b) - V(a^b, z^b)] \\
\mathcal{L}^* g^b(a, z) &= -\partial_a[\eta^b(a^b, z^b) g^b(a^b, z^b)] - \partial_z[\mu(z^b) g^b(a^b, z^b)] \\
&\quad + \frac{1}{2}\partial_{zz}[g^b(a^b, z^b)\sigma^2(z^b)] + \lambda[g(a^b, z''^b) - g(a^b, z^b)] \\
\eta^b(a^b, Z^b) &= (z^b + ra^b - c^b)\mathbb{1}_{\{\tau > t\}} \\
&\quad + (z^b - \delta z^b + ra^b - c^b)\mathbb{1}_{\{\tau \le t\}}
\end{aligned}
$$

Denote $V^b(a_i, z_j) = V_{i,j}$ where debt is indexed by $i = 1, 2, ...I$ and income is indexed by $j = 1, 2, ...J$, so $I \times J$ is the matrix of grid points. This gives a system of $I \times J$ linear equations. Thus, the discretized version of HJB can be written as:

$$\frac{V_{i,j}^{n+1} - V_{i,j}^{n}}{\Delta} + \rho^b V_{i,j}^{n+1} = u(c_{i,j}^n) + \partial_a V_{i,j}^{n+1}(z_j + ra_i - c_{i,j}^n)$$
$$+ \mu_j \partial_z V_{i,j}^{n+1} + \frac{\sigma_j^2}{2}\partial_{zz} V_{i,j}^{n+1} + \lambda[V_{i,j'}^{n+1} - V_{i,j}^{n+1}]$$

where $j' \neq j$. The parameter Δ is the step size which can be arbitrary large given the implicit scheme. If we allowed for multiple jump magnitudes we would have to sum over all those possible values, possibly weighted by different λ. Since we abstract away from such complications by considering one fixed jump size, we largely avoid such computational complexities.

The upwind method involves using either backward or forward difference approximation of the derivatives of the value function depending on the sign of the drift: if drift is positive use forward approximation, if drift is negative use backward approximation. This is mainly relevant with respect to the state variable debt a^b. Since income process is given exogenously, and has a well defined drift, diffusion and jump terms, either approximation should work for the derivatives of value function with respect to income. After defining the forward and backward differences, and following all the relevant steps as outlined in Achdou et al. (2015), we get the discretized version of HJB equation in matrix notation as follows:

$$\frac{V^{n+1} - V^{n}}{\Delta} + \rho^b V^{n+1} + u^n + A^n V^{n+1}$$

where A^n is a $I \times J$ matrix that encodes the evolution of the state variables a^b and z^b. This matrix is indeed the finite dimension analogue of the backward operator $\mathcal{L}V^b(a,z)$. The simplicity offered by this operator approach is that the adjoint operator $\mathcal{L}^* V^b(a,z)$ in the finite dimensional case is analogous to the matrix transpose. Thus, after properly discretizing HJB equation and carefully collecting the appropriate terms together we get the KF equation as well. We first discretize the KF equation appropriately and then require that the discretized density g^b integrate to one. There is no iterative step involved in solving for density. Finally we take the transpose A^T of the matrix A, which immediately gives the stationary KF equation as:

$$A^T g^b = 0$$

Thus, after approximation of the matrix A in HJB equation, the remaining computation is greatly simplified once we recognize, based on simple application of operators, that approximation to KF equation can be obtained by taking the transpose of A (more specifically, the A^n matrix from last HJB iteration step).

A.2 Proofs

Proof. **Proposition (1):** The proof and arguments here are very similar to those in Aiyagari (1994), Miao(2006), and Achdou et al. (2015). This is done by use of fixed point theorem, specifically the intermediate value theorem. We also discuss few other ways to prove the existence in our model.

Define the net savings/debt in the economy: $SD(r) = \Gamma^l(r) + \Gamma^b(r)$, where $\Gamma^l(r)$ is the aggregate savings and $\Gamma^b(r)$ is the aggregate debt. First we need to show that $SD(r)$ is continuous in r and then show that there exists at least one r such that $SD(r) = \Gamma^l(r) + \Gamma^b(r) = 0$.

From the solution to the stationary problem, we know that the savings policy η^l and debt policy η^b are both continuous functions of r. The stationary densities g^h and g^l are also continuous in r by assumption. This means aggregate savings $\Gamma^l(r)$ and aggregate debt $\Gamma^b(r)$ are both continuous in r. Thus, $SD(r)$ is also continuous.

Alternatively, we can also see that $SD(r)$ is a linear combination of two linear operators $\Gamma^l(r)$ and $\Gamma^b(r)$ (both are expectation operators which are linear). From our model set-up, it is easy to see that the operator $SD(r)$ is bounded. Or, we could also require $|a^l|$ and $|a^b|$ to be Lebesgue integrable, which essentially says that the aggregate debt and aggregate loan/credit supply is finite, which is a rather mild assumption. This would imply we are looking at convergence in L^1-space. Then, we can use the well-known result that linear operators are bounded if and only if they are continuous. This means $SD(r)$ is continuous.

In the model, lenders always lend/save and borrowers always borrow with $r < \rho^l < \rho^b$. First consider the case where $r \to \rho^l$, $a^l \to \bar{a}$, where $\bar{a} \in (0, \infty)$. We then get $\lim_{r \to \rho^l} SD(r) = \bar{\Gamma}$ where $\bar{\Gamma} \in (0, \infty)$.

For the case where $r \to -\infty$, $a^b \to \underline{a}^l < 0$, we get $\lim_{r \to -\infty} SD(r) = \underline{\Gamma} < 0$.

We can see that $SD(r)$ lies between $-\infty$ and some positive value. Invoking the intermediate value theorem, there exists at least one r such that $SD(r) = 0$, i.e. a stationary equilibrium exists.

$\blacksquare$

Proof. **Theorem 1:** For reference, the borrowing household's dynamic HJB equation (2.19) is:

$$\rho^b V^b(a^b, z^b, t) = \max_c u(c^b) + \partial_a V^b(a^b, z^b, t)(\eta^b(a^b, z^b, t))$$
$$+ \partial_z V^b(a^b, z^b, t)\mu(z^l) + \frac{1}{2}\partial_{zz} V^l(a^l, z^l, t)\sigma^2(z^l)$$
$$+ \lambda \mathbb{E}_J[V^b(a^b, z'^b) - V(a^b, z^b)] + \partial_t V^b(a^b, z^b, t)$$

The envelope condition for both lenders and borrowers is obtained by differentiating both side of their HJB equation w.r.t. a^b, given by[1]:

$$(\rho^b - r)\partial_a V^b(a^b, z^b) = \partial_{aa} V^b(a^b, z^b) b(a^b, z^b) + \partial_{az} V^b(a^b, z^b)\mu(z^b) + \frac{1}{2}\partial_{azz} V^b(a^b, z^b)\sigma^2(z^b)$$
$$+ \lambda[\partial_a V^b(a^b, z'^b) - \partial_a V^b(a^b, z^b)] + \partial_{at} V^b(a^b, z^b)$$

Also, the first order condition (FOC) w.r.t. consumption c^b is $u'(c^b(a^b, z^b)) = \partial_a V^b(a^b, z^b)$.

Differentiating FOC w.r.t. a^b, we get: $\partial_{aa} V^b(a^b, z^b) = u''(c^b(a^b, z^b))\partial_a c^b(a^b, z^b)$,
w.r.t. z^b: $\partial_{az} V^b(a^b, z^b) = \partial_z u'(c^b(a^b, z^b))$ and $\partial_{azz} V^b(a^b, z^b) = \partial_{zz} u'(c^b(a^b, z^b))$.
w.r.t. time t: $\partial_{at} V^b(a^b, z^b) = \partial_t u'(c^b(a^b, z^b))$.

Also, differentiating the debt policy function (2.21) $\eta^b(a^b, z^b) = z^b_t + r_t a^b_t - c^b_t$, we get $\partial_a \eta^b(a^b, z^b) = r - \partial_a c^b_t(a^b, z^b)$. After default/jump occurs, borrowers' income is reduced by a certain amount, but this condition still holds whether or not default/jump has occurred.

[1]Note we have dropped the expectation operator $\mathbb{E}_J$ as we are assuming some fixed jump size here. This is for simplicity of notation, as including $\mathbb{E}_J$ presents no additional challenge. Our focus is not on the magnitude of the jump, but on the possibility of jump λ. Given the continuous model, any jump necessarily entails credit/default risk. Since we do not analyze the scenario involving various jump magnitude we choose fixed J, but show the optimization equations in their full generality.

Substituting and re-arranging these along with the envelope condition above (and using compact notation $\eta^b(a^b, z^b) = \eta^b$, $c_t^b(a^b, z^b) = c^b$) we get:

$$\eta^b(\partial_a \eta^b - r) = \frac{(r-\rho^b)u'(c^b) + \partial_z u'(c^b)\mu(z^b) + (1/2)\partial_{zz}u'(c^b)\sigma^2(z^b) + \partial_t u'(c^b) + \lambda(u'(c(z')^b) - u'(c^b)}{u''(c^b)})$$

Now, if $\eta^b = \eta^b(a^b, z^b) \to 0$, then we have $c^b(a^b, z^b) \to \underline{c}^b$ as $a^b \to \underline{a}^b$. This implies:

$$\eta^b(\partial_a \eta^b) \to \frac{(r-\rho^b)u'(\underline{c}^b) + \partial_z u'(\underline{c}^b)\mu(z^b) + (1/2)\partial_{zz}u'(\underline{c}^b)\sigma^2(z^b) + \partial_t u'(\underline{c}^b) + \lambda(u'(\underline{c}(z')^b) - u'(\underline{c}^b)}{u''(\underline{c}^b)}$$

Taking Taylor series approximation of $\eta^b(a^b, z^b)^2$ around $\underline{a}^b$ we get:

$$\eta^b(a^b, z^b)^2 \approx 2\eta^b(\underline{a}^b, z^b)(\partial_a \eta^b(\underline{a}^b, z^b))(a^b - \underline{a}^b)$$

Taking square root of this expression after substitution, we get the following dynamics:

$$\eta^b(a^b, z^b) \approx -\mathcal{S}_t(z^b)\sqrt{a^b - \underline{a}^b}$$

where

$$\mathcal{S}_t(z^b) = \sqrt{\left(2\frac{(r-\rho^b)u'(\underline{c}^b) + \partial_z u'(\underline{c}^b)\mu(z^b) + (1/2)\partial_{zz}u'(\underline{c}^b)\sigma^2(z^b) + \partial_t u'(\underline{c}^b) + \lambda(u'(\underline{c}(z')^b) - u'(\underline{c}^b))}{u''(\underline{c}^b)}\right)^+} > 0$$

For the stationary equilibrium, we just drop $\partial_t u'(\underline{c}^b)$ as we don't have the term $\partial_t V^b(a^b, z^b, t)$ in the HJB equation, which correspond to stationary HJB equation (2.30). We then get our desired expression for stationary case:

$$\mathcal{S}(z^b) = \sqrt{\left(2\frac{(r-\rho^b)u'(\underline{c}^b) + \partial_z u'(\underline{c}^b)\mu(z^b) + (1/2)\partial_{zz}u'(\underline{c}^b)\sigma^2(z^b) + \lambda(u'(\underline{c}(z')^b) - u'(\underline{c}^b))}{u''(\underline{c}^b)}\right)^+} > 0$$

For part (b), we integrate the evolution of optimal debt equation in stationary case as:

$$da_t^b = \eta^b dt,$$

$$\int_{a_0^b}^{a_t} \frac{da_s^b}{\sqrt{a_s^b - \underline{a}^b}} \approx \int_0^t -\mathcal{S}(z^b)ds,$$

$$2\left(\sqrt{a_t^b - \underline{a}^b} - \sqrt{a_0^b - \underline{a}^b}\right) \approx -\mathcal{S}(z^b)t,$$

$$a_t^b - \underline{a}^b \approx \left(\sqrt{a_0^b - \underline{a}^b} - \frac{\mathcal{S}(z^b)}{2}t\right)^2$$

Thus, the implications in theorem (1) follow from this equation because as $a_t^b \rightarrow \underline{a}^b$ the right side also goes to zero, which implies that time ($\mathcal{T}^b$) taken for convergence of $a_t^b \rightarrow \underline{a}^b$ is given by $\mathcal{T}^b \rightarrow (2\sqrt{a_0^b - \underline{a}^b}/\mathcal{S}(z^b))$ which is finite. ∎

Proof. **Corollary (1.1):** From theorem (1) we have $\eta(a^b, z^b) \approx -\mathcal{S}(z^b)\sqrt{a^b - \underline{a}^b}$, and thus $da^b \approx \eta^b(a_t, z_t)dt$ near the constraint. Clearly, at the constraint $\underline{a}^b$, η^b vanishes, as $\eta^b = 0$ at $\underline{a}^b$ for all income $z^b \leq z_c^b$ below some threshold level z_c^b.

