MODELS FOR SYSTEMIC RISK

MARKET IMPACT, GAME THEORETIC AND RANDOM FINANCIAL NETWORK MODELS OF SYSTEMIC RISK

By

QUENTIN SHAO

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (2017) McMaster University
(Financial Mathematics) Hamilton Ontario

TITLE: MARKET IMPACT, GAME THEORETIC AND RANDOM FINANCIAL NETWORK MODELS OF SYSTEMIC RISK

AUTHOR: Quentin Shao

SUPERVISOR: Dr. Thomas Hurd

NUMBER OF PAGES: xiv, 141

Abstract

Systemic risk is the risk that an economic shock may result in the breakdown of the fundamental functions of the financial system. It can involve multiple vectors of infection such as chains of losses or consecutive failures of financial institutions that may ultimately cause the failure of the financial system to provide liquidity, stable prices, and to perform economic activities. This thesis develops methods to quantify systemic risk, its effect on the financial system and perhaps more importantly, to determine its cause.

In the first chapter, we provide an overview and a literature review of the topics covered in this thesis. First, we present a literature review on network-based models of systemic risk. Finally we end the first chapter with a review on market impact models.

In the second chapter, we consider one unregulated financial institution with constant absolute risk aversion investment risk preferences that optimizes its strategies in a multi asset market impact model with temporary and permanent impact. We prove the existence and derive explicitly the optimal trading strategies. Furthermore, we conduct numerical exploration on the sensitivity of the optimal trading curve. This chapter sets the foundation for further research into multi-agent models and systemic risk models with optimal behaviours.

In the third chapter, we extend the market impact models to the multi-agent setting. The agents follow a game theoretic strategy that is constrained by the regulations imposed. Furthermore, the agents must liquidate themselves if they become insolvent or unable to meet the regulations imposed on them. This paper provides a bridge between market impact models and network models of systemic risk.

In chapter four, we introduce a financial network model that combines the default and liquidity stress mechanisms into a "double cascade mapping". Unlike simpler models, this model can quantify how illiquidity or default of one bank influences the overall level of liquidity stress and default in the system. We derive large-network asymptotic cascade mapping formulas that can be used for efficient network computations of the double cascade. Finally we use systemic risk measures to compare the results of including with and without an asset firesale mechanism.

Acknowledgements

Professor Thomas R. Hurd has been and always will be a great mentor, friend, and even an accomplice for giving me the opportunity to pursue a PhD. He has been instrumental in giving me guidance whenever I was lost, while giving me the freedom to express and follow my own interests through out my studies. The last five years would not have been possible without his constant encouragements and words of wisdom.

PhiMAC would not have been a family without mentioning the other professors there. I would like to thank Professor Matheus Grasselli for selling financial math as a subject worth pursuing to me when I was in undergrad, and changing my life henceforth. Professor David Lozinski for first teaching me to pursue my interests outside of lectures, perhaps giving me a first look at doing research, and always being open to discuss whatever trouble I may have caused. Professor Traian Pirvu for being there for weekly discussions to playing badminton, you made sure I stayed healthy. Professor Petar Jevtic, I would like to thank you for being so frank and realistic with me, and always willing to discuss any issues I may have. In addition, Professor Nicholas Kevlahan, for the advice you provide and always being on top of the paperwork and procedures that must be done in order for me to complete my PhD, especially in an environment and time when procedures are added and changed constantly.

To the many talented individuals who've helped bring forth the papers included in this thesis, my co-authors Davide Cellai and Sergey Melnik, thank you for your patience and willingness to include a fresh graduate student as part of your project. I am thankful to Grzegorz Halaj of the European Central Bank, for discussions about the EU financial network.

To my friends and classmates; the M-Phimac students of 2011-12, Janice He, Jerry Zhang, Subashish Sengupta, Lei Wang, Hui Li, and the many others, thank you for including me into your lives, and helping me get through the busiest and most demanding time of grad school. To the other graduate students I met along the way, Dr. Sheng Tu, and Dr. James Jiang, thank you for becoming a second family to me, I can never repay you guys for the kindness you have shown. Bernardo Costa Lima, Huibing Chen and Farzad Pourbabaee, thank you for helping me throughout my research. Tuan Quoc Tran, collaborating with you has taught me so much, and the passion and dedication you have for mathematics is truly inspiring.

Finally, to my parents, thank you for raising me, and supporting me to pursue

my own interests, I can never thank you enough for all you've done.

Contents

List of Tables

List of Figures

Declaration of Academic Achievement

This "sandwich thesis " consists of three self-contained and related financial mathematics papers completed during the author's Ph.D. study from 2013 to 2017. As per the requirements of the School of Graduate Studies, the following lists the work and contribution towards each paper by the author and co-authors.

The first paper, "Optimal Portfolios of Illiquid Assets" is a joint work with Dr. Tuan Tran and Prof. Thomas Hurd. This paper builds upon the first generation market impact models as developed in (Almgren and Chriss, 2001), with linear-quadratic market impact factors. The major contributions of this paper include extending the Merton utility optimization problem to the market impact framework, analytical solutions for trading strategies, and time-consistent trading strategies. Contributions by the author include: 1. Data collection, computer programming, and numerical analysis; 2. Equal share of drafting and majority of finalization of the paper; 3. Equal share of the technical derivations of this paper, and numerous studies into alternative solutions. As of Oct 31, 2016, this paper is under review for publication by SIAM Journal of Financial Mathematics.

The second paper, "Asset Fire Sales and Systemic Risk" is a joint work with Dr. Tuan Tran and Prof. Thomas Hurd. This paper builds upon our previous work on Market Impact models and extends it to a multi-agent setting. The aim of this paper is to measure systemic risk while allowing for agent-based behavioural optimization. One further impact of this paper is the analysis of the regulations constraining the agents. Contributions by the author include: 1. computer programming, numerical analysis, and finalization of the paper; 2. Equal share of the original idea of this paper which evolved from our discussions; 3. Equal share of the technical derivations of this paper; 4. Equal share of drafting. As of October 31, 2016, this paper is currently being finalized for journal submission.

The third paper, "Double Cascade Model of Financial Crises" has been published in International Journal of Theoretical and Applied Finance and is a joint work with Prof. Thomas Hurd, Dr. Davide Cellai, and Dr. Sergey Melnik. This paper is a contribution to the field of network models of systemic risk. The impact of this paper is to provide a fast and accurate method of approximating a default contagion as well as a funding liquidity contagion (stress) spreading through a network using analytical results built on infinite sized graphs. The contributions of the author include: 1. Computer programming, numerical analysis. 2. Equal share of technical derivation of this paper, equal share of finalization of this paper, equal share of drafting. 3. In addition to the original paper, the author has added the firesale effect to the cascade mechanisms and studied the systemic risk measure and their implications for the European banking network-based on 2011 data.

Chapter 1

Market Impact and Systemic Risk

The focus and glue that brings all the portions of this thesis together is systemic risk. The definition of systemic risk, much like the risk itself, is not easily determined or agreed upon. Schwarcz (2008) narrows down systemic risk to the risk coming from a trigger such as an economic shock leading to a domino effect such as a chain of losses or chain of defaulted financial institutions. On the opposite side, Haldane (2010) elusively claims systemic risk is a cost to society like emissions from the auto-industry. Regardless of the definition of the risk, what is agreed upon is that systemic risk is catastrophic and is inherent in the financial system. In this thesis, we will see that systemic risk can be created unknowingly by individual institutions and is generated within the fabric which connects financial institutions.

This thesis aims to develop and intertwine two strands of research that explore aspects of systemic risk. The first strand is market impact models for the dynamics of asset prices under realistic market conditions. The idea is to go beyond the traditional assumption of a frictionless market and perfect competition, and to study how they impact market prices purely through actions of individual agents. In an economic sense, if the agents were contained within a financial system, their trades will impact the real economy in a measurable way.

The second strand is random network cascade models. Random networks model networks with low amounts of information, or simply networks that change so fast that exact details cannot be captured or are no longer useful. In recent years, this has become a

popular approach to describe and simulate systemic risk on a banking network via cascade models. This is a useful approach to understanding and measuring systemic risk, but what it lacks is details on the mechanisms which formed the network and methods that link the financial network modelled to the real economy.

1.1 Understanding and Modelling the Financial Crisis

The financial crisis of 2007-08 changed for the worse the perception of the general public and many governments on the effects of the financial sector as a whole (Turner (2010)). Although called a financial crisis, it was very much a contagion that spread across the globe and affected all sectors of the economy. The financial sector is an indispensable part of the economic engine due to its function as financial intermediaries that provide liquidity to the rest of the economy (Harding and Atkins (2014)). This function inherently carries risk (Zazzaro (2002)), and as a result policy makers are faced with understanding and controlling the risk brought on by the financial system. Furthermore, policy makers must try to prevent a crisis from happening in the financial system and try to contain a crisis if it does occur. This problem is further compounded by the heavy lobbying and the lack of consensus among the economists (U.S. House of Representatives (2013)).

In the financial system, contagion is the idea that weak or unstable banks can turn otherwise healthy banks unstable. Hurd (2016) lists the channels for contagion to spread within the financial system as default contagion, market illiquidity and firesales, correlation, and funding liquidity contagion. Cont et al. (2010) sought to understand the threshold above which contagion will take hold in the financial system. Gai and Kapadia (2010a) explored the extent of a default contagion spreading through financial networks and Gai et al. (2011) studied liquidity contagion that is mathematically its mirror image. Market illiquidity and firesales have been studied in the paper of Cifuentes et al. (2005), and even more explicitly in market impact models of Almgren and Chriss (2001).

This concept of contagion, which is borrowed from biological systems, appears in various schools of economic thought. In monetary economic models with endogenous money, for any growth to occur, at least one agent must be incurring additional debt (Wray (2001)). This setup, when combined with real world data from the European Central Bank (ECB), US Department of Commerce, and Office of Statistics, UK, showed that corporate savings increased massively post crisis. This provided direct evidence of liquidity contagion and lack of growth in all sectors. In addition, data showed that household debt levels stayed constant across the board, while various governments around the world continued to implement austerity. Under Neo-Chartalist monetary theory (Modern Monetary Theory), this situation allows for no growth (Mitchell, 2001), and under monetary circuit theory,

financial institutions must pick up additional debt in order to maintain the economy at a certain level (Godley and Lavoie, 2012). This further intertwining of the financial system and sectors of the real economy brings additional dangers and complexity.

More recently, many central banks such as the Bank of Canada and the Bank of England, started using models such as the Macro-Financial Risk Assessment Framework (MFRAF-Anand (2013)). These are multi-period macroeconomic models with a financial network cascade component inserted inside. Then the results of the macroeconomic model are fed into the financial network component. In each period, the result of the financial network cascade component is used for the next cycle of the macroeconomic model simulation. Such models give a more precise understanding of the impact of the financial system onto the real economy, and even more importantly show the feedback between the real economy and the financial system. Such models also promote the idea that the crisis is not a single crisis, but is rather a combined cycle of impacts of the real economy on the financial system and the impact of the financial system on the real economy.

Financial network models are a valuable method of studying potential future financial crises. From an economic perspective, they allow for the monitoring of the flow of funds within the network, while also monitoring the status of the financial institutions through their balance sheets. In addition, financial network models can provide a damage estimate in monetary terms of a crisis in the system. One method of dealing with the lack of precise and timely information of the nature of the financial network is to use random graphs as a foundation for the study of the general properties of a financial network. Cascade mechanisms are used to simulate simple deterministic behaviours such as those enforced by bankruptcy rules and regulations. We shall investigate cascade mechanisms in greater detail in the following sections.

1.1.1 Cascade Mechanisms

Cascade mechanisms are hard-coded deterministic behaviours that are simulated during a crisis scenario within a network model of financial systems. They describe explicitly the process of monitoring and changing the statuses of financial institutions (FI), commonly referred to in graph theory literature as *nodes*, based on the conditions of their immediate neighbours.

Definition 1. *For any $N \geq 1$, the collection of directed graphs on N nodes is denoted $\mathscr{G}(N)$. We consider that the set of nodes $\mathscr{N}$ is numbered by integers, i.e. $\mathscr{N} = \{1, ..., N\} := [N]$. Then $g \in \mathscr{G}(N)$, is a graph on N nodes and is a pair $(\mathscr{N}, \mathscr{E})$, where $\mathscr{E}$ is the set of directed edges and $\mathscr{E} \subset \mathscr{N} \times \mathscr{N}$. Each element $l \in \mathscr{E}$ is an ordered pair $l = (v, w)$ called an edge or link.*

In particular, the cascade process generated must be feasible on any graph (directed or undirected), which simply means that it must work on any arbitrary graph. It should also be noted that while the cascade process can be applied to any graph, the result of the cascade process is dependent on the random graph sampled. We define a random graph as the following,

Definition 2. *A random graph of size N is a probability distribution $\mathbb{P}$ on the finite set $\mathscr{G}(N)$. $\mathbb{P}$ is invariant under permutations of the N node labels.*

The complexity of the analytical solutions that describe the end results of the cascade processes naturally depends on the types of the random graph families that sample random graphs are chosen from.

Many of these cascade mechanisms are forced upon financial institutions by the regulators and are triggered when specific conditions are met. For these cases, each end result of the corresponding cascade process can be an assessment of the regulations that the cascade mechanism is based upon. On the other hand, actions of self-interest and self-preservation can also be modelled by cascade mechanisms, and provide insight into the effects created in the network by an individual's actions. Even more complex are the interactions between the self-interest behaviours and regulation-driven behaviours: depending on the type of interactions, these may lead to intractable analytical solutions. In the following sections, I provide a brief summary of cascade mechanisms applied to financial cascade models.

Watts' Cascade Model

The Watts' cascade model, introduced in Watts (2002), is a well established starting point for introducing cascade models. The model has many practical purposes, such as modelling the spread of information like ideas, and cultural fads through the population, or even the domino effects of infrastructure failure and organizational failure. Aside from the structure provided by the random graph on which the cascade process takes place, each node is also

given a *threshold* value ϕ. This ϕ value of a node represents the proportion of neighbouring nodes with activated state 1 that are needed for its own state to become activated, that is to change from 0 to 1.

The cascade model starts with all nodes in state 0, and changes a small number of nodes to state 1 in order to propagate the cascade (usually a single node). There are a few key features of this cascade model which differentiate it from the models which follow. First, there is no additional information attached to the edges of the random graph, and there is only one additional variable ϕ attached to the nodes. Second, there are only two operating states for all nodes, 0 and 1. The cascade models that will be introduced later on in this section will delve further into these two points.

Watts (2002) also defined the terminology of "vulnerable nodes", which refers to nodes with thresholds ϕ that will be breached as soon as one neighbouring node is activated. These vulnerable nodes play an important role in determining the extent of the cascade propagation when applied to finite *tree graphs* and infinite *configuration graphs*. This is due to the inherent lack of finite loops in both of these random graph families. The mean connectivity also plays a big influence on the size of the final cascade, with the cascade size increasing monotonically for small values of connectivity reaching a "knife-edge" event at which it jumps to nearly zero. This is illustrated in Figure 1.1 where the cascade model is applied to Poisson random graphs of mean degree z, and threshold value of $\phi = 18\%$. As the connectivity increases, there exist more highly connected vulnerable nodes, while the overall proportion of vulnerable nodes decreases. Note that this model provides a very simple view of a network cascade, and cannot reflect a real world network formed by financial institutions, but even at this level of simplicity, the resulting cascades demonstrate complexity and unpredictability.