Define the characteristic-exponent as $\Phi(\theta) := -Log\mathbb{E}(e^{i\theta a^b})$ for the random variable a^b, $\theta \in \mathbb{R}$. For the process $da^b \approx \eta^b(a_t, z_t)dt$, the Fourier Transform of a^b is: $E[e^{i\theta a^b}] = \hat{T} = e^{-\Phi(\theta)} = e^{-i\theta\eta(a^b, z^b)}$, where $i\theta\eta(a^b, z^b)$ is the characteristic-exponent of a^b.

We can also see that $\hat{T}$ is a tempered distribution as it is both linear and continuous functional in the space of Schwartz functions. Actually, Fourier transform defines a bijection from the space of Schwartz functions to themselves, and thus, is an isomorphism in the space of tempered distributions.

As $a^l \rightarrow \underline{a}^b$, we have $\eta(a^b, z^b) \rightarrow 0$, implying $\hat{T} \rightarrow 1$. We know that delta function is a tempered distribution and thus, has a well-defined Fourier transform which is 1 at the point of concentration. Further, inverse Fourier transform of constant $\hat{T} = 1$, a tempered distribution, yields the delta function given by: $\frac{1}{2\pi}\int e^{-i\theta a^b}\hat{T}d\theta = \delta(a^b - \underline{a}^b) = \hat{\delta}(\underline{a}^b)$, i.e. the delta function at $\underline{a}^b$.

■

Proof. **Corollary (1.2):** Using the equation in theorem (1) part (b) proof: $a_t^b - \underline{a}^b \approx \left(\sqrt{a_0^b - \underline{a}^b} - \frac{\mathcal{S}(z^b)}{2}t\right)^2$. For some $\tilde{a}^b \in \underline{\mathcal{B}}$:

$$\mathbb{P}(a_t^b < \tilde{a}^b) = \mathbb{P}\left(\underline{a}^b + \left(\sqrt{a_0^b - \underline{a}^b} - \frac{\mathcal{S}(z^b)}{2}t\right)^2 < \tilde{a}^b\right),$$

$$\mathbb{P}\left((\tilde{a}^b - \underline{a}^b) > \left(\sqrt{a_0^b - \underline{a}^b} - \frac{\mathcal{S}(z^b)}{2}t\right)^2\right)$$

From this expression, using the monotonicity property of measures, we can see that whatever makes the speed $\mathcal{S}(z^b)$ higher also makes the tail near the constraint heavier. From theorem (1) we have: $\mathcal{S}(z^b) \propto \lambda$ and $\mathcal{S}(z^b) \propto \frac{1}{\mathcal{A}^b}$, and the result follows. ■

Proof. **Theorem (2):** The proof of this theorem is almost similar to the proof of theorem (1) above, but with few changes to reflect lenders' problem.

Set $\lambda = 0$ and use the optimal drift of the savings/financial wealth as in equation (2.14):

$\eta^l(a^l, z^l) = z_t^l + r_t a_t^l - s_t(\alpha a_t^l) - z_t^l$.

Following similar steps as in proof of theorem (1), but now consider the case that if $\eta^l \to 0$, then $a^l \to \bar{a}^l$ for some initial $a_0^l > \bar{a}^l$, $c^l(a^l, z^) \to \bar{c}^l$. Thus, we get the analogous expression (just the stationary case shown here):

$\eta^l(a^l, z^l) \approx \mathcal{S}(z^l)\sqrt{\bar{a}^l - a^l}$, where $\mathcal{S}(z^l)$ is:

$$\mathcal{S}(z^l) = \sqrt{\left(2\frac{(r - \rho^l - s\alpha)u'(\bar{c}^l) + \partial_z u'(\bar{c}^l)\mu(z^l) + (1/2)\partial_{zz}u'(\underline{c}^l)\sigma^2(z^l)}{u''(\bar{c}^l)}\right)^+} > 0$$

Also, using similar steps to derive part (b) of theorem (1), we integrate the evolution of optimal saving/lending equation in stationary: $da^l_t = \eta^l dt$. We get: $\bar{a}^l - a^l_t \approx \left(\sqrt{\bar{a}^l - a^l_0} + \frac{\mathcal{S}(z^l)}{2}t\right)^2$.
So, $\mathbb{P}(a^l_t > \tilde{a}^l) = \mathbb{P}\left(\bar{a}^l - \left(\sqrt{a^l_0 - \bar{a}^l} + \frac{\mathcal{S}(z^l)}{2}t\right)^2 > \tilde{a}^l\right)$.
$\mathbb{P}\left((\bar{a}^l - \tilde{a}^l) > \left(\sqrt{a^l_0 - \bar{a}^l} + \frac{\mathcal{S}(z^l)}{2}t\right)^2\right)$

As seen, the parameter $\mathcal{S}(z^l)$ and α are inversely related i.e. $\mathcal{S}(z^l) \propto \frac{1}{\alpha}$. Using the monotonicity of measures, we can see that as CDS purchases α increases, all else equal, the right tail increases. This means the total credit supplied $E(a^l)$ also increases. All else equal, this means with *positive* probability the tail of the debt also increases.

■

Proof. **Theorem (3):** We have a Poisson process here, which is a renewal process. Thus, we will derive the result by using renewal arguments as in Feller (1971) by considering the time of first claim and size. For completeness, we will prove for general $X_0 = x$ first and then set $x = 0$ to get the closed form expression because there is no closed form expression for $x > 0$.

Recall, our general process in stochastic differential equation (SDE hereafter) form is: $dX_t = (\gamma X_t + \xi_t)dt - dY_t$. With $\gamma = 0$, the process considered for this proof is: $dX_t = \xi dt - dY_t$.

Re-writing the process in levels, $X_t = \xi t - Y_t$, for some initial value $X_0 = x$. Being more explicit about the total claims Y_t up to time t, we now write it as $Y_t = \sum_{i=1}^{N(t)} y_i$. Hence, $X_t = \xi t - \sum_{i=1}^{N(t)} y_i$.

Thus, we are now explicitly using the specific case of the subordinator Y_t, i.e. the well-known compound Poisson process: $N(t)$ is the Poisson process with intensity λ, same as that of the borrowers. The claims Y_i are sequence of i.i.d. variables, independent of $N(t)$ and drawn from the equilibrium leverage distribution K across the borrowers.

Note: While the actual leverage distribution of a^b in equilibrium is defined in the negative half-line, for the purpose of the financial sectors problem, we can reflect the same distribution about y-axis, such that now it is defined in the positive half-line. This is just a simple transformation for convenience and without loss of generality, as we can now just subtract this positive values as claims.

We know $E(N(t)) = \lambda t$. So, expected value of capital-surplus is:
$E(X_t) = E(\xi t - \sum_{i=1}^{N(t)} Y_i) = \xi t - E(N(t)E(y) = \xi t - \lambda E(y)t$.

Define $\bar{\pi}(x) = 1 - \pi(x)$ for notational simplicity, the non-failure probability.

Let τ_1 be the time of first claim. So, capital at τ_1 is $X_{\tau_1} = \xi\tau_1 - y_1$. As seen, the financial sector can not fail at least until first claim arrives i.e. not within time interval $(0, \tau_1)$. Then using renewal arguments as in Feller (1971):
$\bar{\pi}(x) = E(\bar{\pi}(x + \xi\tau_1 - y_1)) = \int_0^\infty \lambda e^{-\lambda s} \int_0^{x+\xi s} \bar{\pi}(x + \xi s - y)dK(y)ds$.

Let $z = x + \xi s$. Then, $\bar{\pi}(x) = (\lambda/\xi)e^{\lambda x/\xi} \int_x^\infty \lambda e^{-\lambda z/\xi} \int_0^z \bar{\pi}(z - y)dK(y)dz$.

Differentiating $\bar{\pi}(x)$: $\bar{\pi}'(x) = (\lambda/\xi)\bar{\pi}(x) - (\lambda/\xi) \int_0^x \bar{\pi}(x - y)dK(y)$.

Now, integrating $\bar{\pi}'(x)$ from 0 to t:
$\bar{\pi}(t) - \bar{\pi}(0) = \int_0^t [(\lambda/\xi)\bar{\pi}(x) - (\lambda/\xi) \int_0^x \bar{\pi}(x - y)dK(y)]du$.

Solving above equation and after some algebra, we get:

$\bar{\pi}(x) = \bar{\pi}(0) + (\lambda/\xi) \int_0^x \bar{\pi}(x-y)(1-K(y))dy$

Now the main problem is reduced down to to finding $\bar{\pi}(0)$.

We know that $E(y) = \mathbb{E}(y)$ from $\int_0^\infty (1-K(y))dy = \int_0^\infty \bar{\mathbb{P}}(y)^b dy = E(y)$.

Taking limits as $x \to \infty$, and using monotone convergence theorem on the integral on the right hand side of above equation (to take limits inside the integration), we get:

$\bar{\pi}(\infty) = \bar{\pi}(0) + (\lambda/\xi)E(y)\bar{\pi}(\infty)$

From our model, it is clear that $\bar{\pi}(\infty) = 1$ and thus, $\pi(\infty) = 0$, i.e. the financial sector does not fail if initial capital $x = \infty$. This is as expected and also intuitive.

Now, using $\pi(x) = 1 - \bar{\pi}(x)$, and setting $x = 0$ in above expression, we first get:

$1 - \bar{\pi}(0) = (\lambda/\xi)E(y)\bar{\pi}(\infty)$. Our result then follows as:

$\pi(x) = \frac{\lambda}{\xi} \int_{\mathbb{R}+} \bar{\mathbb{P}}(y)^b dy$.

∎

Proof. **Theorem (4):** Consider the Laplace transform of capital X_t: $E(e^{\theta X_t}) = e^{\psi(\theta)t}$ where $\psi(\theta)$ is the Laplace exponent. for all $\theta, t \geq 0$.

For $\gamma = 0$, the process is similar to that used in proof of theorem (3): $X_t = \xi t - Y_t$.

Then, $\psi(\theta) = \xi\theta - \int_{\mathbb{R}+}(1 - e^{-\theta y})\Pi(dy) = \xi\theta - \lambda \int_{\mathbb{R}+}(1 - e^{-\theta y})dK(y)$.

We see that the aggregate claims dynamics, represented by the drift less subordinator Y_t, is characterized by $\int_{\mathbb{R}+}(1 - e^{-\theta y})\Pi(dy)$ component.

We can see that $\psi(\theta)$ is strictly convex on $(0, \infty)$.

Using $(1 - K(y)) = \bar{\mathbb{P}}^b(y)$, and doing integration by parts gives:

$$\psi(\theta) = \xi\theta - \lambda\theta \int_{\mathbb{R}+} e^{-\theta x}(1 - K(y))dy$$

$$\psi(\theta) = \xi\theta - \lambda\theta \int_{\mathbb{R}+} e^{-\theta x}\bar{\mathbb{P}}^b(y)dy$$

Here, ψ is continuous and strictly convex in $(0, \infty)$. Let ϕ be the right-inverse function of ψ i.e. $\phi(q) = \sup\{\theta \geq 0 : \psi(\theta) = q\}$. Then we know ϕ is a strictly increasing, continuous, concave function. Because ϕ is concave, whatever affects ψ affects ϕ inversely. From this observations, we get our results and interpretations. For this, we need explicit relation showing dependence between $\mathcal{X}_\tau$ and ϕ.

Define $\mathscr{X}_t = \sup_{s \leq t} X_s$. We are considering times when the first claim has not yet arrived, i.e. all $t \leq \tau$, where τ is the first (arrival) time of default.

The process is $X_t = \xi t - Y_t$. And without any claims $Y_t = 0$ at time $t \leq \tau$, X_t is non-decreasing (monotone increasing also works assuming there is always some premium collected every period, but we choose the weaker condition here), and we get $\mathscr{X}_t = X_t$ for all $t \leq \tau$. Note, if $t > \tau$, this is no longer true, as X_t for $t > \tau$ is no longer non-decreasing such that $\mathscr{X}_t \neq \sup_{s \leq t} X_s$.

Then from the theory of Levy processes involving subordinators, $\mathscr{X}_t$, when evaluated/sampled at independently and exponentially distributed time τ, is exponentially distributed (see Bertoin (1996) or Kyprianou (2013) for more details).

In particular, setting $q = \lambda$, if τ_λ is an exponentially distributed random variable with parameter λ, then the distribution of $\mathscr{X}_{\tau_\lambda}$ is also exponentially distributed with parameter $\phi(\lambda)$, i.e. $\mathbb{P}(\mathscr{X}_{\tau_\lambda}) = 1 - e^{-\phi(\lambda)x}$.