Eisenberg and Noe Cascade Model

The cascade model by Eisenberg and Noe (2001) introduced nodes as financial nodes, edges as liabilities, and more importantly was specifically designed as a financial network model. As with the Watts' model, in addition to the skeleton structure provided by the random graph, each node is also assigned a starting equity value that serves as a threshold. Going beyond the Watts' model, each edge is also assigned a positive numeric value to represent

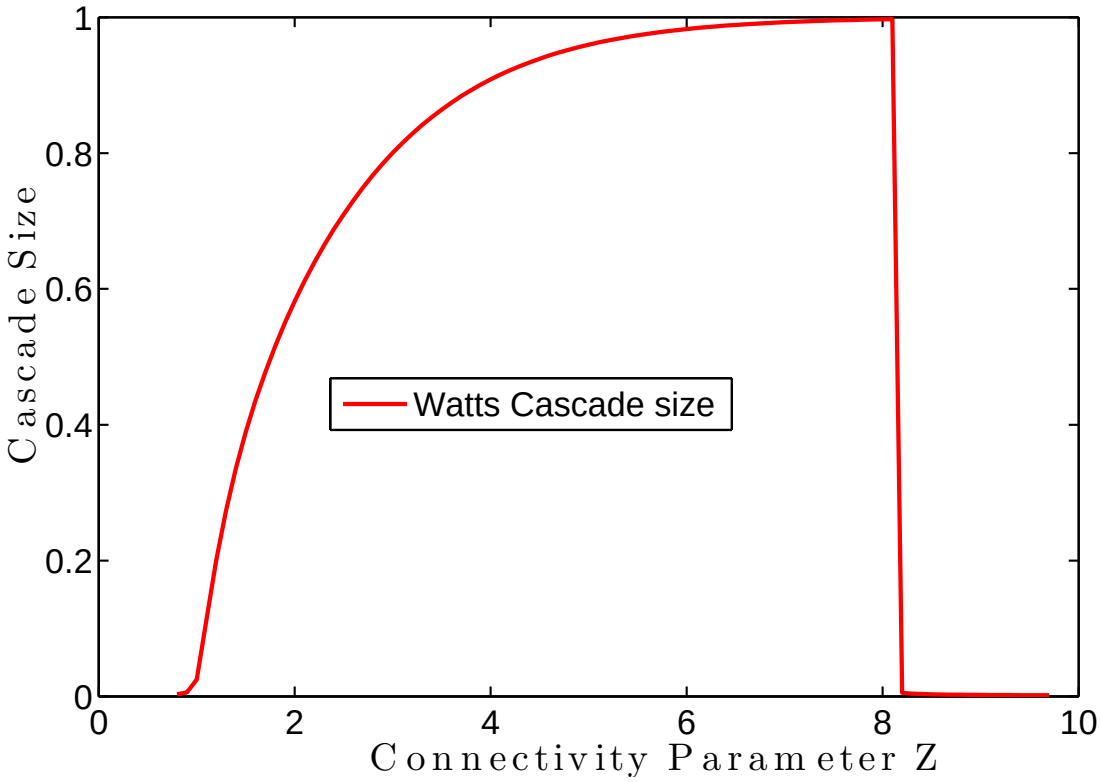


Figure 1.1: Watts' Cascade Model Mean Cascade Size: resulting mean cascade size from Watts' [2002] cascade model applied to Poisson graphs. The connectivity parameter Z refers to the expected number of edges belonging to a node picked at random. Cascade Size refers to the proportion of nodes in the random graph that are activated by the end of the cascade.

the liability between two financial nodes. It is important to note that this cascade process takes place on a directed random graph, where all edges are generated with directions. More generally, any cascade model that includes the flow of liabilities will also be taking place on a directed random graph.

The cascade mechanism for the Eisenberg and Noe (EN) cascade model is a loose representation of the laws for corporate bankruptcy. First, if a financial node cannot meet all of its obligations, it will only be required to hand over all of its current value in assets and no more. Second, obligations are either paid in full or if obligations cannot be met, then the value paid to creditors is in proportion of their nominal claim on the financial nodes assets. A financial node only has two sources of funds to repay its obligations, the first is its liquid assets that are on hand, the second is the honoured obligations from other nodes. The goal of the EN cascade model is to compute a clearing vector, which is a vector consisting of the total obligations each node will honour. This model is set up with an initial clearing vector that is less than or equal to the sum of obligations, generated by summing up the out edges of each node (i.e. one or more nodes cannot meet its obligations). This paper improves on earlier papers by proving the existence of the final clearing vector, and shows how it can be computed.

Gai and Kapadia Cascade Model (GK)

The model by Gai and Kapadia (2010a) further extends the conceptualization of nodes as financial intermediaries by explicitly defining a stylistic balance sheet for each node. Once again, on top of the structure provided by the directed random graph, each edge is given a deterministic value of 1 that explicitly represents an interbank exposure. Based on the interbank assets, a value is assigned to each node for the amount of capital it must hold as a financial intermediary. The interbank liabilities are also given since the interbank asset of one node is the interbank liability of its counterparty. Lastly, the non-interbank asset is set proportionally to the interbank assets and the non-interbank liabilities are set to equal the remaining portion of the balance sheet.

The rules for the cascade mechanism are:

1. a financial intermediary defaults once its capital reduces to zero;

2. once a financial intermediary defaults, a zero recovery rule is implemented and the corresponding asset values of its interbank neighbours are reduced to zero;
3. the capital buffers of banks are updated and checked for defaults
4. the cascade stops when no additional financial intermediaries default.

For the purposes of the Gai-Kapadia paper, all banks have identical interbank assets set to 20% of total assets, and capital requirements set to 4%. The initial shock was the removal of all of the "external" assets of a single random node. The cascade results were generated on the Poisson random graph following the examples set by Watts model. The results were nearly identical to the results of the Watts' model. The additional complexity of the GK model is hidden in the balance sheet that each node carries. Sensitivity analysis was done for the many parameters which form the balance sheet variables, to understand the impact on the balance sheet to the final cascade results. Based on the setup above, as the capital requirements are varied, the extent of the contagion, given that a contagion occurred, remained the same. However the frequency of occurrences is reduced. Similarly, when the recovery fraction is varied, the extent of a cascade remains unchanged but the frequency is reduced. This model contains many more parameters than the Watts model, but the cascade process is mathematically analogous to the Watts' models cascade process on a directed random graph.

Gai, Haldane, Kapadia Cascade Model (GHK)

Gai et al. (2011) extends the balance sheet of the previous model to better reflect realistic balance sheets of financial intermediaries. In this model, the random graph is explicitly referred to as the financial network, and the edges are now formally known as unsecured claims between financial intermediaries. The key difference between this model and the previous model is that this is a model of liquidity shock rather than default shock. Figure 1.2 shows the newly extended balance sheet of each node. The cascade mechanism is triggered for a financial intermediary if the aggregate of all its liquid assets and its collateral assets at a haircut rate is less than its interbank liabilities (i.e. there is a chance that the node will face a liquidity shortage). The financial intermediary then converts a predetermined proportion of its interbank assets into the liquid asset, reducing the interbank loan to its debtors evenly. Some of these debtors themselves become illiquid, causing further liquidity

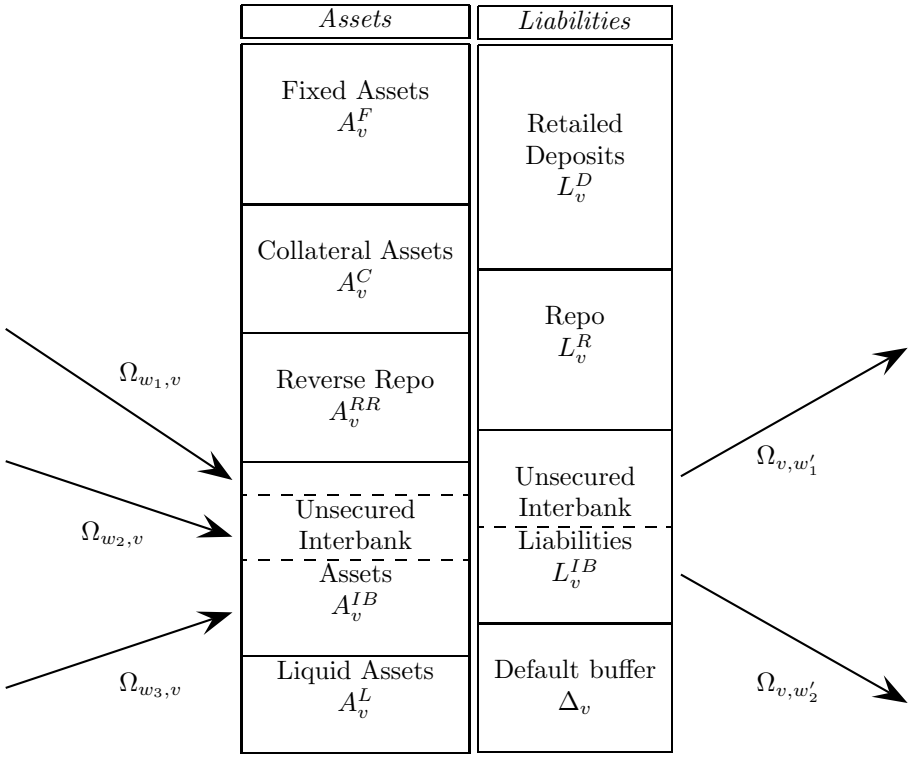


Figure 1.2: Stylized balance sheet: A symmetrical balance sheet of a financial institution on a random financial network. This balance sheet can be used for both default contagion as well as "stress" or funding liquidity contagion.

shocks. The cascade ends when no more interbank assets need to be converted. It is important to note that in this model, financial intermediaries do not default: it is a model that seeks to understand the amount of liquidity hoarding that occurs during a crisis and the corresponding reduction in connectivity of the financial network.

The extensive number of parameters in the balance sheet leads to a multitude of sensitivity analyses on stylized random graphs. For example, raising the liquid asset requirement in this model has an analogous effect to raising the capital requirement of the previous model. The paper also demonstrates that by selectively raising the liquid asset requirements for highly connected banks while maintaining the same average liquid asset requirement overall produces a more stable network. Finally, changing the proportion of interbank assets converted to liquid assets produces the same result as adjusting the recovery fraction parameter of the previous model (i.e. when a default happens the interbank asset of the creditor is lost entirely). These results are analogous to the previous models,

because mathematically the models utilize the same formal cascade process, but with the contagion shocks moving in the opposite direction.

Hurd Gleeson Extended Watts' cascade model

Hurd and Gleeson (2013) developed an extension of the Watts' model, that introduced random edge weights while preserving the cascade mechanism. More importantly, it provided an analytical solution for the cascade process when applied to configuration graphs. The cascade process in Gai and Kapadia (2010a) is mathematically similar to the Watts cascade process on a directed random graph: this allowed the analytic solutions to work for a GK model with random edge weights. Given the symmetry between the GK model and the GHK model, the same analytical solution method works for the GHK model as well. While random graphs are used for financial networks because the data for edges in the real financial network simply does not exist, similarly the data for the edge weights do not exist either. In all previous models, it had been assumed that the edge weights were equal, but there is no evidence to say this is the case at all. By including random edge weights, a more accurate picture of a financial network can be presented, that hopefully leads to more accurate predictions of the susceptibility to contagion.

1.1.2 Random Graph Theory

The foundation of financial network cascade models lies in the families of random graphs that the cascade process is applied to. The specific type of random graphs used can greatly alter the final result of the cascade process and as well as determine the computational methods available. In the following sections, we shall discuss two broad categories of random graphs in detail: configuration graphs and scale-free graphs.

Random Configuration Graphs

A *random graph* is defined as a graph selected with equal probability from the collection of graphs with N nodes and v edges (Bollobás et al. (2003)). Now this definition does not

specify any method to generate such a random graph. A more pragmatic approach is to use a subset of random graphs called random configuration graphs.

A set of *random configuration graphs* with N nodes is defined by a N-length sequence of the degree of nodes, representing the number of edges each node will contain. Each individual random configuration graph is constructed by connecting each edge uniformly until the sequence of degrees is satisfied (Janson (2009)). Unlike the financial network, where self-loops do not carry any meaning, this configuration graph construction process admits self-loops with positive probability. This problem is dealt with in Janson et al. (2014), which showed that in the limit as node numbers increase, simple configuration graphs (graphs without self-loops) have positive probability in the set of configuration multi-graphs. This theorem allows for the extension of many conclusions on configuration multi-graphs to the family of simple configuration graphs.

To form random configuration graphs of infinite size, a sequence of node degree sequences (assuming feasibility) is required. The proportion of nodes of each degree value within each finite random configuration graph must converge. This implies that many properties and consequences of infinite random configuration graph can be analyzed through the construction process of finite random configuration graphs. Hurd and Gleeson (2013) uses the fact that as the size of configuration graphs grow, the density of finite cycles reduces to zero, and created the LTI assumption. They were able to develop a method to analytically calculate cascade processes based on infinite configuration graphs with a specified degree sequence and arbitrary edge connection probability also known as *assortative graphs*. This is very important for financial networks because this allowed nodes, which represent financial institutions, to project preferences for other financial institutions they would like to enter into a counterparty relationship with.

There are two commonly used subsets of random configuration graphs: *k-regular graphs*, and *Poisson graphs*. k-regular random graphs are the simplest set of random graphs: they contain random graphs which only have nodes of degree k. These k-regular graphs are very useful at representing homogeneous networks, such as telecommunication networks or electricity grids where similar equipment are used to form the network (Motter and Lai (2002)). Poisson graphs are a subset of random configuration graphs with the degree sequence taking a Poisson distribution with parameter z, where z is also the mean degree of the graph. The benefit of Poisson graphs is that they can be constructed by

uniformly wiring N nodes until mean degree z is satisfied. In applications, Poisson graphs are used in Watts (2002); Gai and Kapadia (2010a); Gai et al. (2011), and are usually considered as the default standin for a financial network since they provide a wide range of nodes with different connectivity, representing the heterogeneity of financial institutions.

Scale-free Graphs

The family of *scale-free graphs* is a category of random graphs, where the degree distribution follows a power law (Bollobás et al. (2003)). Many studies have suggested that networks arising naturally in the real world display this property (Clauset et al. (2009)). This category of random graphs is important to financial cascade models, due to empirical evidence in Cont et al. (2010) and Boss et al. (2004) that shows the financial networks in Brazil and Austria have a power law degree distribution.

The *preferential attachment* method is often used to generate a sample of a scale-free graph. Under the preferential attachment method, the directed random graph begins as a single node, then a choice is made with probability α that a new node is created and pointing to an existing node, probability β that an edge is formed between existing nodes and γ probability of a new node being created and an edge is formed pointing from an existing node to it (in undirected graphs, α step and γ step are equivalent). The probability of forming an edge with a node v is weighted by:

$$\text{in-node } \mathbb{P}(v = v_i) = \frac{d_{in}(v_i) + \delta_{in}}{t + \delta_{in} n(t)} \tag{1.1}$$

$$\text{out-node } \mathbb{P}(v = v_i) = \frac{d_{out}(v_i) + \delta_{out}}{t + \delta_{out} n(t)} \tag{1.2}$$

Here δ_{in} and δ_{out} are the preferential attachment parameters, $d_{in}(v)$ and $d_{out}(v)$ are the current in-degree and out-degree of node v and lastly $n(t)$ is the number of nodes at time t.

Definition 3. *The in-degree $deg^-(v)$ and out-degree $deg^+(v)$ of a node $v \in \mathscr{N}$ are*

$$deg^-(v) = \sum_{w \in \mathscr{N}} M_{wv}(g) \tag{1.3}$$

$$deg^+(v) = \sum_{w \in \mathscr{N}} M_{vw}(g). \tag{1.4}$$

By increasing the preferential attachment parameters, the bias of selecting less highly connected nodes is created. Through this growth process, the degree sequence distribution will converge in distribution to a power law as shown in Bollobás et al. (2003).

The attributes of the network of the internet has provided a great degree of interest for people to study preferential attachment and scale-free graphs. The border routers of the Internet empirically have a degree distribution that is closely approximated by a power law with tail parameter of 2.1, as shown in Willinger et al. (2002). This network data is accessible to the public, and updated everyday: this accesiblity stands firmly in contrast to the limited information available on the secretive networks of financial institutions. Hurd et al. (2016) constructed a stylized model of a financial network using this method with the power law tail parameter based on Cont et al. (2010). The sample graphs were grown until attaining 1000 nodes, then the top 90 most connected node were extracted to represent the high impact financial institutions, thus giving a theoretical picture of how financial networks would look like as a scale-free graph.