In our model, default arrives at a random time τ with intensity λ and the distribution of the random time is given by the exponential distribution with CDF $P(\tau \leq t) = 1 - e^{-\lambda t}$. Since the random time of default by borrowers is exactly the random time τ with intensity λ, above results hold in our model.

Finally, the mean of the exponential distribution is: $E(\mathscr{X}_{\tau_\lambda}) = \mathcal{X}_\tau = \frac{1}{\phi(\lambda)}$, which gives the result. Also, using the result above that ϕ is the right-inverse function of ψ whose expression is derived above, the rest of the interpretation follows.

$\blacksquare$

Proof. **Lemma (1):** The solution of the SDE $dX_t = (\gamma X_t + \xi_t)dt - dY_t$ can be obtained by using simple variation of constant technique such that we get:

$X_t(x) = e^{\gamma t}\left(x + \int_0^t \xi e^{-\gamma s} ds - \int_0^t e^{-\gamma s} dY_s\right)$.

First integrate the following component of this equation: $\int_0^t \xi e^{-\gamma s} ds$.

For a given ξ, we can take it out of the integration as: $\xi \int_0^t e^{-\gamma s} ds$.

Solving further we get: $\xi[-\frac{e^{-\gamma s}}{\gamma}]_0^t = \xi[\frac{1}{\gamma} - \frac{e^{-\gamma t}}{\gamma}]$.

Setting $\xi = 0$ and re-defining $X_0 = x + (\xi/\gamma)$, we get the equation for X_t with $X_0 = x + (\xi/\gamma)$. Then the result follows.

$\blacksquare$

Proof. **Theorem (5):** Consider the SDE of capital evolution in equation (2.9):
$dX_t = (\gamma X_t + \xi_t)dt - dY_t$

Using simple variation of constant technique, the solution of this SDE is:

$$X_t(x) = e^{\gamma t}\left(x + \int_0^t \xi e^{-\gamma s} ds - \int_0^t e^{-\gamma s} dY_s\right).$$

$X_t(x)$ with x in bracket is to emphasize the general case, i.e. initial value $X_0 = x \geq 0$. Also, consider another equivalent process that we will need below:

$$X_t(x + (\xi/\gamma)) = e^{\gamma t}\left(x + (\xi/\gamma) - \int_0^t e^{-\gamma s} dY_s\right).$$

Using our definition, $\tau_{-(\xi/\gamma)} = \inf\{t > 0 : X_t < -(\xi/\gamma)\}$ for $X_0 = x$, and from above expression we get the following:

$$\pi_t(x) = \mathcal{P}(\tau_{-(\xi/\gamma)} \leq t) = \mathbb{P}(X_t(x) < -(\xi/\gamma)).$$

Using Lemma (1), we get:

$$\mathbb{P}\left(X_t(x) < -(\xi/\gamma))|X_0 = x\right) = \mathbb{P}\left(X_t(x + (\xi/\gamma) < 0|X_0 = x + (\xi/\gamma)\right).$$

Then (dropping the conditional part for notational convenience), we get:

$$\mathbb{P}(X_t(x + (\xi/\gamma)) < 0) = \mathbb{P}(e^{\gamma t}(x + (\xi/\gamma) - \int_0^t e^{-\gamma s} dY_s)) < 0)$$

$$\mathbb{P}(x + (\xi/\gamma) < \int_0^t e^{-\gamma s} dY_s).$$

$$\mathbb{P}(\int_0^t e^{-\gamma s} dY_s > x + (\xi/\gamma)).$$

Set $\mathcal{V}_0 = (x + \xi/\gamma)$ and $\mathcal{V}_t = \int_0^t d(e^{-\gamma s} Y_s) = \int_0^t e^{-\gamma s} dY_s$, the result follows that $\pi_t(x) = \mathcal{P}(\tau_{-(\xi/\gamma)} \leq t) = \mathbb{P}(\mathcal{V}_t > \mathcal{V}_0)$.

Applying the monotonicity property of measures on above equation of $\pi_t(x)$, the rest of the results also follow.

Interpretation: Now we recast/re-interpret the result in terms of present value. As seen, $-(\xi/\gamma)$ is the threshold level at any given time t. Default occurs when the event $(X_t < -(\xi/\gamma)$ occurs. From the usual present value method, for the financial sector to fail, the present value of the all the future flow of the claims till time t must be greater than the present value of the surplus amount over and above what is paid out from the claims.

Thus, with the present value method, we can evaluate all future flow of claims against the the current time. Then the threshold value in terms of the present value is $x + (\xi/\gamma)$, i.e. if the *discounted* present value of the total flow of future claims exceeds the threshold level $x + (\xi/\gamma)$ today, the financial sector fails. Since the event is random, we characterize the failure probability.

Here, $\mathcal{V}_0 = (x + \xi/\gamma)$ denotes the value of initial capital x plus the total *discounted* premium (ξ/γ). This is the solvency threshold in terms of present value, while the solvency threshold $\bar{x} = -(\xi/\gamma)$ applies at any given time t.

Now, $\mathcal{V}_t = \int_0^t d(e^{-\gamma s} Y_s) = \int_0^t e^{-\gamma s} dY_s$ denotes the total *discounted* value of the *flow* of all the future claims till time t, where $(e^{-\gamma s} Y_s)$ is the discounted value of the aggregate claims at time $s \leq t$. By integrating we are adding up all those terms. Then we easily get the expression for the failure probability:

$\pi_t(x) = \mathcal{P}(\tau_{-(\xi/\gamma)} \leq t) = \mathbb{P}(\mathcal{V}_t > \mathcal{V}_0)$.[2]

$\blacksquare$

Proof. **Theorem (6):** Recall that *action* functional is $S[x_t] = \int_{t_i}^{t_f} L(\dot{x}, x, s) ds$. Then

[2] This proof is easily/concretely understood if we think about the claims process as being a compound Poisson process, a special case of the general subordinator Y_t. Since claims arrive at discrete sequence of random time $\tau_1, \tau_2, ...\tau_n, ...$, that would allow for calculation immediately after the n^{th} claim. Then the integral would be replaced by sum over all the future discounted claims against the same threshold $\mathcal{V}_0 = (x + \xi/\gamma)$, yielding the same result and interpretation as before.

the transition probability as in Feynman Path Integral is given by FPI:

$$K(x_i, t_i; x_f, t_f) = \int_{x(0)=x_i}^{x(t)=x_f} \mathcal{D}x(t) e^{-S[x_t]}$$

Now, the entire problem is reduced to finding the Lagrangian $L(\dot{x}, x, s)$ of the stochastic process under consideration, as it summarizes the dynamics of the system. Also, recall that a SDE is well defined when both the continuous/differential expression and the discretization rules are given. Our main evolution equation is $da_t^b = \eta_t^b(a_t, z_t)dt$ (analogous for a^l).

We prove our result in two ways: (i) analogous to the way such problems are represented/solved in physics (this approach can appear slightly heuristic, but they do have strong theoretical foundations), and (ii) by usual mathematical methods.

(1st approach): Here we will use notations/labels that are common in physics for ease of reference. For the process $da_t^b = \eta_t^b(a_t, z_t)dt$, we have $v = \dot{a}_t^b = \eta_t^b$, which is the velocity/speed. Let $p = mv$ be the momentum. Since mass m is a physical property and may not have direct interpretation in economic settings, we can normalize $m = 1$ without loss of generality, which gives $p = v = \eta_t^b$.

Let $T = \frac{p^2}{2m} = \frac{(\eta_t^b)^2}{2}$. This is the kinetic energy, which is the function of p. Also, let V be the potential energy which is the function of position a^b.

Note that Lagrangian characterizes the dynamics of the system. But the Lagrangian need not be unique for a given system. It just needs to generate correct laws (of probability or physics). Since we do not say anything about the uniqueness of our equilibrium/optimal process, it is plausible that other equally correct versions of Lagrangian can be specified. However, all versions must necessarily give rise to the same laws. Thus, the the non-relativistic Lagrangian (as in our model) can be broadly defined as $L = T - V$, which is independent of particular specifications of T and/or V. We elaborate on this below.

The Hamiltonian operator H which represents the total energy of the system is $H = T + V$. This Hamiltonian is the *Legendre transform* of the Lagrangian for a given a_t at time t (the position of the process/particle) such that $p = \eta_t^b$ is the dual variable. Since we are in one dimensional case (a_t^b) we get:

$$H = \dot{a}_t^b \frac{\partial L}{\partial \dot{a}_t^b} - L = \dot{a}_t^b p - L$$

where $\frac{\partial L}{\partial \dot{a}_t^b} = p(\dot{a}_t^b, a_t^b, t)$ is the definition of momentum via Lagrangian. Note, for multiple dimensions (say a 3-dimensional space) we would just sum over $\frac{\partial L}{\partial \dot{a}_t^b}$ for each dimension because velocity can be different for different co-ordinates. Then, substituting using the definition $p = v$, we get: $H = T + V$ and thus, $L = T - V$.

By observation, our process $da_t^b = \eta_t^b(a_t, z_t)dt$ is a function of $p = \eta_t^b$ only. η_t^b is the debt policy function and thus, indirectly depends on a_t^b. But the optimally chosen η_t^b is an endogenous object, and thus the dependence on a_t^b (optimal choice) along with dependence on other variables like income z_t^b is already encoded into this compact and complicated object η_t^b. This can also be seen by defining a new process x_t such that $dx_t = da_t^b$ (we will further explore this aspect in the 2nd approach below), such that velocity $\dot{x}_t = \dot{a}_t^b = \eta_t^b$.

Thus, the process/particle now acts like a *free particle* (though not in an exactly sense because the distribution of a_t^b is not standard normal). This means that there is no additional influence of the potential V, and the dynamics of the system is governed by just the kinetic energy $T = \frac{v^2}{2} = \frac{(\eta_t^b)^2}{2}$. Without the potential V, the Lagrangian is $L = T$, and the result $L = \frac{(\eta_t^b)^2}{2}$ follows.

Note, the result from this 1st approach might appear somewhat heuristic, but this is due to the way we have reduced and simplified our problem to get the desired Lagrangian. While the main dynamics of a_t^b looks like a simple/linear drift process at first glance, this dynamics is really complicated once we focus on the drift term (the policy function) itself which is a highly non-linear equilibrium determined object.

(2nd approach): Here we will work with just mathematical arguments, without relying on the physical interpretations as above. Define a new process x_t whose dynamics is governed by the stochastic process a_t^b. This is analogous to a drift-diffusion process where the stochastic nature is governed by the diffusion/Brownian motion part. The strategy is to first get the Lagrangian for x_t process and then another Lagrangian for the a_t^b process itself as a special case of x_t. We do this longer derivation just for completeness and to show some general structures behind the scene. Thus, define a stochastic process x_t as:

$$dx_t = m(x_t)dt + \sigma(x_t)da_t^b$$

where $da_t^b = \eta_t^b(a_t, z_t)dt$, i.e. $\dot{a}_t^b = \eta_t^b$.

The discretized version of this process can be written as: $\Delta x_t = m(x_t)\Delta t + \sigma(x_t)\Delta a_t^b$.

This can be further generalized by making the discretization depend on some parameter, say ζ, such that we get Itô and Stratonovich stochastic integral interpretations. See Gardiner (1985) for more details. This emphasizes the point we made earlier – any continuous stochastic process is well-defined only when both the continuous version and the discretization rule are given. Consider the following discretization rule:

$$x_{j+1} - x_j \approx (\zeta m(x_{j+1} + (1-\zeta)m(x_j))\epsilon + \sigma(x_t)(a^b_{j+1} - a^b_j)$$

where $\Delta t = \epsilon = t/N = (t_f - t_0)/N$, $t_j = j\epsilon$, $x_j = x(t_j)$.
Taking the limit $N \to \infty$, $\epsilon \to 0$ after substituting , the Lagrangian for such a process is well-know:

$$L(x, \dot{x}, t) = \frac{1}{2\sigma^2(x_t)}(\dot{x}_t - m(x_t))^2 + \zeta \frac{dm(x_t)}{dx}$$

For $\zeta = 0$, we get Itô integral interpretation. For $\zeta = 1/2$, we get Stratonovich integral interpretation.

Thus, to get the Lagrangian for our process, we set $\zeta = 0$, $m(x_t) = 0$ and $\sigma(x_t) = 1$. This is now equivalent to our original process, and the corresponding Lagrangian is:
$\frac{1}{2}\dot{x}_t^2 = \frac{1}{2}(\dot{a}^b_t)^2 = \frac{1}{2}(\eta^b_t)^2$.

Note, if we had a *free particle*, we could take $\dot{x}_t^2$ out of the integral in the action $S[x_t]$. This would yield a normal density (after specifying the appropriate normalization constant), representing a process driven by a Brownian motion only as $dx = dw_t$. This is exactly as we would expect from a Brownian motion (particle/process). However, given the complicated nature of our drift term η^b_t, we cannot take it out of the integral. Thus, we just get the the analytic expression which is sufficient for our purposes.