Scale-free random graphs have a high density of short loops due to the preference of highly connected nodes forming connections with each other. The feedback phenomena introduced by the existence of loops makes analytical solutions to cascade models hard to arrive at. Goh et al. (2003) produced analytical results of a cascade model based on a scale-free graph using branching process techniques, but the key here is that the cascade process itself produced a tree-like path. This special property of the cascade process allowed for the affected nodes to grow into a tree graph, which in turn allowed the branching process techniques to apply. Aside from special cascade processes such as the method used in Goh et al. (2003), the majority of cascade models on scale-free random graphs require Monte Carlo simulations (Motter and Lai (2002)). This is not without its downside, since prior to each cascade process a scale-free random graph has to be sampled, and every sample of a scale-free graph has to be grown one node and one edge at a time. This

requires much more computational power to produce than configuration graphs. Zhang and Moorsel (2009) developed a method to generate fast samples of scale-free graphs using the preferential attachment method, that has been used in generating the scale-free random graphs of Hurd et al. (2016).

1.2 Using Market Impact Models for Systemic Risk

The study of market impact models has a history of less than 30 years. What began with the simple question of studying real trading strategy returns versus their theoretical returns, became a long and arduous task of finding the underlying cause and the theoretical models of the causes for these losses of returns. In this thesis, we propose to investigate two aspects of market impact models. One is a problem first considered by Merton in markets without friction and perfect competition. The second is a much newer problem, where multiple agents dwell in a single market and the market impact is shared between them.

1.2.1 Market Microstructure

Lehalle (2013) presented a detailed look at the market microstructure, as well as the "price discovery" or price formation process. For a market to exist and trades to happen, there must be both buy interests and sell interests. Even more importantly, these interests must be able to meet and exist at the same time. Fundamentally, the purpose of the limit order book is to take down the buy interests, list them according to their taking prices, and symmetrically do so for the sell interests. A trade happens when a sell interest (market order) meets a buy interest (limit order) and both are consumed, or vice versa. In this way, unconsumed buy and sell interests are preserved in the limit order book. This simple trade is governed by strict rules put into place by the market, which define the market microstructure and price formation process:

1. the auction type;
2. tick size;

3. interaction between trading platforms.

The primary equity market is simply listed firms raising capital via selling shares. But once these shares are available to the secondary market for further reselling and buying, the market microstructure takes over. Traditionally, the price formation process consisted of investors providing interest in trades, intermediaries providing advice and connection to exchanges, and finally market operators who will match the interests and create a trade. Today, it is much more common to have a diverse market place, with many intermediaries connected to many exchanges, with investors and traders who utilize High Frequency Trading (HFT) having direct access to certain trading pools, and exchanges bound by regulation to trade at the optimal price. This more complex setup creates opportunities for developing different types of market impact models for many different purposes. A well built market impact model is formulated for the investor setup, intermediary connections, as well as the exchange structure which forms the market microstructure.

1.2.2 Market Impact Models

Fundamentally, a market impact model describes the impact of limited liquidity on the price dynamics of assets. It provides the link between large trades and the usually detrimental influence they have on the prices of the assets. Controversially, a model may also be a basis for certain actions such as a central bank buying government bonds to lower the corresponding interest rates (Gatheral and Schied, 2013).

The goal of a market impact model is to determine how a large trade should be executed with minimal costs and detrimental impact to the traded asset price. This usually means breaking large trades into a sequence of smaller trades and purposely spacing them out. The logic behind this is that there is finite liquidity on the traded asset, that can be seen in the limit order book. As soon as a trade starts walking down the limit order, it in effect is creating a new asset price by drawing out the liquidity away from the market. There are various different assumptions used by market impact models, some where the limit order book will "recover" in finite time and some where the limit order book is permanently consumed. The following literature review of market microstructure describes these models, including the Almgren-Chriss market impact model, and some of their applications.

Arithmetic Brownian Motion was used as part of the price dynamics in the seminal papers by Almgren and Chriss (Almgren and Chriss, 2001),(Almgren and Chriss, 1999). This assumption leads to tractable analytic trading strategies for the optimal liquidation problem. In particular, the expected size of future price changes in dollar values does not depend on past price changes nor initial price. The usual drawback of the ABM model of negative prices still occur in this case, and is sometimes amplified by the market impact functions, as a large sell order could further drive the traded asset price down and towards negative values. However, empirical evidence shows that for short trading horizons, the ABM model closely approximates the GBM model and has minimal probability of negative prices under realistic parameters. Finally, when comparing the optimal liquidation strategies based on the Almgren-Chriss framework, Gatheral and Schied (2011) found that even under extreme values of volatility and trading horizon, there are minimal differences between the trading rate under ABM and GBM. As a result, this framework is used widely by practitioners.

There are three types of market impact which affect the price of a traded asset. The first is the temporary market price impact. It is the shortest market price impact by time. It is assumed to be short enough that it is impossible for any other trades to take advantage of the change in asset price created by the temporary market price impact of another trade. As such, temporary market impact is usually associated with the velocity of trading and directly impacts the cash account of the trader. A usual justification of temporary market impact is walking across the ask-bid spread. Quadratic temporary impact is the simplest convex function that guarantees the lack of price manipulation strategies. Transient market price impact is based on the assumption that the orders in a limit order book will refill at a certain rate after being consumed. It is used in second generation market impact models which model limit book dynamics. Finally there is the permanent market price impact. The idea behind it is that trades can and will reflect changes in the fundemental price of assets. This means that if a trade occurs, and additional information on the fundemental price of the asset is revealed by this trade, then the trading value of this asset will be affected permanently. Linear permanent market impacts greatly reduce the complexity and lead to tractable and analytical optimal trading strategies. Aside from the tractability, linear permanent impact is also necessary to prevent the existence of price manipulation strategies as shown by Gatheral and Schied (2011). There is empirical evidence for both market impacts to be a multiple of a power law function with $\alpha = 0.89$ and $\alpha = 1.6$

respectably, Almgren et al. (2005).

As mentioned above, certain price manipulation strategies may exist in market impact models. Regularity of optimal trading strategies of market impact models involve ruling out strategies which create a kind of arbitrage called "quasi-arbitrage". Unlike traditional arbitrage, there is still risk involved in these strategies, which is to say that the risk penalization is finite, and the expected return can become infinite.

Definition 4. *The expected return on a portfolio is calculated as a weighted average of all possible outcomes of the portfolio weighted by their probability of occurrence. Let X be the returns on a portfolio then,*

$$\mathbb{E}[X] = \int_{\omega \in \Omega} X(\omega) P(d\omega) \tag{1.5}$$

The expected loss is simply the negative of the expected returns.

In market impact models, it is expected that the market impact caused by trading is detrimental to the trader, and furthermore the market friction which is measured by the temporary market impact further reduces the gains of market participants. Yet even with these negative impacts, in many models, "quasi-arbitrage" strategies are possible. First we will define the three fundamental types of "quasi-arbitrage" strategies.

Definition 5. *Let S_t be an asset price process, and q_t be the holding level of that asset. Assuming $q_{T+} = 0$, then q_t describes a liquidation strategy that generates the revenue:*

$$R_T(q) = -\int_0^T S_t^q dq_t. \tag{1.6}$$

Here, S_t^q is the asset price process that is impacted by strategy q. Similarly the liquidation costs, which is the difference between the initial valuation of the asset and the amount of value recovered after liquidation is:

$$C_T(q) = q_0 S_0 - R_T(q). \tag{1.7}$$

Definition 6. *A price manipulation strategy q is a round trip strategy with strictly positive*

expected revenue:

$$\mathbb{E}[R(q)] > 0, \tag{1.8}$$
$$\textit{where } R(q) = -\int_0^T \tilde{S}_t dq_t \textit{ and } q_0 = q_T$$

Definition 7. *Transaction triggered price manipulation strategies are a set of strategies q which generate strictly greater revenue than monotonic strategies (pure buy or pure sell strategies).*

$$\mathbb{E}[R(q)] > sup(\mathbb{E}[R(\tilde{q})|\tilde{q} \textit{ is a monotonic strategy}]) \tag{1.9}$$

Definition 8. *A negative liquidation cost strategy q is such that*

$$\mathbb{E}[R(q)] > q_0 S_0 \tag{1.10}$$

The existence of any of the three strategies above leads to the existence of an optimal trading strategy which can generate unlimited expected returns. As stated previously, the existence of such strategies is ruled out in linear-quadratic market impact models. We present below two continuous time market impact models which follow the linear-quadratic setup.

The following model was introduced in Bertsimas and Lo (1998). The fundamental asset price (unaffected asset price) follows a geometric Brownian motion with μ and Σ as the vector of returns and volatility matrix.

$$\frac{dS_t}{S_t} = \mu dt + \Sigma dB_t. \tag{1.11}$$

Here, B_t is a vector of independent Brownian motions. The affected price that a trader perceives takes the form:

$$\tilde{S}_t = S_t exp(\int_0^t g(\dot{q}_s)ds + h(\dot{q}_t)), \tag{1.12}$$

where g is the permanent impact function and h is the temporary impact function. The B-L model does not admit negative prices, and cannot be forced into a negative price regime

by market impact forces either. This benefit comes at a cost of not being able to derive explicit optimal trading strategies, even with linear impact functions.

The Almgren-Chriss market impact model (Almgren and Chriss, 2001) is one of the best known market impact models. In the continuous time version of this model proposed in Bertsimas and Lo (1998), the unaffected traded asset price has the following dynamics, once again with μ and Σ being the vector of returns and volatility matrix respectively,

$$dS_t = \mu dt + \Sigma dB_t. \tag{1.13}$$

Here, B_t is a vector of independent Brownian motions. The affected price that a trader perceives takes the following form,

$$\tilde{S}_t = S_t + \int_0^t g(\dot{q}_s)ds + h(\dot{q}_t). \tag{1.14}$$

$g(*)$ and $h(*)$ are real functions with $g(0) = h(0) = 0$. In particular $g(*)$ represents the permanent impact on asset price, and $h(*)$ represents the temporary impact. The explicit trading strategy for maximizing expected return under Linear-Quadratic impact functions ($g(x) = \lambda x$ and $h(x) = \gamma x^2$) can be shown easily and is linear.

1.2.3 Optimal Liquidation

For a large trader trading in illiquid assets, the primary assumptions of classical arbitrage pricing theory fail to apply. The first assumption that a competitive market exists, means any trader may trade unlimited quantities of any relevant asset without affecting the asset's market price. The second assumption that the market is frictionless, implies that no transaction costs exist, see Cetin et al. (2004). Neither of these assumptions hold since there are always transaction costs on even the most liquid assets. For a trader exiting a dominant position in a particular asset, adverse impact on price is unavoidable. In particular, for large funds the decision to change even a small fraction of their assets under management can result in very large transactions that comprise a major part of the average daily traded volume of the assets. Similarly any trade would incur costs from crossing the ask-bid spread and possible taxes. When these two assumptions fail, a notion of liquidity risk comes apparent, and traders will aim to minimize such risk.

As shown empirically in Perold (1988), the frictionless "paper" portfolio drastically out-performed the realized portfolio with friction. Liquidity costs can not be viewed as a negligible portion of the costs of trading, and need to be quantified and kept to a minimum. On the other hand, these illiquid assets can be volatile in price and by taking additional time to trade these assets to minimize liquidity costs, the trader will encounter greater price uncertainty. Even more so in optimal portfolio selection, the optimal portfolio choice becomes dependent on both the initial portfolio and the liquidity of the different types of assets, seeAlmgren and Chriss (2001). Like the development of automatic order execution, the problem of finding the optimal trading strategy becomes a mathematical problem, see Schöneborn (2008). Financial Institutions (FI) can use such strategies for rebalancing their portfolios and to achieve a specific target portfolio.

There are two competing methods for measuring the liquidity cost. The first method comes from the models of Bertsimas and Lo (1998); Almgren and Chriss (2001), and various others and can be considered as first generation market impact models. In first generation models, the liquidity cost is modelled directly by temporary and permanent market impact functions. These market impact functions are exogenously given and depend on the velocity of trading and total trade size. The second method involves modelling the limit order book, following the works of Obizhaeva and Wang (2013), Alfonsi and Schied (2010), and others. In limit order book models, statistical parameters such as price resilience rate and density of the limit order book are exogenously provided, and used to compute the liquidity cost of trades.

Of particular interest to this thesis is the continuous time Almgren and Chriss market impact model, where the asset price process is shown in the previous section. In this model, the liquidity costs are directly generated from the permanent market impact from trades as well as the temporary market impact. In addition to minimizing the liquidity cost of the liquidation strategy, (Almgren and Chriss, 2001) set up a penalization term for the variance of the return on the liquidation strategy. This variance penalization term, allows this optimal liquidation problem to be reformulated as a mean-variance optimization problem. Since this variance penalization factor is subjective to the investor's risk tolerance, this allowed for the tailoring of different liquidation strategies for different investors.

A drawback of the mean-variance optimal liquidation problem is that its solutions are not time-consistent. Time-consistency arises when the optimal liquidation strategy

determined at one time will not be viewed as optimal at a different point in time. In the AC (Almgren-Chriss) continuous time market impact model setup, this is dealt with by limiting the admissible set of strategies to static strategies, which are pre-committed. Lorenz and Almgren (2011) utilized a discretization scheme and applied DPP to the continuous time AC model. The single update strategy where the trading strategy is allowed to update at the mid point of the trading period out-performed the pre-commitment strategy. They also showed that the improvement via additional updates diminished at a high rate.

Additional methods of arriving at time-consistent optimal liquidation strategies include the works of Gatheral and Schied (2011) where penalization using quadratic variation instead of variance was implemented. Furthermore in Schied et al. (2010), optimal liquidation strategies with constant absolute risk aversion (CARA) was studied, and the adapted optimal liquidation strategy was found to be a static strategy. This implies the optimal liquidation strategy is also time-consistent. The exponential utility case can also be formulated into a mean-variance optimization problem, hence the static optimal strategy can be found through the mean-variance techniques based on the AC model.

1.2.4 Portfolio Optimization

Portfolio Optimization is a central topic of financial mathematics with seminal papers including the likes of Markowitz (1952) and Merton (1969). Mean-Variance portfolio optimization has a long and detailed history in financial mathematics. Markowitz (1952) set the foundation for modern portfolio theory, by providing a measure of comparison between different portfolios based on the mean and variance on their returns. In particular, portfolios which have the lowest variance for a given level of expected returns form a surface called the efficient frontier. Unsurprisingly, for market impact models, Mean-Variance optimization was also one of the first methods to measure the risk aversion levels of individual traders for optimal trading strategies. In Almgren and Chriss (2001), various strategies were compared and their efficient frontiers were drawn based by varying the risk aversion parameter.

Merton (1969) provided the first look at portfolio optimization in continuous time. The paper explicitly aimed to determine the optimal portfolio composition of an investor at every instance in time. A portfolio selection strategy is optimal when it maximizes

the utility function of the investor for intermediate consumption and terminal wealth. Closed form solutions for the cases of power, logarithmic and exponential utilities were developed. The exponential utility case is equivalent to a mean-variance optimization problem: This directly relates the risk aversion parameter of exponential utility and the variance penalization parameter in the mean-variance framework.

Mathematically, mean-variance optimization allows us to take advantage of the normal innovations of Brownian motion, and produce closed form or analytic results for the optimal trading problems. Furthermore, by being able to efficiently construct optimal trading strategies under the mean-variance optimization framework, and combining with arithmetic Brownian motion driving the underlying asset price dynamics, it becomes feasible to optimize for risk measures such as Value-at-Risk by searching through the efficient frontier (Almgren and Chriss, 2001).

More importantly, when implemented as part of market impact models in general, the variance penalization term further reduces the possibility of the existence of quasi-arbitrage strategies. In optimal liquidation problems, since the trading strategies are bounded due to the given terminal condition, this drastically reduces the possible quasi-arbitrage strategies available to a trader. In optimal portfolio selection, traders have access to strategies which may become unbounded, and variance penalization becomes instrumental in disincentivising these types of strategies.

1.2.5 Multi-Agent Models

Multi-agent market impact models are a relatively new extension of market impact models that accounts for the joint impact generated by all the participants of a market. A bank's optimal trading strategy on the other hand becomes impossible without the bank knowing the positions of all other market participants. As shown in Facchinei and Kanzow (2007), the multi-agent market impact models is a generalized Nash Equilibrium game, where each agent is trying to optimize its own utility which depends on its own actions as well as the actions of the other agents. The Nash equilibrium strategies which an agent may use are defined as strategies which are optimal conditioned on the optimal strategy of all other agents. Specific to market impact models, the connection and impact of one agent's actions to another is directly caused by the market price impact mechanisms.

A paper on predatory trading (Brunnermeier and Pedersen, 2005) uses historical events to justify banks having at least some knowledge of positions of other traders in extreme circumstances. The most famous example of having knowledge of the positions of other traders is perhaps LTCM (Long Term Capital Management) in 1998: many of its counterparties were able to front run its positions, and more evidently "locals" on the Chicago Exchange Board pits were able to exploit knowledge of LTCM's short positions. This leads to the suggestion that, for multi-agent market impact models dealing with predatory trading, it could be possible that the predator traders have nearly full information of the positions of the distressed trader, and the regulations and margin calls they face. With such information, the predator agents are able to formulate optimal strategies, for a given trading period length, using round trip strategies. Brunnermeier and Pedersen (2005), further studied the effect of changing the trading period length, as well as the effect of having mismatched trading periods between the distressed trader and the predator traders.

Indeed, detailed knowledge of the positions of other traders and trading strategies may be unlikly except for one-off distressed traders. For generalized differential games, the unrealistic assumption of perfect information is typically made. In Zhang (2014), all initial positions of traders are known to all other traders. The optimal trading strategy is defined as the Nash Equilibrium strategy whereby no single trader can do better conditioned on the trading strategies of all other traders. This implicitly implies that all traders trading strategies are known and shared by all other traders. Using this perfect information setup, Zhang was able to compute a multiplayer generalization of the AC model with equilibrium strategies for optimal liquidation. Furthermore, the CARA utility and corresponding mean-variance differential game can be solved in a similar method.

Chan and Milne (2014) reviews and analyzes the relationship between the stability of the financial system and the behaviour of the agents in this system, banks, and the households. Although it is not explicitly a market impact model, the model developed in this paper utilizes many common aspects such as asset price depreciation due to liquidation. It also provides an economic basis for using market impact models to study financial stability.

Multi-agent models can be easier to handle when the volatility of the asset price process is reduced to 0. In this case, the assumptions can be fit to a network-based systemic risk model, similar to Cifuentes et al. (2005). Although this model does not optimize or

anticipate the actions of other agents in the system, it does provide a direct linkage to network-based systemic risk models.

In an extension to fit a multi-agent market impact model into a financial network, Tian and Weinan (2014) used both mechanisms similar to that of Eisenberg and Noe (2001) and an illiquid asset price subject to market price impact. Agents in the system are connected via directed edges which represent interbank loans, and each agent also holds a particular amount of illiquid asset which it may be required to sell if it is unable to meet its payment obligations. The illiquid asset price is determined by the total amount of illiquid asset liquidated during one period through an inverse demand function. In particular, agents in this system are only permitted to sell this illiquid asset, and generate proceeds to pay off its creditors. A set of heuristics and checks can be followed in order determine the equilibrium strategy, the equilibrium asset price, and the payment vector of the system. The resulting game is resolved in a single period.

1.3 Contributions of this Thesis

The two strands of research contained in this thesis may seem very distinct at first. However, I want to point out that price impact is fundamentally the same mechanism as asset firesale contagion of systemic risk, and market impact models provide a natural setup to explore the impacts of this contagion.

Network models of systemic risk also often suffer from the criticism of unrealistic behaviour. In this thesis, the second goal of studying market impact models is to understand the effects of optimal behaviour versus pre-determined threshold behaviours. The key here is to understand these differences during a crisis, and on behaviours which are not forced by regulations. One such behaviour is how the price of illiquid asset is determined in the financial system. We can directly compare the difference between the results of an asset firesale mechanism on a financial network, and the result of a multi-agent market impact model.

More directly, single agent portfolio optimization for market impact models provides the necessary tools to study the multiple agent case for market impact models. The phenomenon of asset firesales only exists because of multiple agents in the same market. To

be able to simulate this behaviour accurately in a network model, we first have to be able to generate this effect from transactions of optimizing agents. These are only some of the reasons to attempt to think of market impact models as part of systemic risk, and to integrate their effects into other models for systemic risk.

In chapter two, we develop a market impact model with temporary and permanent impact for a single utility optimizing financial institution. In this model, the financial institution is able to influence the traded asset price through buying and selling, and is penalized by a transaction cost based on the velocity of its trading. We prove the existence of an analytic solution to the optimal trading strategy, and explored the sensitivities of this optimal trading strategy. We created this single financial institution model to capture more accurate behaviors of individual agents and indeed this model showed peculiar behaviors that only exist in single player markets.

In chapter three, we develop an agent-based model utilizing the market impact model from chapter two. We are able to build on the individual self-optimizing behaviors from chapter two, and through game theoretic arguments arrive at analytical solutions which are Nash Equilibrium strategies. These Nash Equilibrium strategies are affected by the balance sheets of each agent in the system and the regulatory constraints. We also explicitly calculate the individual agents' contribution to the final asset price of this system. From a systemic risk perspective, this model is exactly a model of a pure asset firesale contagion, where agents do not have any connections to each other and can only influence the common asset prices.

In chapter four, we develop a random network model for system risk. In this model, the nodes on the directed graph are financial institutions, forming a random financial network with edges being interbank lending exposures. There are two intertwined directed cascade mechanisms in this model. The default cascade mechanism goes from debtor bank to creditor bank as bankruptcies happen. The liquidity cascade goes from creditor bank to debtor bank has liquidity hoarding occurs. In addition, a firesale mechanism was implemented, which globally affects the fixed asset price of every financial institution in the network. As shown in the previous chapter, this additional firesale mechanism provides a source of systemic risk through self-preserving actions. We were able to construct a model of 89 European financial institutions based on EBA stress testing data, and compute various systemic risk measures on the outcome of the cascade process.

Chapter 2

Optimal Portfolios of Illiquid Assets

"Optimal Portfolios of Illiquid Assets" is a joint work with Dr. Tuan Tran and Prof. Thomas Hurd. Contributions by the author include: 1. Data collection, computer programming, and numerical analysis; 2. Equal share of drafting and majority of finalization of the paper; 3. Equal share of the technical derivations of this paper, and numerous studies into alternative solutions. As of Oct 31, 2016, this paper is under review for publication by SIAM Journal of Financial Mathematics. The page layout of this paper has been changed to better fit the a thesis format, as well as merging the bibliography with the thesis bibliography.

2.1 Introduction

Long after such landmark contributions as the Markowitz mean-variance strategy (Markowitz (1952)) and the Merton portfolio model introduced in Merton (1969), our understanding of optimal portfolio selection has continued to develop. We now have learned how to analyze investment in imperfect markets that have frictions such as transaction costs (Davis and Norman (1990), Perold (1988)) and price impact (Almgren and Chriss (2001), Almgren (2003), Schöneborn (2008)), and have complex dynamics such as jumps (Cartea and Jaimungal (2015), Moazeni et al. (2013) and Pham and Tankov (2008)). Indeed, this problem has generated hundreds of research papers. Our goal now is to present a solvable model

of optimal investment for a large financial institution (FI) in a many-asset setting. It is based on the expected utility maximization criterion, and it accounts for *market illiquid*, which means the transaction costs to pay and the fact that trades have a permanent price impact. The underlying investment assets, which may be very illiquid, are assumed to follow Bachelier dynamics, meaning they are modelled by correlated arithmetic Brownian motions. For these assumptions to make financial sense, the optimal strategy should be implemented only over a time horizon $[0, T]$ short enough that the Bachelier dynamics remains a reasonable approximation (we take as a benchmark $T = 1/2$ years in our examples).

The class of optimal strategies we obtain has several remarkable properties. First, the general multidimensional problem has a closed-form solution expressible in terms of a matrix-valued equation that can be efficiently computed with a controllable error. Second, the solution depends on the full range of important parameters: temporary price impact, permanent impact, risk aversion, the initial portfolio weights, the risk free interest rate, and the parameters underlying the Bachelier dynamics. Thirdly, the optimal strategies, which are a priori adapted processes that solve a version of Merton's problem, turn out to be deterministic over a finite time horizon and to solve a version of the Markowitz mean-variance optimization. This property implies that our investment strategies are fully consistent with dynamic programming, despite being deterministic solutions of a time-inconsistent mean-variance optimization problem.

The aim of this paper is to study the effect of market illiquidity on the behaviour of an FI.*Funding illiquidity* (see for example Brunnermeier and Pedersen (2009)) is the distinct effect that the balance sheet of an FI may experience funding shocks caused by unanticipated withdrawals by depositors. To keep the focus of the paper squarely on market liquidity, funding illiquid is ruled out by the assumption that deposits are constant and sufficient to support all asset purchases of the FI.

The proposed model and its solution is closely related to some important contributions to the existing literature. Our solutions reduce to the Markowitz optimal portfolios, or equivalently to Merton's optimal solutions, when permanent and temporary impact are both assumed to be zero. The posed finance problem is inspired by the mean-variance optimal liquidation problem studied by Almgren and Chriss (2001), but differs in that there is no constraint placed on the portfolio holdings at the terminal time T. Finally, under certain initial conditions the FI will seek to liquidate a large position, creating what

has been called an asset fire sale. Our strategies extend to this setting and give natural criteria similar to those discussed by Brown et al. (2010) that solve the problem of the order in which different assets are liquidated.

2.2 Optimal Portfolio Strategies

This paper will investigate the investment strategies of a large financial institution (FI) with CARA risk preferences (CARA is short for constant absolute risk aversion that trades continuously over a finite time horizon $[0, T]$ in a market with imperfect liquidity. This is similar to a problem studied in Zhang (2014). The changes caused by rebalancing a portfolio of a large FI may amount to a large fraction of the total daily volume traded of these assets and significantly impact these assets' prices. It is well understood that this effect will lead the FI to break large orders into small portions spread over time to reduce market liquidity costs, while still aiming to rebalance its portfolio. By taking additional time to reduce liquidity costs, the FI now faces additional uncertainty in the price of the assets. To handle this delicate balance between liquidity costs and price uncertainty, the FI will be inclined to consider utility optimization.

There are sound economic reasons to optimize using an exponential (CARA) utility function: It leads to a tractable time-consistent strategy where additional information does not provide additional utility, and is similar to the original Mean-Variance optimization of Almgren and Chriss (2001). Since the strategy is only implemented over $[0, T]$, at time T the FI will update its information and continue in a similar way to rebalance over the subsequent period. This rebalancing is necessary to account for shortcomings of the model, changes in the balance sheet, and *unanticipated events*that cause fundamental changes to the parameters of the price dynamics.

A number of simplifications will be assumed about this problem. The total information available to the FI up to any given instant of time t is modelled by a filtration $\{\mathscr{F}_t\}_{t\geq 0}$ on a given probability space $(\Omega, \mathscr{F}, \mathbb{P})$. The market consists of one risk-free asset with zero interest rate, and d risky assets whose *true* price process is $S_t = (S_t^{(1)}, ..., S_t^{(d)})'$ and whose *transaction* price process is $\tilde{S}_t = (\tilde{S}_t^{(1)}, ..., \tilde{S}_t^{(d)})'$. Here and in the following, we adopt matrix notation where M' denotes the matrix transpose of M. Let us denote the vector of the

amounts held in risky assets by $(q_u)_{u\in[t,T]}$ and the vector of trading rates of the large trader by $v_u := \dot{q}_u := dq_u/du, u \in [t,T]$.

Like Almgren and Chriss (2001) and others, we suppose that the price of risky assets follows a d-dimensional Bachelier model with both linear permanent and linear temporary market impact, para-parametrized by Λ and Γ respectively:

$$\begin{aligned} dS_t &= (\Lambda v_t + \mu)\, dt + \Sigma\, dB_t\ , \\ \tilde{S}_t &= S_t + \Gamma v_t\ . \end{aligned} \tag{2.1}$$

Here, B_t is a $d-$dimensional Brownian motion and $\Sigma \in \mathbb{R}^{d\times d}$ is the volatility matrix. The drift term $\mu = b - d + \Lambda Q \in \mathbb{R}^d$ is assumed to be constant. It takes into account the trending rate b, dividend rate d and aggregated permanent market price impact due to external traders Q.

A more general formulation of the model that does not require linear market impact is certainly possible, and will not change many of the same basic properties. However, the assumption of linear impact leads to significantly more tractable optimal strategies. Moreover, as shown by Gatheral and Schied (2011), so-called *dynamic arbitrage* is ruled out by choosing the permanent impact to be linear. It is further assumed that the permanent and temporary impact matrices Λ and Γ are symmetric and non-negative definite. The assumption that Γ is symmetric is without loss of generality. On the other hand, Λ is assumed to be symmetric not for economic reasons but for convenience: when it has an anti-symmetric part, a somewhat more complicated explicit solution is obtainable. Models similar to ours have been studied by Almgren and Lorenz (2007), Gatheral and Schied (2011) and Schied and Schöneborn (2009).

2.2.1 The Merton Problem

Merton's problem, introduced in Merton (1969), aims to determine the strategies followed by utility optimizing investors in continuous time market models. To this end, we now consider the most general portfolio strategy, or control process, that trades within the market impact model (2.1) over some time interval $[s,t] \subset \mathbb{R}_+$. In our setting, each possible strategy will be simply a d-dimensional trading rate process $v = (v_u)_{u\in[s,t]}$ that is integrable

and adapted to the information filtration $\{\mathscr{F}_t\}$: We denote the set of such *admissible* strategies by $\Pi^{ad}[s,t]$. The subclass of *deterministic* strategies where each value $v_u, u \in [s,t]$ is $\mathscr{F}_s$ measurable is denoted by $\Pi^{det}[s,t]$.

Given any control process $v \in \Pi^{ad}[0,s]$ for $0 < t < s$, the *cash net of debt owed* $C_t := C_t^v$ and marked-to-market *equity*, or *assets net of debt owed*, $X_t := X_t^v := C_t + q_t' S_t$ are given by:

$$C_t = C_0 - \int_0^t v_u' \tilde{S}_u \, du = C_0 - \int_0^t v_u' S_u \, du - \int_0^t v_u' \Gamma v_u \, du \, , \tag{2.2}$$

$$X_t = X_0 + \int_0^t q_u \, dS_u - \int_0^t v_u' \Gamma v_u \, du \, . \tag{2.3}$$

where the second equation is obtained by integration by parts. Note that here and henceforth, the superscript v that labels processes controlled by v will be omitted.

The interpretation of (2.2) and (2.3) in terms of the firm's balance sheet is that assets are stochastic due to fluctuations of S, while the debt, thought of as deposits, is assumed to be constant and sufficient to fund all trades. In other words, we focus on market illiquidity without funding illiquidity. It is consistent with the Principle of Limited Liability that a firm becomes insolvent when its equity X_t becomes negative. In the following, an insolvent firm with negative equity $X_T < 0$ at a time T, will be declared to be *in default*, implying that the laws of bankruptcy will be applied to the firm.