$\blacksquare$

Proof. **Corollary (6.1):** This proof follows from the discretized version of the action functional: $S[x_t] = \int_{t_i}^{t_f} L(\dot{x}, x, s)ds$ and the corresponding Lagrangian $L(\dot{x}, x, t) = \frac{1}{2}(\eta_t^b)^2$ obtained in the proof of theorem (6) above. Note, we are interchangeably using x and a^b as we have $dx_t = da_t^b$ from theorem (6). This gives the approximated/discretized version of the transition probability before taking the limit $N \to \infty$, given that $K(x_i, t_i; x_f, t_f) = \int_{x(0)=x_i}^{x(t)=x_f} \mathcal{D}x(t)e^{-S[x_t]}$. Thus, the discretized action/Lagrangian for $\Delta t = \epsilon$ is:

$$S(x_j, t_j; x_{j-1}, t_{j-1}) = \epsilon \sum_{i=0}^{n-1} L((x_j - x_{j-1})/\Delta t, x_{j-1}, t_{j-1})$$

$$S(x_j, t_j; x_{j-1}, t_{j-1}) = \frac{1}{2}\left(\frac{x_j - x_{j-1}}{\epsilon}\right)^2 \epsilon$$

where $a_i^b = x_{j-1}$, $a_f^b = x_j$.

For a given a_i, whenever $a_{f'} > a_f$, the difference $a_f - a_i$ is increasing, so the action $S(.)$ above is increasing, and the negative exponential of action, i.e. $e^{-S(x_j,t_j;x_{j-1},t_{j-1})}$, is decreasing. We know that the transition probability $\mathbb{K}$ is directly related to the transition density function $K(x_i, t_i; x_f, t_f)$. In particular, transition density function is the Radon-Nikodym derivative of transition probability measure, i.e. $\frac{d\mathbb{K}}{da_f} = K(x_i, t_i; x_f, t_f)$.

Thus, the transition probability $\mathbb{K}$ is declining and the rate of decline is exponential. Note, this probability is a close approximation of the continuous time case given the discretization. Since we are only comparing two probabilities, and not characterizing the exact probability level, making comparison via approximation is appropriate. Then the result follows.

∎

Proof. **Corollary (6.2):** We know: $\mathbb{P}(a_t^b \notin \underline{\mathcal{B}}|a_0^b \in \underline{\mathcal{B}}) + \mathbb{P}(a_t^b \in \underline{\mathcal{B}}|a_0^b \in \underline{\mathcal{B}}) = 1$

where $\mathbb{P}(a_0 \in \underline{\mathcal{B}}; a_t \in \underline{\mathcal{B}})$ is the probability of staying in $\underline{\mathcal{B}}$ given that we are in $\underline{\mathcal{B}}$.

Also, by definition $\mathbb{P}_{\tau_e}(t) = \mathbb{P}_{\tau_e}(\tau_e < t) = \mathbb{P}(a_t^b \notin \underline{\mathcal{B}}|a_0^b \in \underline{\mathcal{B}})$.

Then, $\mathbb{P}_{\tau_e}(t) = 1 - \mathbb{P}(a_0 \in \underline{\mathcal{B}}; a_t \in \underline{\mathcal{B}}) = 1 - \int_{\underline{\mathcal{B}}} K^b(a_0 \in \underline{\mathcal{B}}; a_t \in \underline{\mathcal{B}}) da^b$.

∎

Proof. **Corollary 6.3:**

The transition density function $K(a_i, t_i; a_f, t_f)$ (with associated probability $\mathbb{K}$) determines the evolution of the spatial distribution from any given initial condition K_0 (with associated probability $\mathbb{K}_0(a_i)$). The superscript b is dropped from $\mathbb{K}^b$ for convenience.

Note that K (shorthand for $K(a_i, t_i; a_f, t_f)$) solves the Kolmogorov Forward equation associated with the evolution of the optimal debt a_t^b. Given that our economy is driven by Markov process, the initial position does not have to be deterministic, and we can always consider some other later time as the initial time and take the distribution at that time as the initial distribution.

Let $\partial_t \mathbb{K}$ denote the rate of change of transition probability $\mathbb{K}$ which is directly related to the transition density function K by the Radon-Nikodym theorem as discussed in the proof of corollary (6.1) above.

Thus, without loss of generality, we could have also taken the partial derivative of K to get the results. To avoid writing the integral $\int_{a_i}^{a_f} \mathcal{D}a(t)$ repeatedly, we abbreviate the relation with proportionality symbol $\propto$.
Then, for $K(a_i, t_i; a_f, t_f) = \int_{a_i}^{a_f} \mathcal{D}a(t) e^{-S[a_t]}$, we get
$\partial_t \mathbb{K} \propto \frac{\partial e^{-S[x_t]}}{\partial t}$, i.e. the time derivative of the exponential of the negative of action functional.

Given the limits of the integration from a_i to a_f, and that the Lagrangian satisfies measurability and integrability conditions, we can just use the simpler version of Leibniz's integral rule and take partial derivative inside the integral as: $\frac{\partial e^{-S[x_t]}}{\partial t}$.

We know that the total derivative of the action S is $\frac{dS}{dt} = \frac{\partial S}{\partial t} + \frac{\partial S}{\partial a}\dot{a}$
The evolution of debt $\dot{a} = \eta_t^b$ is determined optimally in equilibrium. Also, we have fixed the initial position (x_i, t_i) and the final position (x_f, t_f) in space-time. Thus, we get $\frac{\partial S}{\partial a} = 0$ by using calculus of variations techniques as the path described by $\dot{a}$ is an optimal/extremum path.

Let L be the abbreviation for Lagrangian $L(\dot{a}, a, t) = (1/2)(\eta_t^b)^2$.
The action is $S[a_t] = \int_{t_i}^{t_f} L(\dot{a}, a, s)ds$

Then, $\frac{\partial e^{-S[x_t]}}{\partial t} = e^{-S[x_t]}(-dS/dt) = -Le^{-S[x_t]}$.
To see if the transition speed $\partial_t \mathbb{K}$ is increasing or decreasing in η_t^b, differentiate with respect to η_t^b. This gives:
$\frac{L}{e^{2S[x_t]}}(Le^{\int \eta ds} - \eta e^{S[x_t]})$. This expression is positive for $\eta_t^b < 0$ and the result follows.

For part (b) with $\eta_t^b > 0$, assume that $\frac{\eta e^{\int \eta_s ds}}{2e^{S[x_t]}} < 1$, then the transition speed $\partial_t \mathbb{K}$

is decreasing.

■

Proof. **Corollary 6.4:** We get the following *Euler-Lagrange* equation by applying the variational principle to the action $S[a_t]$:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{a}}\right) = \frac{\partial L}{\partial a}$$

Using the definition of the Lagrangian L in this equation and differentiating, result follows.

■

Proof. **Theorem 7:** The evolution of loan/credit is $da_t^l = \eta_t^l dt$ where η_t^l is the optimally chosen drift.
The Lagrangian is also similar: $L(\dot{a}, a, t) = (1/2)(\eta_t^l)^2$.
So the action is: $S[a_t] = \int_{t_i}^{t_f} L(\dot{a}, a, s)ds$.
Lenders' density is similarly given by $K(a_i, a_i; a_f, a_f) = \int_{a_i}^{a_f} \mathcal{D}a(t)e^{-S[a_t]}$.
Also recall: $\eta_t^l = z + ra - s(\alpha a) - C$, where the time subscripts are dropped for convenience as we are at a given time t.
Differentiating $E(a^l)$ with respect to α, we get $\frac{d}{d\alpha}\int_{a_i}^{a_f} \mathcal{D}a(t)ae^{-S[a_t]}$. Again, using simpler version of Leibniz's integral rule as in the proof of corollary 6.3 above, we can differentiate under the integral.

Since α is exogenous, for a given a^l, all else equal, the result is obtained by just differentiating $e^{-S[a_t]}$ and evaluating the sign of the derivative:
$-e^{-S[a_t]}(\frac{dS}{d\eta_t}\frac{d\eta_t}{d\alpha}) = -e^{-S[a_t]}(\int \eta_t dt(-sa)) = \frac{(sa)\int \eta_t dt}{e^{S[a_t]}}$

Thus, the results (a) and (b) follow based on whether $\eta_t^l > 0$ or $\eta_t^l < 0$.

For (c), steps are similar, the only difference is that instead of the expectation of a^l, we now look at the measure over set $\bar{\mathcal{B}}$ to the right tail. As above, we arrive at the final step $\frac{(sa)\int \eta_t dt}{e^{S[a_t]}}$, such that the measure of the set $\bar{\mathcal{B}}$, i.e. $\mathbb{P}(a^l \in \bar{\mathcal{B}})$, is increasing in α.

$\blacksquare$

A.3 Operators

This appendix discusses some basics of operators used in deriving some of our results. We make use of Kolmogorov backward and forward equations (KFE and KBE hereafter). These partial differential equations (PDE) are central to quantitative/mathematical finance and related literature. For example, the famous Black-Scholes-Merton formula can be expressed via *Feynman-Kac* PDE.

In particular, KBEs are useful for problems where final state is known, such as exit time problems or stopping time problems, or options pricing via standard Black-Scholes-Merton formula. Many exit and stopping time problems make use of KBEs since these equations characterize the perturbations of the initial condition when the final state/condition for exit or stopping is known. On the other hand, KFEs is extremely useful in characterizing the evolution of PDEs forward in time given some initial state/condition. For example, evolution of transition density function (as we have in this chapter) can be compactly characterized via KFEs.

Let $\mathcal{L}$ be the differential operator and $\mathcal{L}^*$ be another differential operator that is the adjoint of $\mathcal{L}$ in L^2-space. For any C^2 functions (or $C^{2,1}$ with explicit space-time representation) $f(x)$ and $h(x)$ we have the following relation:

$$\langle f(x), h(x) \rangle = \int_{R^n} f(x)h(x)d\mu(x)$$
$$\langle \mathcal{L}f(x), g(x) \rangle = \langle f(x), \mathcal{L}^* g(x) \rangle$$

Note $d\mu(x)$ can be substituted with the Lebesgue measure dx depending on the setting. Here $\mathcal{L}$ is the infinitesimal generator and $\mathcal{L}^*$ is the adjoint. Also, if we are in finite dimension R^n, $\mathcal{L}$ can be represented by a matrix, and $\mathcal{L}^*$ is the corresponding transpose $\mathcal{L}^T$.

Specifically (as in our model), the infinitesimal generator of the jump-diffusion process for a C^2 function $f(.)$ is:

$$\mathcal{L}f(z) = \mu(z)\partial_z(f(z)) + \frac{1}{2}\sigma(z)^2\partial_{zz}(f(z)) + \lambda\mathbb{E}_J[f(z') - f(z)]$$

The usual Kolmogorov backward equation is then: $\partial_t f(z) + \mathcal{L}g(z) = 0$. Then, the Kolmogorov forward equation is: $\partial_t f(z) - \mathcal{L}^* f(z) = 0$, where $\mathcal{L}^*$ is the adjoint of $\mathcal{L}$:

$$\mathcal{L}^* f(z) = -\partial_z[\mu(z)f(z)] + \frac{1}{2}\partial_{zz}[\sigma(z)^2 f(z)] + \lambda\mathbb{E}_J[f(z - J) - f(z)]$$

If we consider a constant jump J, the last term will be $\lambda[f(z'') - f(z)]$. Once the drift (μ) and diffusion (σ) terms are specified (generally held constant, corresponding to some calibrated or closely approximated values), computing the equilibrium and understanding/illustrating various model properties become simpler and very efficient.

We note one useful and remarkable property of operators. Any linear operator is bounded if and only if it is continuous. Often times, boundedness is much more easier to check than continuity. Once boundedness is confirmed, then continuity is guaranteed. Further, the eigenvalue of the operator can also be related to the speed of convergence as in Gabaix et al. (2016). Bounded self-adjoint operators are particularly interesting because their spectrum is real and lies in a bounded closed interval. These are some of the few aspects of operators that make them extremely appealing for economic applications. We do not employ all these high level methods in this chapter because our equilibrium structure makes it harder to exogenously impose some crucial assumptions needed for using the results from operator theory. That is why Feynman Path Integral method appears more suitable to our needs. We refer the interested readers to the various references contained in Gabaix et al. (2016) regarding operator theory. Other accessible references include Kreyszig (1978) and Schechter (2002).

A.4 Theory of Lévy Processes

The concepts and results from the theory of Lévy Process used in chapter 2 are fairly standard and accessible. As such, we only briefly review some basics of this theory, and refer the readers to some standard texts on the topic, such as those by Kyprianou (2014), Bertoin (1996), Applebaum (2009) and Cont and Tankov (2004). In short, the basic idea is to transform the problem from the real plane to the complex plane where problems are relatively easier to handle. After some mathematical manipulations and analytical continuations, we transform the problem back to the real plane.