The FI can now try to solve Merton's optimal problem of a CARA investor with constant absolute risk aversion parameter $\lambda > 0$ over any period $[t,T]$. For each t, they may express the value function J_t achieved in terms of a *certainty equivalent value* W_t,

$$J_t := -e^{-\lambda W_t} := \sup_{v \in \Pi^{ad}[t,T]} \left(-\mathbb{E}[-e^{-\lambda X_T} | \mathscr{F}_t] \right) \, . \tag{2.4}$$

If the supremum exists, it is achieved by adopting an *optimal control* denoted by $v^*(t) = (v_u^*(t))_{u \in [t,T]}$, which will be an adapted process over $[t,T]$. The CARA investment problem in general always satisfies the *dynamic programming principle* (see Schied et al. (2010)), which means that for any $s \le t \le T$, $v_u^*(t) = v_u^*(s)$ for all $u \ge t$ and

$$-e^{-\lambda W_s} = \sup_{v \in \Pi^{ad}[s,t]} \left(-\mathbb{E}[e^{-\lambda W_t} | \mathscr{F}_s] \right) \, . \tag{2.5}$$

An investor restricted to deterministic strategies over $[t,T]$ cannot achieve a higher certainty equivalent value than equation (2.4). Therefore, if $\widetilde{W}_t$ is defined by

$$-e^{-\lambda \widetilde{W}_t} := \sup_{v \in \Pi^{det}[t,T]} \left(-\mathbb{E}[-e^{-\lambda X_T}|\mathscr{F}_t]\right) \tag{2.6}$$

then $\widetilde{W}_t \leq W_t$. The first result of this paper, stated next, is that (2.4) is always optimized by deterministic strategies and therefore $W_t = \widetilde{W}_t$ for $t \geq 0$. Moreover, it will be found in subsequent sections that the optimal control and value functions can be expressed in closed forms involving one-dimensional integrals that solve a system of ordinary differential equations of Riccati type. First, however, a note about notation: Because v^* and q^* turn out to be deterministic, we henceforth replace the stochastic process notation v_u by function notation $v(u)$ and moreover suppress the dependence on the investment period $[t,T]$.

Theorem 1. *Under the above modelling assumptions, there is a (possibly infinite) maximal time $T^* \in \mathbb{R}_+ \cup \{\infty\}$) such that for any finite time horizon $[t,T]$ with $0 \leq t \leq T \leq T^*$:*

1. *The optimal strategy $v^*(u), u \in [t,T]$ exists, is unique and $\mathscr{F}_t$ measurable, hence deterministic.*

2. *The value function $\widetilde{W}_t$ achieved over $[t,T]$, when restricted to deterministic strategies, equals W_t.*

3. *The value function has the form $W_t = X_t + V(T-t,q)$ where $V(\tau,q), \tau = T-t$ solves the non-linear partial differential equation*

$$-\partial_\tau V + q'\mu - \frac{\lambda}{2} q'\Sigma\Sigma' q + \frac{1}{4}(\Lambda q + \partial_q V)'\Gamma^{-1}(\Lambda q + \partial_q V) = 0, \quad V(0,q) = 0 \tag{2.7}$$

   on the domain $[0,T] \times \mathbb{R}^d$.

4. *Given initial holdings q at time t, the optimal portfolio holdings $q^*(u)$ for $u \in [t,T]$ solves the system of ODEs:*

$$\frac{dq}{du} = \frac{\Gamma^{-1}}{2}\left(\partial_q V(T-u,q)' + \Lambda q\right) , \quad q(t) = q \ . \tag{2.8}$$

The proof of this theorem is found in the Appendix. As we shall see in Section 2.3, $V(\tau,q)$ is a quadratic form in q with time-dependent coefficients and thus the ODE (2.8)

for q^* is linear and can be solved explicitly.

2.2.2 Mean, Variance, Probability of Default and Time-Consistency

From equations (2.2) and (2.3) we can deduce that, if v is deterministic, then for any $0 \leq s \leq t \leq T$, the equity X_t conditioned on $\mathscr{F}_s$ is normally distributed with mean and variance given by

$$\mathbb{E}[X_t|\mathscr{F}_s] = X_s + \int_s^t \Big(q'(u)(\Lambda v(u) + \mu) - v'(u)\Gamma v(u)\Big)\, du, \tag{2.9}$$

$$\mathrm{VaR}[X_t|\mathscr{F}_s] = \int_s^t q'(u)\Sigma\Sigma' q(u)\, du\ . \tag{2.10}$$

In particular, the fact that $X_T|\mathscr{F}_t$ is always normal implies that

$$\mathbb{E}[e^{-\lambda X_T}|\mathscr{F}_t] = e^{-\lambda\left(\mathbb{E}[X_T|\mathscr{F}_t] - \frac{\lambda}{2}\mathrm{VaR}_{[X_T|\mathscr{F}_t]}\right)} \tag{2.11}$$

and hence from (2.6) and Theorem 1 one deduces that

$$W_t = \widetilde{W}_t = \sup_{v\in\Pi^{det}[t,T]} \left(\mathbb{E}[X_T|\mathscr{F}_t] - \frac{\lambda}{2}\mathrm{VaR}[X_T|\mathscr{F}_t]\right)\ . \tag{2.12}$$

This demonstrates the well-known equality of the certainty equivalent value for CARA optimization with the value function for Markowitz' mean-variance (M-V) optimization, as well as the coincidence of their optimal strategies, when the optimal equity processes under consideration are all normally distributed.

In practice, the firm's*default probability* (DP), meaning the probability that $X_T < 0$, may be preferable to variance as a risk measure for institutional investors, as it gives more information about bad scenarios that need to be controlled. In the Bachelier model, the normality that follows for deterministic strategies implies that over any time horizon $[t, T]$, the Mean-Variance (M-V) criterion

Problem M-V

$$W_V(t,T,q,x,E) := \min_{v\in\Pi^{det}[t,T]} \mathrm{Var}[X_T|\mathscr{F}_t] \quad \text{subject to } \mathbb{E}[X_T|\mathscr{F}_t] = E\ , \tag{2.13}$$

and the Mean-Default Probability (M-DP) criterion

Problem M-DP

$$W_{DP}(t,T,q,x,E) := \min_{v\in\Pi^{det}[t,T]} \mathbb{P}[X_T < 0|\mathscr{F}_t] \quad \text{subject to } \mathbb{E}[X_T|\mathscr{F}_t] = E\ , \tag{2.14}$$

are both solved by the same optimal trading strategy when $\mathbb{E}[X_T|\mathscr{F}_t] = E > 0$. This is because $\mathbb{P}[X_T < 0|\mathscr{F}_t]$ is strictly increasing in $\mathrm{Var}[X_T|\mathscr{F}_t]$ as long as $\mathbb{E}[X_T|\mathscr{F}_t] > 0$ is fixed. Moreover, if $v^*(\lambda)$ denotes the optimizer of (2.12), and $X_T^*(\lambda)$ is the optimal equity it achieves, then $v^*(\lambda)$ also optimizes problems (2.13) provided $E = E(\lambda) := \mathbb{E}[X_T^*(\lambda)|\mathscr{F}_t]$, and also (2.14) if in addition $E(\lambda) > 0$.

Since Merton's optimal problem (2.4) satisfies Bellman's Dynamic Programming Principle at all times, its optimal strategies are "time-consistent", which means that the optimal strategies computed for any two periods $[t,T]$ and $[s,T]$ always coincide on the intersection $[s\vee t, T]$. On the other hand, it is known that mean-variance optimization is generally time-inconsistent and optimal *adapted* strategies starting at one time do not usually appear optimal at a later time. Surprisingly, Theorem 1 combined with equation 2.12 implies that in the present context, both the mean-variance and mean-default probability problems (2.13) and (2.14) are in fact time-consistent, provided the optimization is restricted to deterministic strategies. The following result summarizes these relationships.

Corollary 2. *For any fixed time horizon $[t,T]$, let $E(\lambda) = \mathbb{E}[X_T^*(\lambda)|\mathscr{F}_t]$ be the expected value of equity computed for the optimal strategy $v^*(\lambda)$ of the CARA investment problem* (2.12) *with risk aversion parameter λ. Let $\underline{E}$ and $\overline{E}$ be the infimum and supremum of $E(\lambda)$ when λ varies over $[0,\infty)$. Then:*

1. *For any $E \in (\underline{E}, \overline{E})$, there exists a unique $\lambda = \lambda(E)$ such that $E(\lambda) = E$.*

2. *For all possible values of $E(\lambda)$, the deterministic optimal strategies v^* of Problem M-V coincide with the unique adapted optimal strategy $v^*(\lambda)$ of* (2.12).

3. *The optimal strategies computed for any two periods $[t,T]$ and $[s,T]$ are time-consistent, meaning they coincide on the intersection $[s \vee t, T]$.*

4. *If $E(\lambda) > 0$, the deterministic optimal strategies v^* of Problem M-V and Problem M-DP also coincide with each other.*

2.3 Explicit Optimal Strategies

We now exploit the tractability of Merton's problem in the market impact setting to obtain closed formulas (involving matrix algebra and one dimensional integration) for the optimal trading curve of the financial institution (FI). The techniques invoked in this section are closely related to the methods developed for the optimal liquidation problem in Almgren and Chriss (2001).

Proposition 3. *Under the modelling assumptions of Theorem 1, for any finite T with $T \leq T^*$, the value function $W_t := W(t,T,q,x) = x + V(\tau,q), \tau = T-t$ for any $t \in [0,T]$ has the form*

$$V(\tau,q) = q'A(\tau)q + B(\tau)q + C(\tau) \tag{2.15}$$

where A, B, C are matrix valued functions of dimension $[d,d]$, $[1,d]$ and $[1,1]$ respectively with A symmetric. These functions satisfy Riccati-type ODEs for $\tau > 0$:

$$\frac{\partial A}{\partial \tau} - (A+\Lambda/2)'\Gamma^{-1}(A+\Lambda/2) + \frac{\lambda}{2}\Sigma\Sigma' = 0, \quad A(0) = 0, \tag{2.16}$$

$$\frac{\partial B}{\partial \tau} - B\Gamma^{-1}(A+\Lambda/2) - \mu' = 0, \quad B(0) = 0, \tag{2.17}$$

$$\frac{\partial C}{\partial \tau} - \frac{1}{4}B\Gamma^{-1}B' = 0, \quad C(0) = 0. \tag{2.18}$$

The next theorem, whose proof is given in Appendix A, solves this system of Riccati equations in closed form in terms of $E := \Gamma^{-1/2}(\Lambda/2)\Gamma^{-1/2}$ and the symmetric square root D of

$$D^2 := \frac{\lambda}{2}\Gamma^{-1/2}\Sigma\Sigma'\Gamma^{-1/2} .$$

It also provides a closed form for the optimal strategy q^*.

Theorem 4. *1. The solution of the system of Riccati equations* (2.16)–(2.18) *over the maximal interval* $[0, T^*]$ *is given by*

$$A(\tau) = \Gamma^{\frac{1}{2}}\Big(V(\tau)U(\tau)^{-1} - E\Big)\Gamma^{\frac{1}{2}} \tag{2.19}$$

$$B(\tau) = \overline{\mu}'\big(E - V(\tau)\big)U^{-1}(\tau)\Gamma^{\frac{1}{2}} \tag{2.20}$$

$$C(\tau) = \frac{1}{4}\overline{\mu}'\left(\int_0^\tau \big(E - V(s)\big)\big(U'(s)U(s)\big)^{-1}(E - V'(s))ds\right)\overline{\mu} \tag{2.21}$$

where $\overline{\mu} := D^{-2}\Gamma^{-1/2}\mu$ *and the matrix valued functions* U, V *are given by*

$$U(\tau) \quad = \ \cosh(D\tau) - D^{-1}\sinh(D\tau)E \tag{2.22}$$

$$V(\tau) \quad = -\sinh(D\tau)D + \cosh(D\tau)E. \tag{2.23}$$

2. The maximal time horizon T^* *is*

$$T^* = \ \inf\{\tau > 0 : U(\tau) \text{ is not invertible}\}\ .$$

T^* *is finite if* $D < E$ *and* ∞ *if* $D > E$.

3. For any (t, T, q, x), *the optimal trading curve* $q^*(u)$ *over the period* $[t, T]$ *is*

$$q^*(u) \ = \ \Gamma^{-1/2}U(T-u)U^{-1}(T-t)\Gamma^{1/2}q \tag{2.24}$$

$$+\frac{1}{2}\Gamma^{-1/2}U(T-u)\int_t^u U^{-1}(T-r)\Gamma^{-1/2}B'(T-r)dr \tag{2.25}$$

4. For any (t, T, q, x), *the expected value and variance of the optimal terminal equity are:*

$$\mathbb{E}[X_T^*(\lambda)|\mathscr{F}_t] \ = \ x + q'\Big(A(T-t) - \frac{\lambda}{2}L(T-t)\Big)q \tag{2.26}$$

$$+\ \Big(B(T-t) - \frac{\lambda}{2}M(T-t)\Big)q + C(T-t) - \frac{\lambda}{2}N(T-t)\ ,$$

$$VaR[X_T^*(\lambda)|\mathscr{F}_t] \ = \ q'L(T-t)q + M(T-t)q + N(T-t)\ , \tag{2.27}$$

where formulas for L, M, N *are given in Appendix A.*

In the special case when D and E are commuting matrices, these formulas decouple into d one-dimensional problems, each of which is similar to the single risky asset case we next discuss.

2.3.1 The Case of a Single Risky Asset

In the single risky asset case, one can verify that the scalar functions A, B, C and the optimal trading strategy $q^*(u)$ have comparatively simple formulas obtained by reducing those given in Theorem 4. Notice that several distinct possibilities are determined by the relation between $D = \Sigma\sqrt{\frac{\lambda}{2\Gamma}}$ and $E = \frac{\Lambda}{2\Gamma}$.

Proposition 5. *In the single asset case,*

1. *When $D > E$ or $2\lambda\Gamma\Sigma^2 > \Lambda^2$, denote $K = \tanh^{-1}(E/D)$ we have $U(\tau) = \frac{\cosh(D\tau - K)}{\cosh K}$. The formulas can be rewritten in terms of hyperbolic functions as follows*

$$\begin{aligned} A(\tau) &= -\Gamma D\left(\tanh(D\tau - K) + \tanh K\right) && (2.28)\\ B(\tau) &= \frac{\mu}{D}\left(\frac{\sinh K}{\cosh(D\tau - K)} + \tanh(D\tau - K)\right) && (2.29)\\ C(\tau) &= \frac{\mu^2}{4\Gamma D^3}\left((\sinh^2(D\tau - K) - 1)(\tanh(D\tau - K) + \tanh K) - D\tau\right) && (2.30)\\ &+ \frac{\mu^2}{2\Gamma D^2}\left(\tanh K - \frac{\sinh K}{\cosh(D\tau - K)}\right). && (2.31) \end{aligned}$$

The optimal trading strategy is given by

$$\begin{aligned} q(u)^* &= \frac{\cosh(D\tau_u - K)}{\cosh(D\tau_t - K)}q + \frac{\mu}{2\Gamma D^2}\left(1 - \frac{\cosh(D\tau_u - K)}{\cosh(D\tau_t - K)}\right) && (2.32)\\ &+ \frac{\mu \sinh K \cosh(D\tau_u - K)}{2\Gamma D^2}\left(\tanh(D\tau_u - K) - \tanh(D\tau_t - K)\right), && (2.33) \end{aligned}$$

where $\tau_s = T - s$.

2. *When $D = E$ or $2\lambda\Gamma\Sigma^2 = \Lambda^2$*

$$A(\tau) = 0 \tag{2.34}$$
$$B(\tau) = \frac{\mu}{D}\left(e^{D\tau} - 1\right) \tag{2.35}$$
$$C(\tau) = \frac{\mu^2}{2D\lambda\Sigma^2}\left(\frac{1}{2}e^{2D\tau} - 2e^{D\tau} + D\tau + \frac{3}{2}\right). \tag{2.36}$$

3. *When $D < E$ or $0 < 2\lambda\Gamma\Sigma^2 < \Lambda^2$, denote $K = \coth^{-1}(E/D)$ we have $U(\tau) = -\frac{\sinh(D\tau - K)}{\sinh K}$. The formulas can be rewritten in terms of hyperbolic functions as follows*

$$A(\tau) = -\Gamma D\left[\coth(D\tau - K) + \coth K\right] \tag{2.37}$$
$$B(\tau) = \frac{\mu}{D}\left(\frac{-\cosh K}{\sinh(D\tau - K)} + \coth(D\tau - K)\right) \tag{2.38}$$
$$C(\tau) = -\frac{\mu^2}{4\Gamma D^3}\left((\cosh^2(D\tau - K) + 1)(\coth(D\tau - K) + \coth K) + D\tau\right) \tag{2.39}$$
$$+ \frac{\mu^2}{2\Gamma D^2}(\coth K + \frac{\cosh K}{\sinh(D\tau - K)}). \tag{2.40}$$

The optimal trading strategy is given by

$$q(u)^* = \frac{\sinh(D\tau_u - K)}{\sinh(D\tau_t - K)}q + \frac{\mu}{2\Gamma D^2}\left(1 - \frac{\sinh(D\tau_u - K)}{\sinh(D\tau_t - K)}\right) \tag{2.41}$$
$$-\frac{\mu\cosh K\sinh(D\tau_u - K)}{2\Gamma D^2}\left(\coth(D\tau_u - K) - \coth(D\tau_t - K)\right), \tag{2.42}$$

where $\tau_s = T - s$.

4. *When $\lambda = 0$, we have $U(\tau) = 1 - E\tau$ and $V(\tau) = E$. Moreover*

$$A(\tau) = \frac{\Lambda}{2}\left(\frac{1 - U(\tau)}{U(\tau)}\right) \tag{2.43}$$
$$B(\tau) = \frac{\mu\Gamma}{\Lambda}\left(\frac{1 - U(\tau)^2}{U(\tau)}\right) \tag{2.44}$$
$$C(\tau) = \frac{\mu^2\Gamma^2}{6\Lambda^3}\left(\frac{-U(\tau)^4 + 6U(\tau)^2 - 8U(\tau) + 3}{U(\tau)}\right). \tag{2.45}$$

The optimal trading strategy is given by

$$q(u)^* = U(\tau_u)\left(\frac{q}{U(\tau_t)} + \frac{\mu}{4\Gamma E^2}\left(U(\tau_t) + \frac{1}{U(\tau_t)} - U(\tau_u) - \frac{1}{U(\tau_u)}\right)\right). \qquad (2.46)$$

The third and fourth cases are the cases where $T^* < \infty$, and one finds the solutions become unbounded: $\lim_{\tau \to T^*} A(\tau) = \infty$. In cases 1 and 2, the solutions are bounded for all τ, and $T^* = \infty$.

2.3.2 Small Perturbations of Merton's Solution

In his original paper Merton (1969), Robert Merton presented the exact solution to the problem of optimal investment in a frictionless market for an asset price that follows a geometric Brownian motion. His solution technique also leads to an exact solution of our present model in the limit of zero market impact, $\Lambda = 0, \Gamma = 0$, which we will call the "Merton solution". It is of some interest to consider the explicit general solution from the previous section as a perturbation of the Merton solution, and to investigate the nature of its convergence as market impact goes to zero. We suppose that $\Lambda = \epsilon\Lambda_1, \Gamma = \epsilon\Gamma_1$ with small ϵ and denote by $W(t, T, x, q, \epsilon)$ the certainty equivalent value function with its dependence on ϵ. For simplicity, we confine our attention to the single asset case of the previous section.

The Merton solution over the period $[t, T]$ with initial conditions $X_t = x, q_t = q$ involves an instantaneous trade that incurs no trading cost, to the optimal value $q^M := \frac{\mu}{\lambda\Sigma^2}$. This portfolio is then held constant. One can show that this strategy achieves the certainty equivalent value function $W(t, T, x, q, 0) = x + \frac{\mu^2(T-t)}{2\lambda\Sigma^2}$ which we note is independent of q.

Now, for small ϵ, the general solution of our model is given by Case 1 of Proposition 5, which leads to the following perturbative expansion

$$W(t, T, x, q, \epsilon) = W(t, T, x, q, 0) + \epsilon^{1/2} Q(q) + o(\epsilon^{3/2}). \qquad (2.47)$$

where

$$Q(q) \;:=\; \Gamma_1 D_1 q^2 + \frac{\mu}{D_1} q + (2E - 1) D_1. \qquad (2.48)$$

Here we define $D_1 = \Sigma\sqrt{\frac{\lambda}{2\Gamma_1}}$ which does not depend on ϵ and we have $D = \epsilon^{-1/2}D_1 \to \infty$ as $\epsilon \to 0$. It is obvious that E does not depend on ϵ either. Thus the value function of our problem converges to the value function of the Merton solution with rate of convergence $\epsilon^{1/2}$.

The optimal holding at the terminal time is given by

$$q_T^* = q^M + (q_t - q^M)\frac{\cosh K}{\cosh(D\tau - K)} + Eq^M\frac{1 - \tanh(D\tau - K)}{D(1 - \tanh^2 K)}. \tag{2.49}$$

Here $\tau := T - t$. It is straightforward that $\lim_{\epsilon\to 0} K = 0$, hence $\lim_{\epsilon\to 0} q_T^* = q^M$.

Let $\tilde{A}(\tau) := \frac{U(\tau)}{V(\tau)} = D\tanh(D\tau - K)$, the trading rate at the initial time t is given by

$$\begin{aligned} v_t &= \lim_{s\to t}\dot{q}_s && (2.50)\\ &= q_t\tilde{A}(\tau) + U(\tau)\frac{\mu}{2\Gamma D^2}\left[\frac{E(D^2 - (\tilde{A}(\tau))^2)}{D^2 - E^2} - \frac{\tilde{A}(\tau)}{U(\tau)}\right] && (2.51)\\ &= (q_t - q^M)D\tanh(D\tau - K) + \frac{\mu E\cosh K}{2\Gamma_1 D_1^2\cosh(D\tau - K)}. && (2.52) \end{aligned}$$

Note that $\lim_{\epsilon\to 0}\tanh(D\tau - K) = 1$ and $\lim_{\epsilon\to 0}\cosh(D\tau - K) = \infty$, we have $\lim_{\epsilon\to 0} v_t = \pm\infty$ depending on if $q_t > q^M$ or $q_t < q^M$, i.e. the optimal strategy is to trade rapidly in the beginning. We then conclude that the optimal trajectory $q^*(u, \epsilon)$ converges to an L–shaped or Γ–shaped curve when the market impact tends to zero.

This result implies that when market impact is low, the firm will follow an optimal trading strategy very close to the constant holding strategy of the Merton problem. A more surprising fact is the portfolio which starts at the Merton portfolio will remain constant if permanent impact has $\Lambda_1 = 0$, and all strategies regardless of initial portfolios will move towards the Merton portfolio for sometime initially. In the following subsection, we will show similar results for the multi-asset case.

2.4 Numerical Investigations

We now consider the investment behaviour of a hypothetical unregulated financial institution, such as a hedge fund or mutual fund. The firm trades a single risky asset, with initial price $S_0 = \$100$, in a market with a 0% risk free rate of return. They use our CARA optimal investment model to trade over non-overlapping half-year trading periods: we focus here on the period $[0, T]$, $T = 1/2$. The CARA risk aversion parameter λ is chosen to be consistent with a target default probability of 1% for each period. Thus the firm will trade aggressively to maximize their expected return with a quite high tolerance to the potential of default.

The calibrated parameters of the model given in Table 3.2 are taken to be fixed at the beginning of the period $t = 0$. Note that the firm uses the Bachelier model only for a short period, and expects to recalibrate at the beginning of the each successive period. Since the risky asset is illiquid, there is market impact related to the velocity of trading and the total amount traded: these are assumed to give the temporary and permanent market impact parameter estimates $\Gamma = \$10^{-7}$years/ (units traded)^2 and $\Lambda = \$4 * 10^{-8}$/unit.

Balance sheets for a small, medium and large firm will be considered, all with a risky asset-to-equity ratio of $4 : 1$. The initial stock holdings are $q_0 = [50000, 200000, 800000]$ from which the initial firm equity and cash net of debt are then determined to be $X_0 = 0.25 * q_0 * S_0, C_0 = -0.75 * q_0 * S_0$. In all three cases, as indicated above, the financial institution targets a fixed default probability under the optimal strategy. This is implemented by choosing the internal risk aversion parameter λ so that the $W_{DP}(0, T, q_0, X_0, E(\lambda)) = 0.01$. Note that even though the optimal strategy does not depend on X_0 for fixed λ, this specification of λ depends on X_0. Thus firms that differ only in X_0 do adopt different investment strategies.

2.4.1 The Efficient Frontier

Figure 2.1 shows for each of the three firms how the expected rate of return on equity (ERR) and default probability (DP) for their CARA/MV/DP optimal strategies depend as λ varies over the set of feasible values $[0, \infty)$. These quantities are computed by the

Table 2.1: Benchmark Parameters

Calibrated Parameter	**Model Parameter Value**
Initial Stock Price	$S_0 = \$100$
Trading Period	$[0, T],\ T = 0.5$ year
20% Annualized Volatility	$\Sigma = \$20/\text{unit}/\sqrt{\text{year}}$
4% Annual Growth	μ =4/unit/year
Temporary Market Impact	$\Gamma = \$10^{-7}\text{year}/\ (\text{unit})^2$
Permanent Market Impact	$\Lambda = \$4 * 10^{-8}/\text{unit}$
λ such that Probability of Default = 0.0005	λ varies
Initial Holdings	$q_0 = [50000, 200000, 800000]$
Initial Cash net Debt owing	$C_0 = -0.75 * q_0 * S_0$
Initial Equity	$X_0 = 0.25 * q_0 * S_0$

formulas

$$\text{ERR}(\lambda) = \frac{1}{T}\left(\frac{E(\lambda)}{X_0} - 1\right)\ , \quad \text{DP}(\lambda) = N\left(-\frac{E(\lambda)}{\sqrt{V(\lambda)}}\right) \tag{2.53}$$

where $N(\cdot)$ denotes theCDF of the standard normal and $E(\lambda), V(\lambda)$ are given by (2.26) and (2.27) . Such a graph is called an *efficient frontier*, and it summarizes the results a firm may achieve by adopting different possible risk aversion parameters.

As explained earlier, the three firms each select the optimal investment strategy given by the value λ that leads to $\text{DP}(\lambda) = 0.01$: with the benchmark parameters given in Table 3.2, the three values they compute are $\lambda = [2.56 \times 10^{-7}, 6.7 \times 10^{-8}, 1.83 \times 10^{-8}]$. While Figure 2.1 suggests that, ceteris paribus, larger firms have a lower efficient frontier, this ordering can be made to reverse by increasing the permanent impact parameter.

2.4.2 Properties of the Optimal Trading Curve

To better understand the properties of the optimal investment strategies that result from our method, we now investigate how the three hypothetical firms' optimal trading in the single asset case compare as important model parameters are varied away from the benchmark parameters of Table 3.2. Figures 2.2 and 2.3 summarize the results of four experiments, and show how the firms' optimal trading strategies over the time period $[0, 1/2]$ years change as the asset rate of return, asset volatility, temporary market impact and permanent market impact are made to vary one at a time. In each figure, the red curve denotes the benchmark parametrization, while the other two curves show the result as one specific parameter is

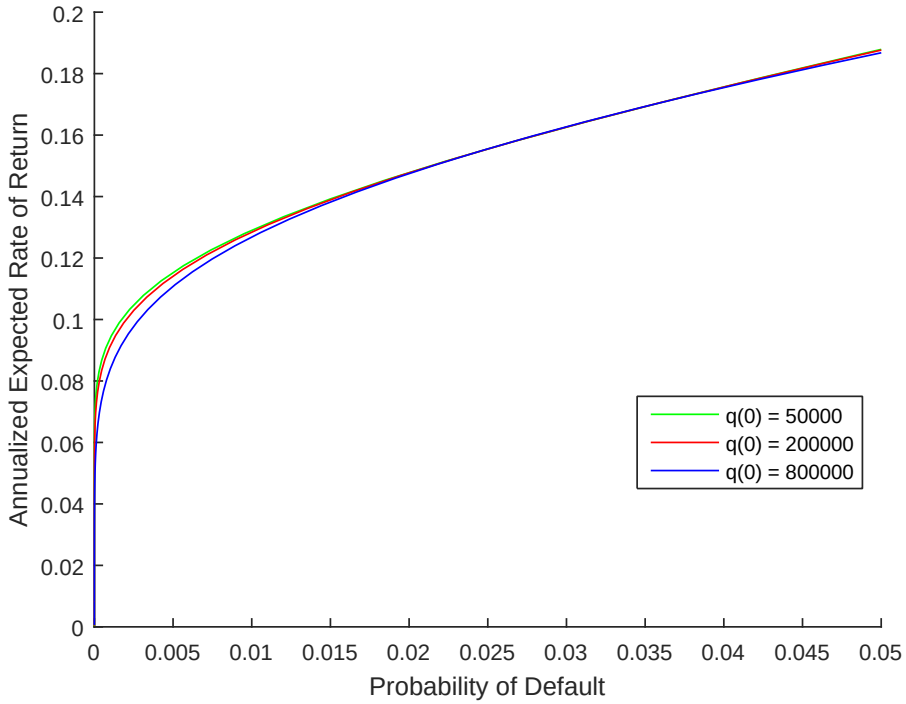
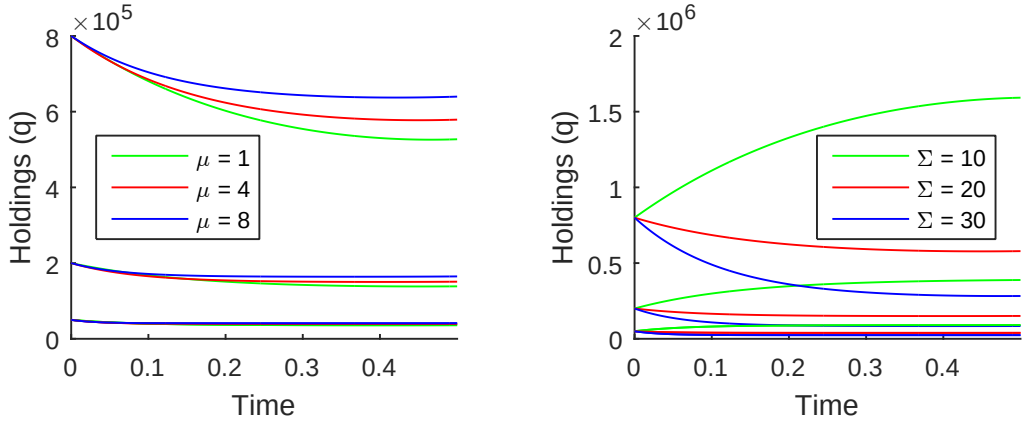


Figure 2.1: The efficient frontier for three firms with parameters given in Table 3.2, showing their default probability and expected rate of return on equity, when adopting their optimal portfolio with risk aversion parameters λ varying over $[0, \infty)$.

varied upwards (blue curve) and downwards (green curves).

One point needs to be reiterated: for each choice of a set of parameters excluding the risk aversion parameter λ, λ is computed to ensure that the firm's default probability (DP) is exactly 1%. Thus each curve in these figures corresponds to a different value of λ.

The effect on the optimal strategy of varying the asset rate of return μ and volatility Σ is shown in Figure 2.2. It is not a surprise to observe that the optimal strategy will include more of the risky asset as the rate of return is raised, or as the volatility is lowered. There is a threshold value of Σ below which the firm switches from sell strategies to buy strategies. Although not shown in the graph, one finds the reverse is the case for μ. Finally, the velocity of selling strategies seems to retain a similar shape over time under these variations. Each of these observations are borne out by more extensive investigations of the dependence on



(a) (b)

Figure 2.2: Effects on the benchmark optimal trading curve (red curve) for three firms as one parameter changes upwards (blue curve) and downwards (green curves). (a) shows the effect of changing μ, the mean rate of return of the risky asset; (b) shows the effect of changing Σ, the volatility of the risky asset. The vertical axis shows q, the amount of the risky asset being held at any time during the trading period. Both figures were computed using a trading period of half a year, while maintaining a probability of default of 1% for all trading curves.

these parameters.