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ denote the filtered probability space characterizing randomness in the economy with filtration $\mathbb{F} = \{\mathcal{F}\}_{t\geq 0}$. A real valued stochastic process

$(X_t)_{t\geq 0}$ defined on this space is called a *Lévy Process* if:

(i) the sample paths are adapted, cadlag i.e. right-continuous with left limits existing a.s. $\mathbb{P}$, and

(ii) it has stationary, independent increments i.e. for any $0 \leq s \leq t$, the increments $X_t - X_s$ are independent and drawn from the same distribution.

Thus, Lévy process can be represented in general form as:

$$X_t = \mu t + \sigma W_t + Y_t$$

where Y_t is the jump process. Details of each component in this general form is given by Lévy-Itô decomposition. Some standard examples of Lévy process are deterministic linear drift, Brownian motion, Poisson and compound Poisson processes. Sum of Lévy process is again another Lévy process. This is how we obtain our Lévy jump-diffusion process for the income of the borrowers – by combining Brownian motion with drift and compound Poisson process. This is like a generalization of Black-Scholes-Merton model of stock prices. For the capital evolution process of the financial sector, we just omit the Brownian motion component. In financial applications, such as modeling of stock price evolution, prices can be represented as the exponential of some Lévy process. This ensures that the price is positive while still allowing for log-returns to be independent and stationary.

The characteristic exponent of a Lévy process with Lévy triple (μ, σ, Π) is given by the *Lévy- Khintchine* formula:

$$\Psi(u) = iu\xi + \frac{1}{2}\sigma^2 u^2 + \int_{\mathbb{R}} (1 - e^{iux} + iux\mathbb{1}_{(|x|<1)})\Pi(dx)$$

where $u, \mu, \sigma \in \mathbb{R}$ and Lévy measure Π is concentrated on $\mathbb{R} \setminus \{0\}$ such that $\int_{\mathbb{R}}(1 \wedge x^2)\Pi(dx) < \infty$. An important subclass of Lévy process is called *subordinators* that have non-decreasing sample paths. Examples include the Poisson and the inverse Gaussian process. Another special case are Lévy process with negative jumps called *spectrally negative* Lévy process. Such process have jumps only in one direction and are very useful for many calculations. Further, various risk processes can be modeled as a *spectrally negative* Lévy process, as we have done in this chapter. This allows for applying results of this theory to understand various systemic risk features.

For a mathematical exploration of Lévy type stochastic processes similar to that in our model, see Patie (2005) and Loeffen and Patie (2010). For a mathematical treatment of risk processes/ruin theory, see the survey paper by Paulsen (2008) and books by Albrecher and Asmussen (2010) and Kyprianou (2013).

We now discuss application of *Lévy- Khintchine* formula in our model. Let us recall the evolution of the financial sector's surplus-capital X_t:

$$dX_t = (\gamma X_t + \xi_t)dt - dY_t$$

The distribution of X_t can be characterized by the *Lévy- Khintchine exponent (i.e. characteristic exponent)* $\Psi(u)$:

$$\Psi(u) = iu\xi + \int_{\mathbb{R}-}(1 - e^{iuy})\Pi(dy)$$

$$\mathbb{E}[e^{iuX_t}] = e^{-t\Psi(u)}, u \in \mathbb{R}$$

Or, equivalently, via the related Laplace exponent $\psi(\theta)$:

$$\psi(\theta) = \theta\xi - \int_{\mathbb{R}+}(1 - e^{-\theta y})\Pi(dy)$$

$$\mathbb{E}[e^{\theta X_t}] = e^{t\psi(\theta)}, \theta \geq 0$$

where $\Pi(dy) = \lambda K(dy)$ is the Lévy measure, for some claims distribution F on $\mathbb{R}^+$. In our model, the claims distribution is the equilibrium leverage distribution across borrowers. The aggregate claims dynamics is represented by the driftless subordinator Y_t, which can be characterized by $\int_{\mathbb{R}+}(1 - e^{-\theta y})\Pi(dy)$ component of the Laplace exponent $\psi(\theta)$. Here, ψ is continuous and strictly convex in $(0, \infty)$. Let ϕ be the right-inverse function of ψ i.e. $\phi(q) = \sup\{\theta \geq 0 : \psi(\theta) = q\}$. Then we know ϕ is a strictly increasing, continuous concave function.

These simple mathematical structures, coupled with relevant model structures, are used for deriving various results in this chapter that have generated deeper insights. The flexibility/generality of the techniques of Lévy processes and their natural applicability to economics/finance are also clearly reflected in the simplicity of the proofs.

A.5 Feynman Path Integral

In the main body of chapter 3, we provided some discussions of quantum mechanics methods, that were largely intuitive. In this appendix we will build upon those discussions, highlighting relevant mathematical aspects. To characterize transition dynamics analytically, we need at the very least, the analytic formulation of transition probability. Except for few simple cases, obtaining such an analytic expression for many general process that describe the dynamic economy is either very difficult or even impossible. Fortunately, some quantum mechanics methods, namely the Hamiltonian and Feynman Path Integral (FPI hereafter), can be immensely helpful in this regard for addressing various economic problems.

Hamiltonian is generally used for simpler, non-linear cases. It is a differential operator, a partial differential equation describing the time evolution of quantum dynamical system. It appears in context of $Schrödinger$ equation which is analogous to Kolmogorov equations for diffusions. FPI is generally used for more complex cases where the entire trajectory of a process can be considered. Path Integral can be defined as a limit of the sequence of finite dimensional integrals, similar to the way Riemann integral is defined. Further, FPI includes outcomes of Hamiltonian formalism as a special case. These two widely used formalisms of quantum mechanics can gives us transition probabilities necessary for charactering the dynamics of economic system as it evolves from some given initial conditions, possibly towards some long run stationary equilibrium.

Here, we briefly review some important concepts from the Path Integral formalism because it is most suitable in our modeling context, while also noting the Hamiltonian formalism wherever possible. We only attempt to provide a broad but brief overview, just enough for getting the essence of the formalism and how that is used in this chapter. Application of this formalism to this chapter is sufficient to highlight the general applicability of these methods in a wider range of economic settings. For more details, we refer to some excellent standard texts on the subject: Feynman and Hibbs (1965, classic book on the subject), Sakurai (1994, standard graduate textbook), Shankar (2011, standard graduate textbook) and Griffiths (2004, standard undergraduate textbook). Feynman (1948) is the classic paper that introduced Path Integral as a new approach to non-relativistic Quantum Mechanics. Baaquie (2004) shows how the mathematics and ideas of Quantum Mechanics can be used for financial applications (eg. modeling of interest rates and option pricing), independent of the commonly used stochastic calculus methods. Linetsky (1998) reviews applications of Path Integral to options pricing and financial modeling. Goovaerts, Schepper and Decamps (2004) show how transition density for general diffusion processes can be expressed via FPI, allowing for exact calculation for various classical financial models.

Let $K(t_i, x_i; t_f, x_f)$ be the transition density function of going from x_i at some initial time t_i to $x_f \in \mathcal{B}$ (a Borel set) at time $t_f > t_i$. We only need to care about the elapsed time $\tau = t_f - t_i$. So we re-write this expression as $K(x_i, \tau; x_f)$.

Hamiltonian: Now we express our problem first via Hamiltonian just for completeness and then move to FPI. The time evolution of the state of the system i.e. the wave function $\psi(t)$ is governed by the *Schrödinger equation* given by:

$$i\hbar \frac{\partial \psi(t)}{\partial t} = H\psi(t)$$

where H is the Hamiltonian operator (representing the total energy of the system) and $\hbar = \frac{h}{2\pi}$ is the reduced Plank constant. The solution can be written as $\psi(t) = U(t, t_0)\psi(t_0)$ where $U(.)$ is a unitary operator called the evolution operator. Such operators form the continuous one-parameter group generated by the operator $-iH/\hbar$ such that $\psi(t) = e^{-\frac{i}{\hbar}H(t-t_0)}\psi(t_0)$. Also, the wave function $\psi(t)$ describes the *probability amplitude* to go from one point to the other, and its square gives the probability density of finding the particle in certain region of space, given by:

$$\int_R |\psi(x,t)|^2 dx = 1$$

From Kolmogorov equation we have:

$$\frac{\partial K(t,x)}{\partial t} = -\mathcal{L}K(t,x)$$

Denote $-\mathcal{L} \equiv H$. Using Dirac's notation and after some algebraic manipulations we get the transition density from Hamiltonian as: $K(x_i, \tau; x_f) =< x_i|e^{-\tau H}|x_f >$. For example, in a finite dimensional case, the right hand side of this equation can be read, in order, as row vector times matrix times column vector[3]. If the Hamiltonian operator (and the associated Kolmogorov forward operator) is allowed to be fully non-linear, Hamiltonian method can be too cumbersome to yield any clear insights or interpretations. However, Feynman's approach (discussed below), can work smoothly in such complicated cases as well.

Now, after performing Fourier transform of $|x>$ basis to momentum space, using completeness equation for momentum basis, denoted as $|p>$. After some mathematical manipulations of Hamiltonian using Dirac's notation (see Baaquie (2004) for more details), we get the expression for the kernel of transition density as: $K(x_i, \tau; x_f) =< x_i|e^{-\tau H}|x_f >= \frac{1}{2\pi}\int e^{ip(x_i - x_f + \tau\mu)}dp$, where μ is the drift of a simple Brownian diffusion with drift process $dx_t = \mu(x,t)dt + \sigma(x,t)dW_t$, where $\sigma = 0$.

[3]Dirac (1930), one of the early pioneers of quantum mechanics, introduced a very compact and efficient formulation of the mathematics to describe the quantum systems by using new notation for vectors and inner products. Dirac's notation of *bra* vector, denoted as $< a|$ and *ket* vector, denoted as $|a>$, are simply row and column vectors respectively. These vectors can be of any size and can be generalized to their continuous counterparts. The linear operator between the *bra* and *ket* vectors is essentially a square matrix in finite dimension with analogous counterpart in continuous case like the Hamiltonian operator above. The *ket* vector $|a>$ lies in the Hilbert space, and *bra* vector $< a|$ lies in the dual space. We call $< a|$ the *adjoint* of $|a>$ which corresponds to the complex conjugate in the inner-product space. Thus, the inner product can be represented as $< x_i|x_j >=< x_j|x_i >^*$. Hence the name *bracket*.

Feynman Path Integral (FPI): FPI relates the entire path/trajectory of a given stochastic process to the transition density. As intuitively discussed in main body of the chapter, we associate a Lagrangian to a given stochastic process x_t. This Lagrangian is a function of the position x_t and time derivative $\dot{x}_t$ (i.e. the velocity) of the process, describing the dynamics/trajectory of the process. The process can take all possible paths between the initial position x_0 and the final position x_t at time t, and each path contributes to the total probability. The Lagrangian implicitly associates certain probability to each possible path. Then time integral of the Lagrangian function is evaluated, i.e. the integral over time of the trajectory of particle/process, which is called the *action* functional, such that it is an extremum path:

$$S[x_t] = \int_{t_i}^{t_f} L(\dot{x}, x, s) ds$$

where $L(\dot{x}, x, s)$ is the Lagrangian of the process and the square bracket emphasizes that the *action* is a functional as it depends on the trajectory of the particle/process. For example, the trajectory could be described by the simple diffusion process: $dx_t = \mu(x, t)dt + \sigma(x, t)dW_t$.

Thus, FPI is defined in terms of transition probability, i.e. the amplitude to go from one point to another point where each possible path is assigned a certain probability given by:

$$K(x_i, t_i; x_f, t_f) = \int_{x(0)=x_i}^{x(t)=x_f} \mathcal{D}x(t) e^{-S[x_t]}$$

where $\mathcal{D}x(t)$ is the notation that denotes integral over all paths (implicitly embedding appropriate normalization constant C). The exponential of the negative of *action* gives weight to each path.