In Figure 2.3a, the main effect of decreasing temporary impact Γ is seen to be to move more quickly to the final holding level early in the period. This can be understood as a change in the optimal balance between reducing temporary impact costs and price uncertainty due to the asset volatility. To a lesser extent, one also sees in these examples that the level of the final holdings decreases slightly as Γ increases.

The effect of permanent impact Λ on the strategy is more subtle. From Figure 2.3b, a higher permanent impact parameter Λ leads to an optimal strategy ending with a higher holding level. It also causes more curvature for the trading strategies, especially towards the closing time where all trading curves seem to have positive slope. Indeed, directly from (2.8), the general formula for the trading velocity, one verifies that at the close of the period $\frac{dq}{du}\,|_{u=T} = \frac{1}{2}\Gamma^{-1}\Lambda$. This means that as long as Λ is positive, every trader holding long positions, whether leveraging up or down, will always end the period by buying more shares. The reason is because permanent impact gives any trader a small opportunity to

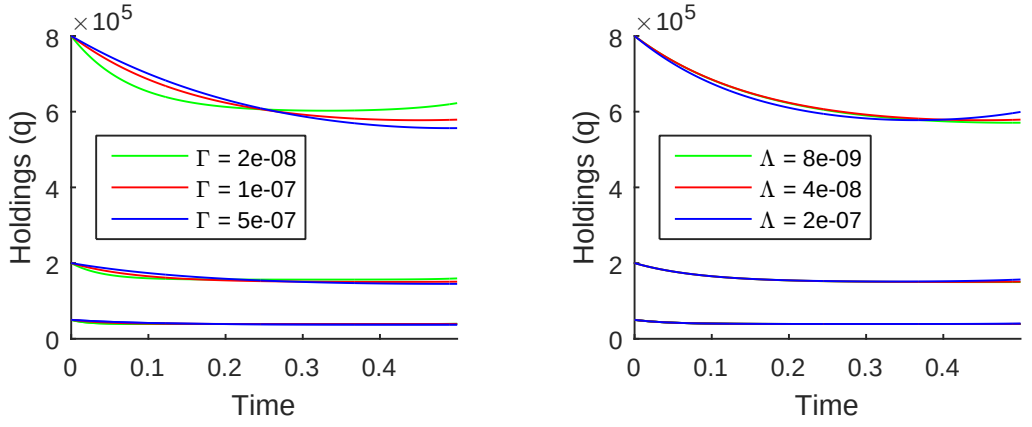


Figure 2.3: Effects on the benchmark optimal trading curve (red curve) for three firms as one parameter changes upwards (blue curve) and downwards (green curves). (a) shows the effect of changing the temporary impact parameter Γ; (b) shows the effect of changing the permanent impact parameter Λ. The vertical axis shows q, the amount of the risky asset being held at any time during the trading period. Both figures were computed using a trading period of half a year, while maintaining a probability of default of 1% for all trading curves.

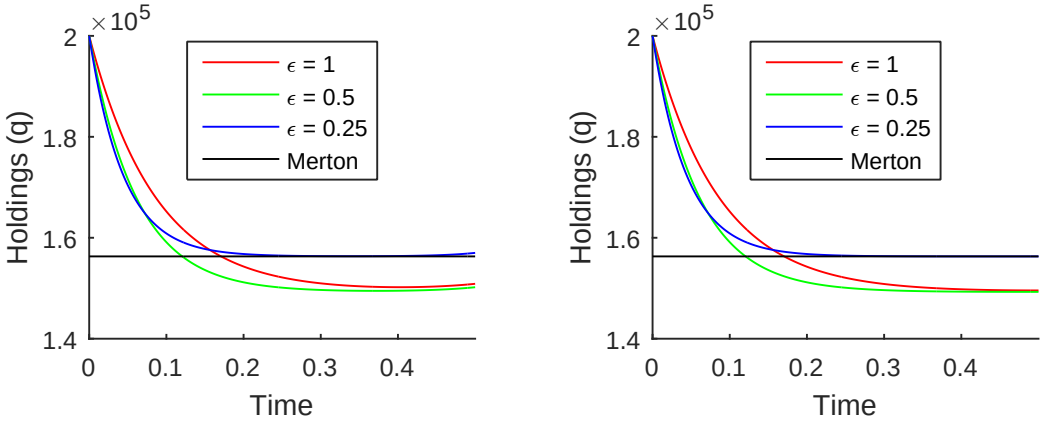
push the asset prices in a favourable direction at the last moment. We call this the *Ponzi property* of our market impact model: the gains it implies cannot be converted to cash without bursting the small price bubble the trader has created.

2.4.3 Small Market Impact

The perturbative analysis of Section 2.3.2 provides an alternative framework for understanding the effect of permanent and temporary market impact. We investigate the middle-sized firm with $q_0 = 2 \times 10^5$ and market impact parameters $\Gamma(\epsilon) = \epsilon \times 10^{-7}$, $\Lambda(\epsilon) = \epsilon \times 4 \times 10^{-8}$ for a sequence of values $\epsilon_n = 10^{-n}, n = 0, 1, \ldots$ approaching zero. Figure 2.4(a) shows how the optimal strategies converge for $\epsilon \to 0$ to the constant Merton solution for $u \in (t, T)$, but show rapid transient effects for u near both endpoints. The small Ponzi effect near $u = T$ can be turned off by taking $\Lambda(\epsilon) = 0$, as shown in Figure 2.4(b).

These figures suggest that for reasonable parameter values and small market impact,

our model will deliver strategies that are effectively similar to the Merton solution. The observed relationship between the optimal strategies and the Merton solution, valid for small market impact, actually remains true for intermediate levels of market impact such as our benchmark parametrizations. One observes in Figures 2.2 and 2.3 that all strategies tend to flatten as u approaches T, albeit with a small Ponzi effect at the end of the period. It will be well worth studying the extent that the value of the holdings at which the strategy flattens is well approximated by the Merton solution. As the market impact parameters decrease, the flat portion of the curve becomes wider, and closer to the Merton solution.

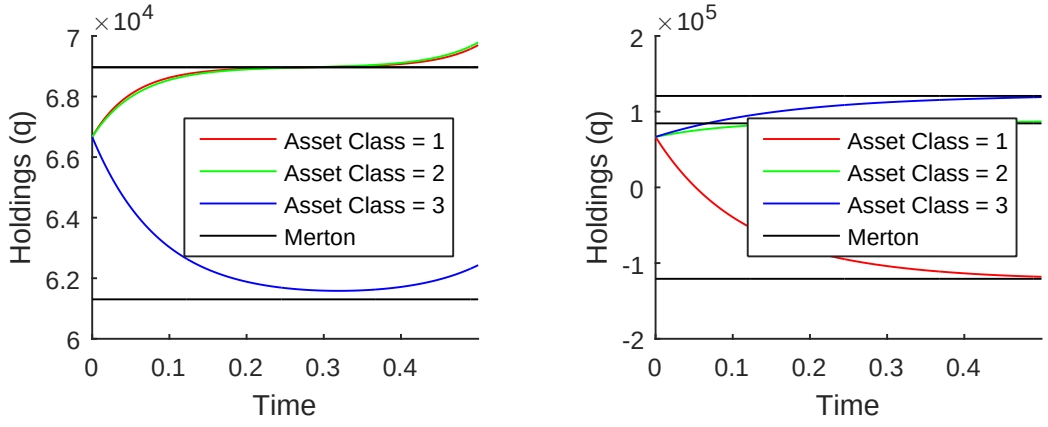


(a) $\Lambda = \epsilon\Lambda_1, \Gamma = \epsilon\Gamma_1$ (b) $\Lambda = 0, \Gamma = \epsilon\Gamma_1$

Figure 2.4: The behaviour of the optimal trading strategy for a decreasing sequence of market impact parameters as described in Section 2.4.3. They show convergence to the constant Merton solution.

An analysis similar to that of Section 2.3.2 allows us to understand the multi–asset investment problem in the small market impact regime. Figure 2.5, we used the standard asset parameters as Asset 2, with Asset 1 being the asset with the lower perturbed parameters from the previous cases, and Asset 3 having the higher perturbed parameters from the previous cases. Figure 2.5(a), compares the uncorrelated case to the Merton solution. Figure 2.5(b), compares the case of constant pairwise correlation $\rho = 0.5$ to the Merton solution. In both cases we can see the behaviour similar to the one asset case above. It should be noted that unlike the single asset case, a hedging strategy can be utilized for when multiple assets are available, hence short selling of a illiquid asset class can be optimal.

We observe again in the multi-asset problem that when the market impact is small,



(a) 3 Uncorrelated Assets

(b) 3 Positively Correlated Assets

Figure 2.5: The behaviour of the optimal three asset trading strategy in the uncorrelated and correlated cases, when compared to the Merton solution.

the general optimal strategy is close to the Merton solution.

2.4.4 Bounded Optimal Trading Strategies

We have seen in Section 2.4.2 that in the single asset case, positive Λ creates the Ponzi property that gives any trader an opportunity to push the price in their favour near the end of the period. Case 1 of Proposition 5 shows that as long as $\Lambda < \sqrt{2\lambda\Gamma}\Sigma$, the optimal strategies computed over any finite period $[0, T]$ remain bounded. However, when $\Lambda > \sqrt{2\lambda\Gamma}\Sigma$, Case 3 of Proposition 5 implies that for the period $[0, T^*]$ with $T^* = \tilde{K}/D$, the optimal strategy $q^*(u)$ and the value function W both blow up at $u = 0$.

Similar possibilities arise in the multi-asset investment problem. As Λ increases, eventually the matrix function $U(t)$ becomes singular for some finite $t = T^*$. Again, one then finds that for the period $[0, T^*]$, the optimal strategy $q^*(u)$ and the value function W both blow up at $u = 0$.

2.5 Remarks and Conclusions

The three hypothetical financial institutions studied in Section 2.4 face a typical investment problem, namely to maximize their return on equity subject to an upper bound on the downside risk, which is defined here as the probability of default. We have presented an analytically tractable version of the optimal portfolio problem that can be justified in three different ways: as utility optimization, as mean-variance optimization and as mean-default probability optimization. Numerical evidence shows that the solutions generated by the method have desirable and interesting features. Perhaps most importantly, we have learned that these strategies closely track the classic Merton solution arising in the zero market impact model.

We reiterate that while our optimal portfolio problem is closely related to the optimal liquidation problem, they are quite distinct. However, it is clear that following the Almgren-Chriss strategy to achieve a fixed target portfolio at time T will coincide with our strategy for $t < T$ if the portfolios coincide at T. Since our strategies come close to the Merton constant solution q^M, as long as the temporary and permanent impact is small enough and q_0 is close to q^M, our optimal strategies are typically well approximated by the Almgren-Chriss strategy that liquidates the amount $q_0 - q^M$ over the period $[0, T]$.

The three benchmark firms have efficient frontiers shown in Figure 2.1 that quantify by how much their rate of return will increase if they raise their tolerance to default. We have observed that optimal trading strategies that account for market impact tend to move over the trading period toward the Merton solution. If they are initially close to the Merton solution, they will tend to remain close, which means the Merton solution is robust to perturbations. The speed of approach increases as the temporary impact parameter Γ decreases. In addition, the main effect of the permanent impact Λ is the Ponzi property that is manifested by some amount of buying near the end of the period. This Ponzi effect is typically small, but as Proposition 5 shows, it will dominate the character of the solution when Λ becomes large enough to cause an asset price bubble.

Left to themselves, there is little incentive for such FIs to limit risk seeking. By choosing a low value of λ, or equivalently, accepting a high leverage ratio, they can achieve a high rate of return on capital. Since lower temporary impact and higher permanent impact

are both relatively more advantageous to larger firms, one has situations where large firms implement aggressive Ponzi style strategies. In scenarios where the assets perform badly, there is a likelihood of serious asset price feedback that may adversely affect other financial institutions holding common assets. Such asset price feedback, both bubbles and bursts, has been identified in the literature, notably Cifuentes et al. (2005), as a critical channel of systemic risk, popularly known as the *asset fire sale channel*. One application of our model, yet to be explored in detail, will be its use to specify the natural behaviour of the banks and financial institutions in a large financial system, and then to see how systemic risk measures are affected by asset fire sales due to market impact. In this systemic risk context, it will also be important to introduce the effects of funding illiquidity by modelling the stochastic nature of deposits.

If large banks were permitted to act in their own self interest without regard to their systemic effects, they would pose an unacceptable threat to financial stability. For that reason, all banks are subjected to a regime of strict financial regulation, of which the most important are limits to their capital asset ratio and liquidity coverage ratio. Under such regulatory constraints, FIs' investment strategies will differ dramatically from the optimal strategies produced in the present paper. The optimal behaviour of such regulated financial institutions will be the target of future modelling studies.

Appendix 2.A Proofs of Main Results

Proof of Theorem 1: In this proof we fix T to be finite. The existence of a maximal T^* is a consequence of solving (2.7), which is analyzed in the proof of Proposition 3. The Hamilton-Jacobi-Bellman (HJB) equation associated to (2.5) arises from the DPP by assuming Markov controls $v_t = v(t, T, X_t, q_t)$ and value function $W_t := W(t, T, X_t^v, q_t^v)$ for deterministic functions v, W. For simplicity of exposition, we have omitted the potential for dependence on the stock price S_t: the standard verification result used at the end of this argument shows this is consistent.

Under these assumptions, the DPP implies that $-e^{-\lambda W(t,T,X_t^v,q_t^v)}$ is a supermartingale

for all v and a martingale for the optimal v^*, which leads to the HJB equation for W

$$\begin{aligned}\partial_t W + \partial_X W q'\mu + \frac{1}{2} q'\Sigma\Sigma' q[\partial^2_{XX} W - \lambda(\partial_X W)^2] \\ + \sup_v[(\partial_q W' + \partial_X W q'\Lambda)v - v'\Gamma v \partial_X W] = 0. \\ W(T,T,X,q) = X.\end{aligned}$$

The ansatz $W(t,T,X,q) = X + V(t,T,q)$ leads to the equation for V

$$\begin{aligned}\partial_t V + q'\mu - \frac{\lambda}{2} q'\Sigma\Sigma' q + \sup_v[(\partial_q V' + q'\Lambda)v - v'\Gamma v] = 0. \\ V(T,T,q) = 0.\end{aligned}$$

The optimal feedback control is thus $v^* = \frac{\Gamma^{-1}}{2}(\partial_q V' + \Lambda q)$, which is independent of X and the price process, and hence deterministic. Using this control leads to

$$\partial_t V + q'\mu - \frac{\lambda}{2} q'\Sigma\Sigma' q + \frac{1}{4}(\partial_q V + \Lambda q)'\Gamma^{-1}(\partial_q V + \Lambda q) = 0 \ . \tag{2.54}$$

As we will shortly see in the proof of Proposition 3, this ODE has a unique smooth solution which is deterministic, over any finite time interval $[t,T]$ for T less than a possibly infinite maximal T^*. Therefore, by the classical verification theorem, we have $W = \tilde{W}$ and the other statements of the theorem follow.

□

Proof of Proposition 3: By Theorem 1 , the value function for Merton's problem over $[t,T]$ has the form $W(t,T,X,q) = X + V(t,T,q)$, where V satisfies the ODE (2.54). This ODE and the form (2.15) leads to Riccati equations with initial conditions for A, B, C

$$\partial_\tau A \quad -\frac{1}{4}(A + A' + \Lambda)\Gamma^{-1}(A + A' + \Lambda) + \frac{\lambda}{2}\Sigma\Sigma' = 0, \quad A(0) = 0 \tag{2.55}$$

$$\partial_\tau B \quad -\frac{1}{2} B\Gamma^{-1}(A + A' + \Lambda) - \mu' = 0, \quad B(0) = 0 \tag{2.56}$$

$$\partial_\tau C \quad -\frac{1}{4} B\Gamma^{-1}B' = 0, \quad C(0) = 0. \tag{2.57}$$

Notice that if A is a solution of (2.16), then so is A': By the uniqueness theorem for solutions of ODEs, $A = A'$ and therefore A is symmetric.