As originally described by Feynman (1948), the path integral can be defined as a limit of the sequence of finite-dimensional multiple integrals (analogous to Riemann integral being defined as the limit of the sequence of finite sums):

$$K(x_i, t_i; x_f, t_f) = \underbrace{\int_{x(0)=x_i}^{x(t)=x_f} \mathcal{D}x(t)}_{\text{sum over all paths}} \qquad \underbrace{e^{-S[x_t]}}_{\text{amplitude/weight for each path}}$$

where the notation is

$$\int_{x(0)=x_i}^{x(t)=x_f} \mathcal{D}x(t) = \lim_{n\to\infty} C_n \int dx_1 \int dx_2 ... \int d_{n-1} = \lim_{n\to\infty} C_n \prod_{n=1}^{n-1} \int dx_n$$

Since for calculations we need to use limiting procedure, we can write $K(.)$ as:

$$K(x_i, t_i; x_f, t_f) = \lim_{n\to\infty} C_n \int dx_1 \int dx_2 ... \int d_{n-1} \quad e^{-S(x_j, t_j; x_{j-1}, t_{j-1})}$$

where

$$S(x_j, t_j; x_{j-1}, t_{j-1}) = \Delta t \sum_{i=0}^{n-1} L((x_j - x_{j-1})/\Delta t, x_{j-1}, t_{j-1})$$

where time is discretized as $t_i = t_0 < t_1 < ... < t_n = t_f$, $\Delta t = (t_f - t_i)/n$, $t_j = t_{j-1} + \Delta t$, C_n is a normalization constant which depends on the elapsed time $t_f - t_i$. For further details, see the references given above, particularly the original paper by Feynman (1948) and the book by Feynman and Hibbs (1965).

For example, if the process is driven by zero-drift Brownian motion, say $x_t = \sigma dW_t$, then normalization constant is $C_n = \frac{1}{(\sqrt{2\pi\sigma^2 \Delta t})^n}$. The Lagrangian for such a process is well-known as $L(\dot{x}, x, s) = \frac{1}{2\sigma^2}\dot{x}_s^2$. In physics, it describes the evolution of a *free particle*. In probability theory, this describes the diffusion dynamics.

Evaluating the action of this Lagrangian, we get $S[x_t] = \frac{1}{2(t_f - t_i)}(x_f - x_i)^2$. Taking the limit as $n \to \infty$, we get the familiar transition probability of the process driven by standard Brownian motion:

$$K(x_i, t_i; x_f, t_f) = \frac{1}{\sqrt{2\pi\sigma^2(t_f - t_i)}} e^{-\frac{(x_f - x_i)^2}{2\sigma^2(t_f - t_i)}}$$

which is, as expected, a normal density. Note that this is also the fundamental solution of the zero drift diffusion equation $dx_t = \sigma dW_t$, whose density evolution can be characterized by the Kolmogorov forward equation:

$$\frac{\partial K}{\partial t} = \frac{1}{2}\sigma^2 \frac{\partial^2 K}{\partial x^2}$$

This is also the well known one-dimensional heat-equation in physics, where $K(.)$ describes the distribution of temperature.

Now consider the usual diffusion process: $dx_t = \mu dt + \sigma dW_t$. The Lagrangian associated with this stochastic process is known in the literature by: $L(\dot{x}, x, s) = \frac{1}{2\sigma^2}(\dot{x}_s - \mu)^2$.

For example, setting $\mu = 0$, we get a zero-drift Brownian motion $dx_t = \sigma dW_t$, and the Lagrangian as $L(\dot{x}, x, s) = \frac{1}{2\sigma^2}\dot{x}_s^2$, which corresponds to the normal density as discussed above.

See Goovaerts, Schepper and Decamps (2004) for the FPI and transition density expressions associated with various classical financial models, such as Geometric Brownian motion, Vasicek model, Cox-Ingersoll-Ross model etc. Note that the Lagrangian constructed and related examples in this appendix are for reduced form equations/dynamics. When we impose the equilibrium conditions such that either drift or diffusion coefficients (or both) are determined endogenously (and potentially non-linear, as is the drift coefficient in this chapter), it might not be possible to get closed form or even simple analytical expressions of the associated density function. Thus, extra care needs to be taken in defining and deriving the equilibrium process in order to get the most our of FPI formalisms. In this chapter, we focus on the dynamics of the debt/leverage and financial wealth, a_t^b and a_t^l respectively. These are the key equilibrium/endogenous objects whose evolution encodes a lot of information. This simple observation and adaptation has allowed us to efficiently use the FPI formalism, as clearly evident from the simplicity of the proofs of various results.

Another property of the FPI is that it satisfies the *Chapman-Kolmogorov equation* for Markov processes, making it extremely useful in economic and financial applications:

$$K(x_i, t_i; x_f, t_f) = \int_{\mathbb{R}} K(t_s, x_s; t_f, x_f) K(t_i, x_i; t_s, x_s) dx_s$$

Thus, FPI can be applied to a wide variety of economic settings, helping answer many interesting questions and explore different/new economic phenomenon, especially those relating to the transition dynamics and evolution of the entire system. This paper represents one such application of these powerful methods.

A.6 Additional Details on Financial Innovation

In this appendix we provide some additional details on financial innovation. Financial innovation is an important function of the financial sector and it can also be an important source of systemic risk. Financial innovations include various financial products like Equity Options, Credit Default Swaps (CDS), Mortgage Backed Assets (MBS), Collateralized Debt Obligations (CDOs), Interest Rate Swaps, Currency Swaps, and many more. These types of financial products have emerged mainly since the mid-to late 1980s due to our increased understanding of various complex financial derivatives/instruments and how to price them. The rigorous understanding/pricing of such products was made possible after the development of standard option pricing theory, starting from the Black-Scholes-Merton model (Black and Scholes (1973) and Merton (1973)). It was followed by the theory of the term structure of interest rates by Heath, Jarrow and Morton (1992), which forms the basis to understanding derivatives. These initial advances paved the way for a large body of research in this area, making it possible to price a variety of complex financial securities. This ultimately offered a powerful set of risk management tools to market participants. Once the problem of pricing complex securities became feasible, we saw the proliferation of a wide range of such financial innovation products from relatively simple equity options to sophisticated CDS and MBS.

In chapters 2 and 3, the model of financial sector is motivated by the CDS market structure. Notwithstanding the complexities of constructing and pricing a CDS contract in actual practice, a CDS is essentially an insurance against credit risk: a negotiated contract between buyer and seller of insurance protection where the reference entity, notional, premium (also called spread), and maturity are specified. The buyer makes regular payments to the seller who compensates the buyer if a credit event occurs. In other words, CDS is an insurance contract written on the outstanding notional value of the bond or loan on an entity[4]. In fact, market participants can buy and sell CDS against vari-

[4]Credit derivatives began trading in the early 1990s and CDS was first traded by JP Morgan in 1995 (Jarrow 2009). CDS trade in Over-the-Counter (OTC) markets and not in standardized exchanges as stocks, ETFs, and equity options do.

ous eligible reference entities in the market, effectively allowing them to bet or hedge against the solvency and insolvency of the big players in the market, like big banks and corporations. In the recent crisis, the government stepped in to stop the failure of AIG as it was a large counterparty to most of the big CDS contracts. For further illustration of the CDS market are contained in Jarrow (2011) and Stulz (2010). The basic structure of the CDS market is also visually captured in Figure (2.1). Our model essentially generalizes this structure.

CDS represents an important instance of financial innovation. CDS is thought to be directly linked to the financial/credit crisis of 2007-2009. There have been claims that trading in CDS was one of the key factors making the recent credit crisis even more severe. Thus, recent financial regulations have been sharply aimed at regulating such OTC derivatives. In chapters 2 and 3, our goal is to understand the underlying mechanisms and channels through which such innovations can lead to systemic risk implications for the entire economy. For example, innovations like CDS could potentially be welfare enhancing as economic agents have the option to optimally allocate their risks. The market is one step closer to completion. Trading of innovations like CDS can even reduce market imperfections in debt trading, thus lowering the debt costs and allowing more access to debt capital. However, default possibilities of the counterparty could potentially annul the benefits of such innovations, leading to increased systemic risk and possibly the failure of the financial markets (Jarrow, 2011). Thus, it is important to understand the structure and underlying mechanisms with respect to financial innovation if we are to understand systemic risk.

To understand the role of the financial sector in the recent credit crisis, involving CDS market, here is a paragraph from Stulz (2010):

On one side, they seem like straightforward financial derivatives that serve standard useful functions: making it easier for credit risks to be borne by those who are in the best position to bear them, enabling financial institutions to make loans they would not otherwise be able to make, and revealing useful information about credit risk in their prices. On the other side, in trying to understand the credit crisis, many observers have identified credit default swaps to be a prominent villain. One segment of the "60 Minutes" television show on October 26, 2008, called credit default swaps on subprime mortgages the "bet that blew up Wall Street."

If CDS acts like insurance, why did it become the focus of such a massive carnage on Wall-Street that ultimately affected the entire country (and the world)? This excerpt from Jarrow (2009a) should give us a clear idea about the nature of CDS as essentially being a double-edged sword. This also highlights why the model of the financial sector in first two chapters is strongly motivated by the economic structure of the CDS market:

Similar to a put option, all derivatives have characteristics of insurance contracts. For example, the credit derivatives -credit default swaps (CDS) -that were so visible during the crisis are precisely insurance contracts written to insure the value of bonds issued by various corporations and financial trusts...As insurance contracts, these derivatives enable the insured to hedge financial risks. And everyone understands the benefits of obtaining insurance..

The provider of the insurance, the insurer, is willing to sell the insurance and bear the risk of a loss because they receive a risk premium for this activity. Insurance contracts are good, and they increase society's welfare as long as the insurer can fulfill the terms of the contract if the insured event occurs. This occurs, of course, if the insurer is well capitalized and has sufficient funds set aside to cover any losses incurred by the *insured. If not, then insurance doesnt reduce risk, but it magnifies it. When the loss event occurs, both the insured and the insurer suffer losses. In this later case, insurance although good in concept, fails in practice.*

Thus, it is clear that financial innovation serves a beneficial function of providing various risk management tools, but the benefits are conditional upon the financial sector remaining solvent. Systemic risk can arise from the potential failure of the financial sector, which eliminates the first-order beneficial effects of insurance. See Stulz (2010) for an intuitive discussion of CDS and how it may have contributed to the recent credit crisis. For a more technical treatment, including the pricing of CDS and other credit derivatives, see the review by Jarrow (2011).

A.7 Empirical Observations

In this section, we will further discuss various empirical observations relevant to chapters 2 and 3. Since the late 1970s and early 1980s, the leverage ratio (debt-to-income ratio) of U.S. households increased steadily as shown in Figure (A.1). US households entered the financial crisis of 2007-2009 with high leverage ratio (data from FRED database). Mian and Sufi (2011a) find similar pattern using a slightly different data set where the corporate sector (excluding financial sector) did not seem to show any such trend increase. The size of the financial sector also grew considerably during this time from around 5% (as value added/GDP) in the 1980s to around 8% in 2007. Figure (A.2) shows the scatter plot of U.S. household debt-to-GDP ratio and GDP share of the financial sector from 1970-2009 (Philippon, 2014).

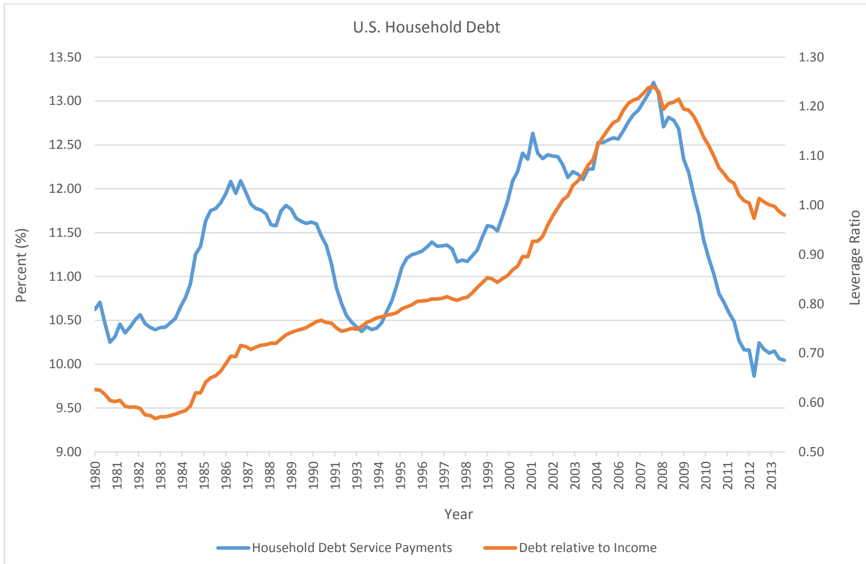


Figure A.1: Household Leverage

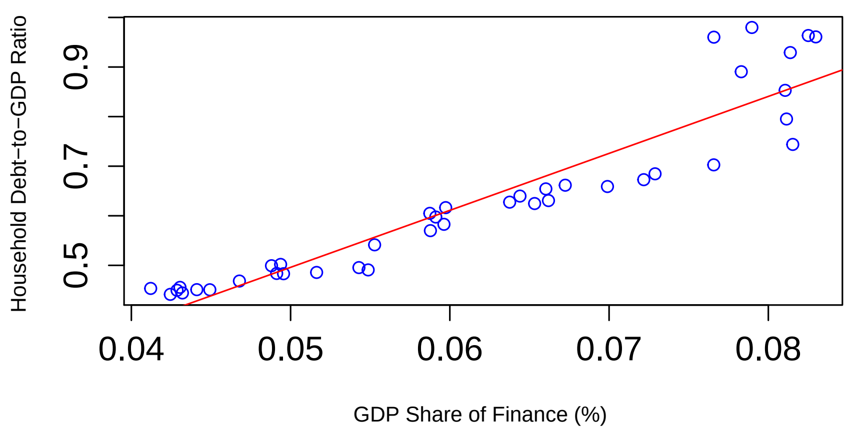


Figure A.2: Debt-to-GDP Ratio Vs. GDP Share of Financial Sector

Another important feature of the data is the household heterogeneity in marginal propensity to consume, save and borrow at or around the borrowing limit/constraint. This is especially important when a large portion of the population is facing the constraint. In the credit card utilization rates study by Gross and Souleles (2002), income is found to be highly correlated with credit utilization. Agents close to their spending limit on their credit card spend a much higher portion of their increased limit relative to agents who are far from their constraints. Agents who are at their credit limit spend about $0.45 for every $1 increase in their limit. But agents with large breathing room in their cards spend only about $0.07 for every $1 increase in their borrowing limit. This suggests that we need to look at not just the borrowing constraint but also the mass of the population that is facing the constraint because the behavior of such agents is very different than those agents far from their constraint. Heterogeneity matters a great deal in this type of situation. We will explicitly relate this empirical observation to the notion of systemic risk in our model.

Now we examine the financial derivatives markets, and specifically the CDS market for our context. The entire OTC derivatives market has exhibited exponential growth over the last decade as shown in Figures (A.3) and (A.4), going from less than $100 trillion to as high as $700 trillion in just about a decade's time. Compare that to the US GDP of about $18.5 trillion in 2016. But the CDS market feature stands out in particular. In 2001 the CDS market was just under $1 trillion. It reached a peak of around $62 trillion in 2007, right at the dawn of the credit crisis. Since then it steadily declined to less than $20 trillion in 2014. However, all other derivatives have remained fairly stable after an initial decline in the first few years of the credit crisis. Also, the size of the CDS market dwarfs that of the equity derivatives market. It is normal that the total Interest Rate and

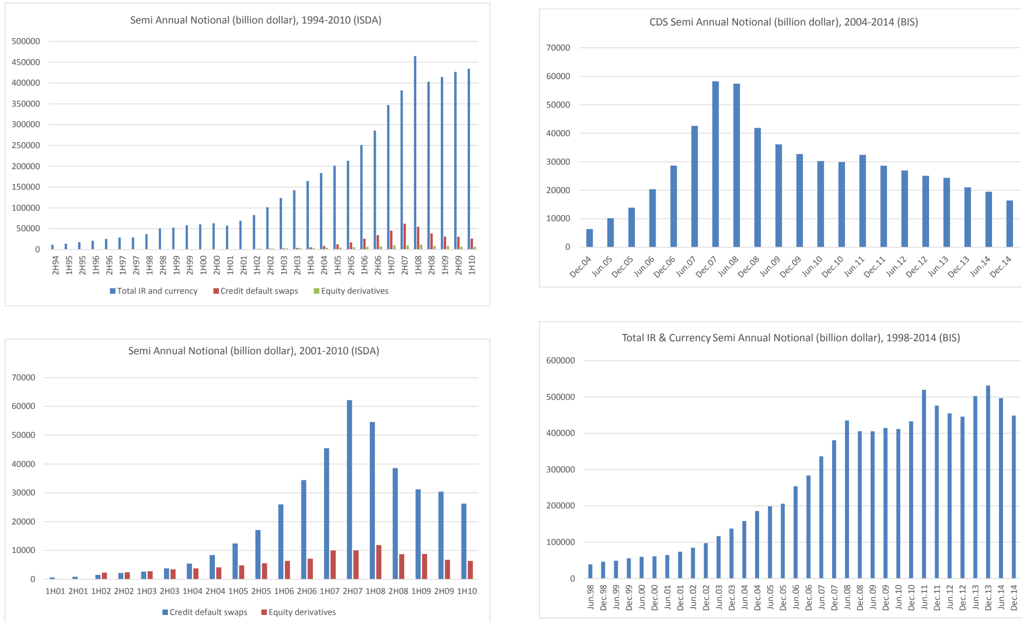


Figure A.3: Derivatives Data 1

Currency Derivatives markets dominate the CDS market because those markets are the biggest in the world, clearly overshadowing even the total value of the entire U.S. equities market.

From these observations, we can see connections between financial innovation and leverage. One could argue that one of the roles of CDS, or other similar derivatives for that matter, has been to increase the supply of credit. The existence and availability of such financial instruments, originally designed to protect against defaultable securities, might itself make market participants willing to supply even more credit than they would otherwise. This gives rise to an interesting situation where a lending party can buy sufficient protection against their existing loans with the goal to use more of their existing capital to make out newer loans. This mechanism highlights the relationship between increasing overall debt/leverage and increasing CDS market size before the crisis and

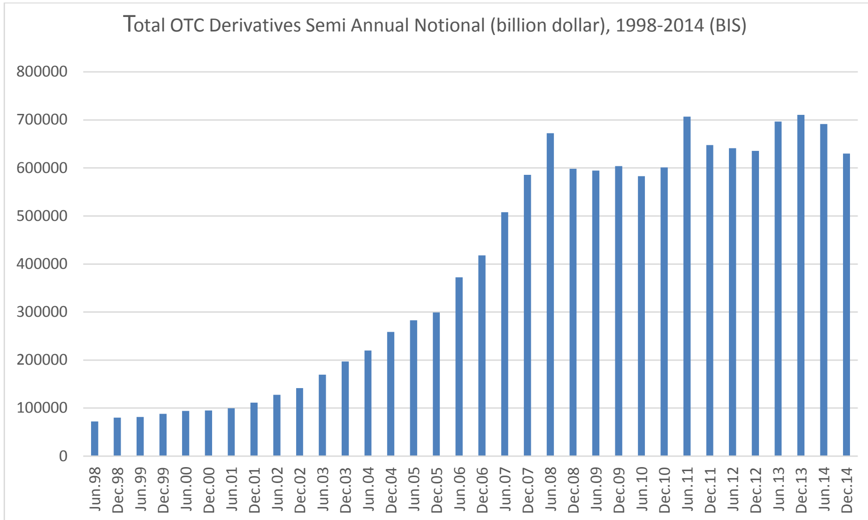


Figure A.4: Derivatives Data 2

the subsequent decline in both markets after the crisis. In chapters 2 and 3 we will analyze a variety of such macro-finance links and mechanisms through a rigorous and robust framework.

BIBLIOGRAPHY

Achdou, Y., J. Han, J. Lasry, P. Lions, and B. Moll. 2015. "Heterogeneous Agent Models in Continuous Time." Princeton University Working Papers.

Achdou, Y., F. Camilli, and I. Capuzzo-Dolcetta. 2012. "Mean field games: convergence of a finite difference method." *ArXiv e-prints*.

Acharya, V. and L. Pedersen. 2005. "Asset Pricing with Liquidity Risk. *Journal of Financial Economics*, 77: 375–410.

Acharya, V., L. Pedersen, T. Philippon and M. Richardson. 2010. "Measuring Systemic Risk." Working Paper, New York University.

Ait-Sahalia, Y. 1999. "Transition densities for interest rate and other nonlinear diffusions." *The Journal of Finance*, 54(4), 1361–1395.

Aiyagari, S.R. 1994. "Uninsured Idiosyncratic Risk and Aggregate Saving." *The Quarterly Journal of Economics*, 109(3): 659–84.

Albrecher, H., and S. Asmussen. 2010. *Ruin Probabilities*, 2nd edition, World Scientific, New Jersey.

Allen, F., and D. Gale. 1994. *Financial Innovation and Risk Sharing*. MIT Press, Cambridge, MA.

Applebaum, D. 2009. *Lévy Processes and Stochastic Calculus*, 2nd edition, Cambridge University Press.

Barles, G., and P. E. Souganidis. 1991. "Convergence of approximation schemes for fully nonlinear second order equations." *Asymptotic Analysis*, 4: 271–283.

Baaquie, B. 2004. *Quantum Finance: Path Integrals and Hamiltonians for Options and Interest Rates*. Cambridge University Press.

Benhabib, J., A. Bisin, and S. Zhu. 2011. "The Distribution of Wealth and Fiscal

Policy in Economies With Finitely Lived Agents." *Econometrica*, 79(1): 123157.

Benhabib, J., A. Bisin, and S. Zhu. 2013. "The Distribution of Wealth in the Blanchard-Yaari Model." NYU Working Papers.

Benhabib, J., A. Bisin, and S. Zhu. 2014. "The Wealth Distribution in Bewley Models with Investment Risk." NYU Working Papers.

Bernanke, B.S., and M. Gertler. 1989. "Agency Costs, Net Worth, and Business Fluctuations." *American Economic Review*, 79 (1): 14–31.

Bernanke, B.S., M. Gertler, and S. Gilchrist. 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework." *Handbook of Macroeconomics.* Vol. 1, edited by John B. Taylor and Michael Woodford, 134193. Amsterdam: Elsevier Science, North-Holland.

Bertoin, J. 1996. *Lévy Processes*, Cambridge University Press.

Bewley, T.. 1986. "Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers." *Contributions to Mathematical Economics in Honor of Gerard Debreu*, edited by Werner Hildenbrand and Andreu Mas-Collel. Amsterdam: North-Holland.

Billio, M., M. Getmansky, A.W. Lo, and L. Pelizzon. 2010. "Measuring Systemic Risk in the Finance and Insurance Sectors." MIT Sloan School Working Paper 4774.

Bisias, M. Flood, A.W. Lo, S. Valavanis. 2012. "A Survey of Systemic Risk Analytics," Office of Financial Research." Working Paper 0001, January, 2012.

Black F., and M. Scholes. 1973. "The pricing of options and corporate liabilities." *Journal of Political Economy*, 81: 637–59.

Board of Governors of the Federal Reserve System, *Financial Accounts of the United States - Z.1*, 25 Sep. 2017. URL: https://www.federalreserve.gov/releases/z1/.

Brunnermeier, M., and Nagel, S. 2004. "Hedge funds and the technology bubble.” *Journal of Finance*, 59(5): 2013–2040.

Brunnermeier, M., and L. Pedersen. 2008. “Market Liquidity and Funding Liquidity.” *Review of Financial Studies*, 22(60): 2201–2238.

Brunnermeier, M., and Y. Sannikov. 2014. “A Macroeconomic Model with a Financial Sector.” *American Economic Review*, 104(2): 379–421.

Brunnermeier, M., and M. Oehmke. 2012. “Bubbles, Financial Crises, and Systemic Risk.” NBER Working Paper No. 18398.

Brunnermeier, M.K., T.M. Eisenbach, and Y. Sannikov. 2013. “Macroeconomics with Financial Frictions: A Survey.” In *Advances in Economics and Econometrics: Tenth World Congress of the Econometric Society.* Vol. 2, edited by Daron Acemoglu, Manuel Arellano and Eddie Dekel,394. New York: Cambridge University Press.

Candler, G.V. 1999. “Finite-Difference Methods for Dynamic Programming Problems.” In *Computational Methods for the Study of Dynamic Economies*. Cambridge, England: Cambridge University Press.

Capuzzo-Dolcetta, I., and P. L. Lions. 1990. “Hamilton-Jacobi Equations with State Constraints.” *Transactions of the American Mathematical Society*, 318(2): 643+.

Census Bureau, U.S. , *New Residential Sales*, 22 Sep. 2017. URL: https://www.census.gov/construction/nrs/index.html.

Cetin, U., R. Jarrow, and P. Protter. 2004, “Liquidity Risk and Arbitrage Pricing Theory.” *Finance and Stochastics*, 8: 311–341.

Cetin, U. and L. C. B. Rogers. 2007, “Modeling Liquidity Effects in Discrete Time.” *Mathematical Finance*, 17 (1): 15–29.

Chatterjee, S., D. Corbae, M. Nakajima and Rios-Rull. 2007. “A Quantitative Theory of Unsecured Consumer Credit with Risk of Default.” *Economterica*, 75(6): 1525–1589.

Chjebbi, S. and H. Soner. 2013. "Merton Problem in a Discrete Market with Frictions." *Nonlinear -Analysis: Real World Applications*, 14: 179–187.

Christiano, L., M. Eichenbaum, and C. Evans, 2005. "Nominal rigidities and the dynamic effects of a shock to monetary policy." *Journal of Political Economy*, 113(1): 1–45.

Christiano, L., M. Trabandt, and K. Walentin. 2010. "DSGE Models for Monetary Policy Analysis." NBER Working Paper No. 16074.

Clark, S., and Coggin, D. 2011. "Was there a US house price bubble? An econometric analysis using national and regional panel data," *Quarterly Review of Economics and Finance*, 51: 189–200.

Cont, R., and P. Tankov. 2004. *Financial Modelling with Jump Processes*, 1st edition, Chapman and Hall/CRC Financial Mathematics Series.

Cont, R., and E. Voltchkova. 2005. "A finite difference scheme for option pricing in jump-diffusion and exponential Lévy models." *SIAM Journal on Numerical Analysis*, 43(4):1596–1626.

Cuoco, C. and H. He. 1994. "Dynamic Equilibrium in Infinite-Dimensional Economies with Incomplete Financial Markets," working paper, Wharton.

Cuoco, C., and H. He. 2001. "Dynamic Aggregation and Computation of Equilibria in Finite-Dimensional Economies with Incomplete Financial Markets." *Annals of Economics and Finance*, 2: 265–296.

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

Dirac, P. 1930. *The Principles of Quantum Mechanics*. Oxford University Press.

Duffie, D., and M. Huang. 1996. "Swap rates and credit quality." *Journal of Finance*, 51(3): 921–49.

Duffie, D., and K. Singleton. 1999. “Modeling term structures of defaultable bonds.” *Review of Financial Studies*, 12(4): 687–720.

Eggertsson, G.B., and P. Krugman. 2012. “Debt, Deleveraging, and The Liquidity Trap: A Fisher-Minsky-Koo Approach.” *The Quarterly Journal of Economics*, 127: 1469–1513.

Feller, W. 1971. *An Introduction to Probability Theory and Its Applications*. Vol. II, Wiley, New York.

Feynman, R. 1948. “Space-Time Approach to Non-Relativistic Quantum Mechanics.” *Review of Modern Physics*, 20(2): 367–387.

Feynman, R. and A. Hibbs. 1965. *Quantum Mechanics and Path Integrals*. McGraw-Hill.

Gabaix, X., J.M. Lasry, P.L. Lions, and B. Moll. 2016. “The Dynamics of Inequality.” *Econometrica*, 84(6): 2071–2111.

Garber, P. 1989. “Tulipmania.” *Journal of Political Economy*, 97(3): 535–60.

Garber, P. 1990. “Famous first bubbles.” *Journal of Economic Perspective*, 4(2): 35–54.

Gardiner, C.W.. 1985. *Handbook of Stochastic Methods*, 2nd Edition, Springer-Verlag.

Gennaioli, N., A. Shleifer, and R.W. Vishny. 2012. “Neglected Risks, Financial Innovation and Financial Fragility.” *Journal of Financial Economics*, 104: 452–468.

Gross, D.B., and N.S. Souleles. 2002. “Do Liquidity Constraints and Interest Rates Matter For Consumer Behavior? Evidence from Credit Card Data.” *The Quarterly Journal of Economics*, 149–185.

Goovaerts, M., A.D. Schepper, M. Decamps. 2004. “Closed-form approximations for diffusion densities: a path integral approach.” *Journal of Computational and Applied Mathematics*, 164–165: 337–364.

Griffiths, D. 2004. *Introduction to Quantum Mechanics*, 2nd edition, Pearson Prentice Hall.

Guerrieri, L. and M. Iacoviello. 2017. "Collateral constraints and macroeconomic asymmetries." *Journal of Monetary Economics*, 90: 28–49.

Guvenen, F. 2011. "Macroeconomics With Heterogeneity: A Practical Guide." NBER Working Papers 17622.

Hall, R. 2011. "The Long Slump." *American Economic Review*, 101: 431–469.

Hand, L.N, and J.D. Finch. 1998. *Analytical Mechanics*. Cambridge University Press.

Harrison, M., and D. Kreps. 1978. "Speculative investor behavior in a stock-market with heterogeneous expectations." *Quarterly Journal of Economics*, 92, 323–336.

Hayashi, F. 1982, "Tobin's marginal q and average q: A neoclassical interpretation." *Econometrica*, 50(1): 213–224.

He, Z., and A. Krishnamurthy. 2012. "A Model of Capital and Crises." *Review of Economic Studies*, 79(2): 735–77.

He, Z., and A. Krishnamurthy. 2013. "Intermediary Asset Pricing." *American Economic Review*, 103(2): 732–70.

He, Z., and A. Krishnamurthy. 2014. "A Macroeconomic Framework for Quantifying Systemic Risk." NBER Working Papers 19885.

Heath D., R. Jarrow, and A. Morton. 1992. "Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation." *Econometrica*, 60(1): 77–105.

Heathcote, J., K.Storesletten, and G.L. Violante. 2009. "Quantitative Macroeconomics with Heterogeneous Households." *Annual Review of Economics*, 1(1): 319–354.

Holmes, R. 1970. *A Course on Optimization and Best Approximation*. Lecture Notes in Mathematics 257, Springer.

Hopenhayn, H.A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica*, 60(5): 1127–50.

Huggett, M. 1993. "The risk-free rate in heterogeneous-agent incomplete-insurance economies." *Journal of Economic Dynamics and Control*, 17(5-6): 953–969.

Jarrow, R. 2015. "Asset Price Bubbles." *Annual Review of Financial Economics*, 7: 201–18.

Jarrow, R. 2009a. "An expensive Education." *Canadian Investment Review*, Winter 2009: 9–15.

Jarrow, R. 2009b. "Credit Risk Models." *Annual Review of Financial Economics*, 1: 37–68.

Jarrow, R. 2011. "The Economics of Credit Default Swap." *Annual Review of Financial Economics*, 3: 235–257.

Jarrow, R. 2016. "On the Existence of Competitive Equilibrium in Frictionless and Incomplete Stochastic Asset Markets." *Mathematics and Financial Economics*, 11(4): 455–477.

Jarrow, R. 2016. "Asset Market Equilibrium with Liquidity Risk." forthcoming, *Annals of Finance*.

Jarrow, R. 2017a. *Continuous Time Asset Pricing Theory: A Martingale Based Approach*. Manuscript, Cornell University.

Jarrow, R. 2017b. "Capital Asset Market Equilibrium with Liquidity Risk, Trading Constraints, and Asset Price Bubbles." Working Paper, Cornell University.

Jarrow, R., and P. Protter. 2008. "Liquidity Risk and Option Pricing Theory." *Handbooks in OR&MS, vol. 15*, eds J. R. Birge and V. Linetsky, Elsevier B. V.

Jarrow, R., and S. Turnbull. 1995. "Pricing derivatives on financial securities subject to credit risk." *Journal of Finance*, 50(1): 53–85.

Jarrow, R., and S. Turnbull. 1997. "When swaps are dropped." *Risk Magazine*, 10(5): 70–75.

Jarrow, R., and Fan, Y. 2001. "Counterparty Risk and the Pricing of Defaultable Securities." *Journal of Finance*, 56(5): 1765–1799.

Jarrow, R., D. Lando, and F. Yu. 2005. "Default risk and Diversification: Theory and Empirical Implications." *Mathematical Finance*, 15(1): 1–26.

Jarrow, R., and S. Lamichhane. 2018. "Asset Price Bubbles, Market Liquidity and Systemic Risk." Working Paper, Cornell University.

Jermann, U., and V. Quadrini. 2012. "Macroeconomic Effects of Financial Shocks." *American Economic Review*, 102(1): 238–271.

Kaplan, G., B. Moll, and G. Violante. 2016. "Monetary Policy According to HANK." Princeton University Working Paper.

Kashakari, N. 2017. "Monetary Policy and Bubbles." URL: https://www.minneapolisfed.org/ and-events/messages/monetary-policy-and-bubbles.

Krusell, P., and A. Smith. 1998. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, 106 (5): 867–896.

Kiyotaki, N., and J. Moore. 1997. "Credit Cycles." *Journal of Political Economy*, 105 (2): 211–48.

Kreyszig, E. 1978. *Introductory Functional Analysis with Applications*. John Wiley & Sons. Inc.

Kyprianou, A.E. 2013. *GerberShiu Risk Theory*. Springer.

Kyprianou, A. 2014. *Fluctuations of Lévy Processes with Applications: Introductory Lectures*. 2nd edition, Springer.

Lando, D. 1998. "On Cox processes and credit risky securities." *Review of Deriva-*

tives Research, 2: 99–120.

Lasry, J.M, and P.L. Lions. 2007. "Mean field games." *Japanese Journal of Mathematics*, 2: 229–260.

Linetsky, V. 1998. "The Path Integral Approach to Financial Modeling and Options Pricing." *Computational Economics*, 11: 129–163.

Lamichhane, S. 2018. "Systemic Risk and Financial Innovation: A Macro-Finance Perspective." Working Paper, Cornell University.

Loeffen, R., and P. Patie. 2010. "Absolute Ruin in the OrnsteinUhlenbeck type Risk Model." Working Paper, Cornell University.

Madan, D. and H. Unal. 1998. "Pricing the risks of default." *Review of Derivatives Research*, 2: 121–160.

Merton, R. 1970. "A dynamic general equilibrium model of the asset market and its application to the pricing of the capital structure of the firm." Working Paper, MIT.

Merton, R.. 1973. "The theory of rational option pricing." *Bell Journal of Economic Management Science*, 4: 141–83.

Merton, R. 1974. "On the pricing of corporate debt: the risk structure of interest rates." *Journal of Finance*, 29: 449–70.

Mian, A., and A. Sufi. 2011a."House Prices, Home Equity-Based Borrowing, and the U.S.Household Leverage Crisis." *American Economic Review*, 101: 2132–2156.

Mian, A., and A. Sufi. 2011b."What Explains High Unemployment? The Aggregate Demand Channel." University of Chicago, Working Papers.

Minsky, H.P. 1986. *Stabilizing an Unstable Economy*. New Haven: Yale University Press.

Nikolau, K. 2009. "Liquidity (Risk) Concepts: Definitions and Interactions."

Working Paper No. 1008, European Central Bank.

Patie, P. 2005. "On a martingale associated to generalized Ornstein-Uhlenbeck processes and an application to finance." *Stochastic Processes and their Applications*, 115 (4): 593–607.

Patie, P. 2012. "Law of the Absorption Time of Some Positive Self-Similar Markov Processes." *The Annals of Probability*, 40 (2): 765–787.

Paulsen, J. 2008. "Ruin models with investment income." *Probability Surveys*, 5: 416–434.

Pennanen, T. 2014, "Optimal Investment and Contingent Claim Valuation in Illiquid Markets," *Finance and Stochastics*, 18: 733–754.

Philippon, T.. 2014. "Has the U.S. Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation." NYU Working Papers.

Protter, P. 2005. *Stochastic Integration and Differential Equations*. Second Edition, Version 2.1, Springer.

Radner, R. 1982. "Equilibrium under Uncertainty." *Econometrica* 36(1): 31–58.

Rockafeller, R.T. 1970. *Convex Analysis*. Princeton U. Press.

Ross, S. 1976. "Options and efficiency." *Quarterly Journal of Economics*, 90(1): 75–89.

Ruszczynski, A. 2006. *Nonlinear Optimization*. Princeton University Press.

Sakurai, J. 1994. *Modern Quantum Mechanics*, revised edition, Addison Wesley.

Schechter, M. 2002. *Principles of functional analysis*. 2nd edition, American Mathematical Society.

Shankar, R. 2011. *Principles of Quantum Mechanics*. 2nd edition, Plenum Press.

Soner, H.M. 1986a. "Optimal Control with State-Space Constraint I." *SIAM Journal on Control and Optimization*, 24(3): 552–561.

Soner, H.M. 1986a. "Optimal Control with State-Space Constraint II." *SIAM Journal on Control and Optimization*, 24(6): 1110–1122.

Stiglitz, J.E. 2012. "Macroeconomic Fluctuations, Inequality,and Human Development." *Journal of Human Development and Capabilities*, 13(1): 31–58.

Tobin, J. 1969. "A general equilibrium approach to monetary theory." *Journal of Money, Credit and Banking*, 1(1): 15–29.

Tuy, H. 1998. *Convex Analysis and Global Optimization*. Kluwer Academic Publishers, Norwell MA.

Vath, V., M. Mnif, and H. Pham. 2007. "A Model of Optimal Portfolio Selection under Liquidity Risk and Price Impact." *Finance and Stochastics*, 11: 51–90.

White, E. 1990. "The stock market boom and crash of 1929 revisited." *Journal of Economic Perspective*, 4 (2): 67–83.

Xiong, W. 2013. "Bubbles, Crises, and Heterogeneous Beliefs." NBER Working Paper No. 18905.

Xing, X., W. Zhang, Y. Wang. 2009. "The Stationary Distributions of Two Classes of Reflected Ornestein-Uhlenbeck Processes." *Journal of Applied Probability*, 46(3): 709–720.

Zhu, S. 2010. "Wealth distribution under idiosyncratic investment risk." NYU Working Papers.