□

Proof of Theorem 4: Part 1: Note that $\frac{\lambda}{2}\Gamma^{-1/2}\Sigma\Sigma'\Gamma^{-1/2}$ is positive definite and define D to be its symmetric square root. If

$$\tilde{A}(\tau) := \Gamma^{-1/2}(A(\tau) + \Lambda/2)\Gamma^{-1/2}$$

then (2.16) becomes

$$\partial_\tau \tilde{A} - \tilde{A}^2 + D^2 = 0, \quad \tilde{A}(0) = E := \Gamma^{-1/2}(\Lambda/2)\Gamma^{-1/2} \ . \tag{2.58}$$

One can now check that the solution to (2.58) has the form $\tilde{A} = VU^{-1}$, where U, V satisfy the following linearODE with terminal condition

$$\left[\begin{array}{c} \partial_\tau U \\ \partial_\tau V \end{array} \right] = \left[\begin{array}{cc} 0 & -\mathbb{1} \\ -D^2 & 0 \end{array} \right] \times \left[\begin{array}{c} U \\ V \end{array} \right], \quad \left[\begin{array}{c} U(0) \\ V(0) \end{array} \right] = \left[\begin{array}{c} \mathbb{1} \\ E \end{array} \right] \ .$$

By block-diagonalization using

$$Q = \left[\begin{array}{cc} \mathbb{1} & \mathbb{1} \\ D & -D \end{array} \right], \quad Q^{-1} = \frac{1}{2} \left[\begin{array}{cc} \mathbb{1} & D^{-1} \\ \mathbb{1} & -D^{-1} \end{array} \right]$$

one finds

$$\left[\begin{array}{cc} 0 & -\mathbb{1} \\ -D^2 & 0 \end{array} \right] = Q \left[\begin{array}{cc} -D & 0 \\ 0 & D \end{array} \right] Q^{-1}$$

and therefore, the solution of the matrix ODE is

$$\left[\begin{array}{c} U(\tau) \\ V(\tau) \end{array} \right] = Q \left[\begin{array}{cc} e^{-D\tau} & 0 \\ 0 & e^{D\tau} \end{array} \right] Q^{-1} \times \left[\begin{array}{c} \mathbb{1} \\ E \end{array} \right] \ .$$

From the explicit forms

$$U(\tau) \quad = \ \cosh(D\tau) - \ \sinh(D\tau) D^{-1} E \tag{2.59}$$

$$V(\tau) \quad = - \sinh(D\tau) D + \ \cosh(D\tau) E \tag{2.60}$$

one finds $A(\tau) = \Gamma^{1/2}(\tilde{A}(\tau) - E)\Gamma^{1/2}$ where

$$\tilde{A}(\tau) = [- \sinh(D\tau) D + \ \cosh(D\tau) E][\cosh(D\tau) - \ \sinh(D\tau) D^{-1} E]^{-1}. \tag{2.61}$$

The Riccati equation (2.17) for B can be solved by noting that $\tilde{B} = B\Gamma^{-1/2}$ solves the ODE

$$\partial_\tau \tilde{B} - \tilde{B}\tilde{A} - \mu'\Gamma^{-1/2} = 0.$$

Since $\partial_\tau U = -\tilde{A}U$, we find $\partial_\tau(\tilde{B}U) = (\partial_\tau \tilde{B} - \tilde{B}\tilde{A})U = \mu'\Gamma^{-1/2}U$ which can be integrated to give $\tilde{B}(\tau)U(\tau) = \mu'\Gamma^{-1/2}(\int_0^\tau U(s)ds)$ and thus

$$B(\tau) = \mu'\Gamma^{-1/2}\left(\int_0^\tau U(s)ds\right)U^{-1}(\tau)\Gamma^{1/2} \ .$$

It is straightforward that $\int_0^\tau U(s)ds = D^{-2}[E - V(\tau)]$ which gives the desired formula

$$B(\tau) = \mu'\Gamma^{-1/2}D^{-2}[E - V(\tau)]U^{-1}(\tau)\Gamma^{1/2} \ .$$

In a similar fashion, one finds

$$C(\tau) = \frac{1}{4}\int_0^\tau B(s)\Gamma^{-1}B'(s)ds \tag{2.62}$$

$$= \frac{1}{4}\overline{\mu}'\left(\int_0^\tau (E - V(s))(U'(s)U(s))^{-1}(E - V'(s))ds\right)\overline{\mu}, \tag{2.63}$$

where $\overline{\mu} := D^{-2}\Gamma^{-1/2}\mu$.

Part 2: This part is straightforward.

Part 3: From part 4 of Theorem 1, the optimal control $q^*(u)$ over the period $[t, T]$ solves

$$\partial_u q - \Gamma^{-1}(A(T-u) + \Lambda/2)q = \frac{1}{2}\Gamma^{-1}B'(T-u)$$

When this linear ODE is multiplied on the left by the integrating factor $U^{-1}(T-u)\Gamma^{1/2}$, the left-hand side becomes an exact derivative:

$$\partial_u\left[U^{-1}(T-u)\Gamma^{1/2}q\right] = U^{-1}(T-u)\Gamma^{1/2} \times \frac{1}{2}\Gamma^{-1}B'(T-u) \ .$$

Integration of this equation over $[t, u]$ gives

$$U^{-1}(T-u)\Gamma^{1/2}q(u) - U^{-1}(T-t)\Gamma^{1/2}q = \frac{1}{2}\int_t^u U^{-1}(T-r)\Gamma^{-1/2}B'(T-r)dr$$

which leads to the desired formula.

Part 4: The Variance is calculated directly as follows

$$\mathrm{Var}_t(X_T^*) = \int_t^T q^*(s)'\Sigma\Sigma' q(s)^* ds = q' L(T-t) q + M(T-t) q + N(T-t) \ .$$

Rewrite $q^*(u) = \tilde{U}(T-u)\tilde{U}^{-1}(T-t)q + \frac{1}{2}\tilde{U}(T-u)I(u)$, where $\tilde{U}(T-u) := \Gamma^{-1/2}U(T-u)$ and $I(u) := \int_t^u \tilde{U}^{-1}(T-r)\Gamma^{-1}B(T-r)dr$. Explicit forms for L, M, N are calculated as follows.

$$L(T-t) = (\tilde{U}^{-1}(T-t)')\Big(\int_t^T \tilde{U}(T-r)'\Sigma\Sigma'\tilde{U}(T-r)dr\Big)\tilde{U}^{-1}(T-u) \ .$$

By using Fubini's formula, we have

$$\begin{aligned} M'(T-t) \quad &= \int_t^T \tilde{U}^{-1}(T-t)'\tilde{U}(T-r)'\Sigma\Sigma'\tilde{U}(T-r)I(r)dr \\ &= \int_t^T \Big(\int_s^T \tilde{U}^{-1}(T-t)'\tilde{U}(T-r)'\Sigma\Sigma'\tilde{U}(T-r)dr\Big)\tilde{U}(T-s)^{-1}\Gamma^{-1}B(T-s)ds \\ &= \tilde{U}(T-t)'^{-1}\int_t^T \tilde{U}(T-s)L(T-s)\Gamma^{-1}B(T-s)ds \ . \end{aligned}$$

Similarly

$$\begin{aligned} N(T-t) \quad &= \frac{1}{4}\int_t^T I(r)'\tilde{U}(T-r)'\Sigma\Sigma'\tilde{U}(T-r)I(r)dr \\ &= \frac{1}{4}\int_t^T \Big(\int_s^T I(r)'\tilde{U}(T-r)'\Sigma\Sigma'\tilde{U}(T-r)dr\Big)\tilde{U}^{-1}(T-s)\Gamma^{-1}B(T-s)ds \\ &= \frac{1}{4}\int_t^T M(T-s)\Gamma^{-1}B(T-s)ds \ . \end{aligned}$$

□

Chapter 3

Asset Fire Sales and Systemic Risk

"Asset Fire Sales and Systemic Risk" is a joint work with Dr. Tuan Tran and Prof. Thomas Hurd. Contributions by the author include: 1. computer programming, numerical analysis, and finalization of the paper; 2. Equal share of the original idea of this paper which evolved from our discussions; 3. Equal share of the technical derivations of this paper; 4. Equal share of drafting. As of October 31, 2016, this paper is currently being finalized for journal submission.

3.1 Asset Fire Sales and Systemic Risk

Recent research in systemic risk, notably Cifuentes et al. (2005), has recognized that *market illiquidity*, manifested in so-called *asset fire sales*, is a critically important source of financial instability. At a time of crisis, banks may be weakened to the extent that they need to shrink their balance sheets by liquidating assets. This can lead to strong downward price spirals that further impact the entire financial system. To understand this better, a number of recent contributions, for example Caccioli et al. (2014), have proposed network models with the fire sale contagion propagated indirectly from bank to bank through their trading impact on the prices of assets they both own. Such models adapt the methodology of Eisenberg and Noe (2001) that describes the end result of a stylized financial crisis as a fixed point of a cascade mapping on the collection of balance sheets. From one point of

view, such a crisis proceeds through stepwise iteration of the cascade mapping that distills how banks update their balance sheets each step, ultimately approaching a fixed point or cascade equilibrium.

As reviewed in Chapter 2 of Hurd (2016), such cascade models are called *static*, meaning they make an oversimplification that no uncertainty occurs from time step to time step. They also propose that each FI has a simple behaviour at each step that depends deterministically on the state of their balance sheet, that at a moment in time has a stylized form that might look like Table 4.2. It is assumed that during a crisis, their behaviour is mostly self-protection, and should account for the need to stay solvent, to avoid deposit runs, and avoid having to sell illiquid assets in a collapsing market. In fact, given the regulatory regime that banks now face, one can assume they act in the simplest way they can to remain compliant with important regulatory constraints.

Assets	illiquid assets A^F	cash C	
Liabilities	long term debt D^ℓ	short term debt D^s	equity E

Table 3.1: Stylized balance sheet:

Basel III, the current regulatory regime for banks worldwide, differs most essentially from earlier regulatory regimes by recognizing the need to go beyond controlling microprudential risk (the risk of an individual bank failure), to trying to account for macroprudential risks (the risk of system wide failures). However, in most existing models of systemic risk, banks' assumed behaviour does not explicitly anticipate the actions of other banks. Such models can therefore give only a very incomplete picture of macroprudential risks in the system. The goal of the present paper is to improve understanding of how a regulated bank may behave at any time during a fire sale crisis when it anticipates the actions of other banks. In particular, we propose that each FI adopts strategies that during a financial crisis will anticipate the potential for large scale deleveraging by all other FIs. With this in mind, in addition to assuming the usual accounting principles, we make additional simplifying modelling assumptions:

A1 There is a single common illiquid asset of which banks always hold non-negative amounts, and the banks' trades impact the market asset price;

A2 Banks have perfect information about the balance sheets and hence strategies of all other banks;

A3 Banks' balance sheets are subject to a set of one or more regulatory constraints, and banks must remain compliant;

A4 Banks trade each period to maximize their one-period return, subject to the regulatory constraints at the end of the period;

A5 Banks that cannot become compliant liquidate 100% of their illiquid assets and can then be removed from the financial system;

A6 Small traders that are price takers and non-strategic clear the asset market each period.

In addition to these economic and behavioural assumptions, Section 3.2 will also introduce additional assumptions for mathematical convenience. Like other static cascade models that are step-wise deterministic, we assume a simple deterministic asset price dynamics. We propose a simple market clearing mechanism each period. We also focus on two specific examples of regulatory constraints: the leverage ratio (LR) constraint and the liquidity coverage ratio (LCR) constraint.

Some of these assumptions require additional justification and discussion. First, we reiterate that the assumed balance sheets omit all interbank assets that are commonly the focus of systemic risk networks: extending our model in this direction will be an interesting future research project. The assumption of complete information is obviously counterfactual, but is typical in game theory. We will comment on its implications in the conclusions of the paper. A final point is that the idea that the crisis proceeds stepwise, even if commonplace in the systemic risk literature, still begs the question of the time scale involved: The picture we have is that the time step is short relative to one month, that is to say, less than one week.

As we have just hinted, we will find that at any cascade step, determining the banks' collective actions amounts to a classic example of a multi-player game (for a review of game theory in economics, see Facchinei and Kanzow (2007)) whose solutions can be comprehensively analyzed. We will find there will be a finite collection of Nash equilibrium solutions, amongst which is a unique Pareto optimal solution. In a certain sense, only

the Pareto optimal solution should have economic meaning. Such Pareto solutions correspond to bank strategies that are easy to compute, and can be taken as the natural bank behaviour assumption that determines the cascade mapping at the core of new network models of systemic risk. There is scope in the resultant network models for banks to exhibit *tipping point* or *threshold* behaviour because, as we shall show, the Pareto optimal solution is generally discontinuous in the underlying parameter space of the model.

The general modelling framework formulated in Section 2 is analyzed in Section 3, leading to the main technical result of the paper, Theorem 8, that characterizes all Nash equilibria in terms of the final asset price determined by the market clearing condition. In Section 4 the paper also investigates more detailed properties of the model system under the liquidity coverage ratio constraint, and in particular, we find an efficient explicit algorithm to compute all Nash equilibriums. Section 5 explores certain numerical features of the model. First, we look more closely at the nature of the tipping points that arise in the case of $N = 2$ banks. Then we move on to larger networks, and investigate the impact on network stability of banks that are in danger of non-compliance. Finally, the conclusions in Section 6 discuss in more detail how these results change our view of the fire sale contagion mechanism.

3.2 The multi-agent model

At a moment in time, say $t = 0$ or $t = 1$, the financial system under study will be thought of as a collection of N banks, with balance sheets $\mathscr{B}_t$, that are large strategic traders subject to certain regulatory constraints. The collection of initial bank balance sheets $\mathscr{B}_0 := \{B_0^i\}, i \in \{1, \ldots, N\} := [N]$, as shown in Table 4.2, are composed of fixed assets A_0^i, cash C_0^i and debt $D_0^i = D_0^{s,i} + D^{\ell,i}$ separated into short-term and long-term parts. Their equity $E_0^i := A_0^i + C_0^i - D_0^i$, assets minus liabilities, is positive unless the bank is *insolvent*. We also assume banks each hold q_0^i units of a single illiquid asset class, whose unit price, S_0, is assumed to remain constant until an auction at time $t = 1$. Then the values of fixed assets are $A_0^i = q_0^i S_0$. All this information about the initial state of bank i is recorded in the vector

$$B_0^i = (A_0^i, C_0^i, D_0^{s,i}, D_0^{\ell,i}, E_0^i, q_0^i).$$

Note in particular that our model does not permit interbank crossholdings.

Each bank is subject to a *regulatory constraint* $H^i(B_t^i) \geq 0$ that is checked on certain dates. Usually the *regulatory function* H^i is linear in the state variables B_t^i: Three examples are given toward the end of this section. In our one-period setting, since $t = 1$ will be the only monitoring date considered, the constraint will be checked on B_1^i.

The external economy will be described by a large collection of *small traders* that are non-strategic price takers: their combined market effect is to purchase Y shares determined from the price S by an *aggregate demand function* $Y = D(S)$. This means that over the period $(0, 1)$, if each large trader trades $|x^i|$ units, a buy order if $x^i > 0$ or a sell order if $x^i < 0$, the new price S_1 will be determined by the aggregate demand supplied by the banking system $X := \sum_{i\in[N]} x^i$ through the *market clearing condition* $X + Y = 0$, which implies

$$S_1 = D^{-1}(-X) \ . \tag{3.1}$$

The goal of this model is to determine the natural strategic trading behaviour of such regulated banks over the single period from $t = 0$ to $t = 1$, represented by $x^i = q_1^i - q_0^i$, the number of units they trade, and the final asset price S_1 that results from these trades.

To have a well-posed and tractable problem, we make final assumptions, in addition to those outlined in Section 3.1:

B1 The external demand function is linear, $D(S) = \frac{1}{\lambda}(S_0 - S)$, with λ called the *market illiquidity parameter*. Hence the market clearing condition (3.2) at time $t = 1$ implies

$$S_1 = S_0 + \lambda X \tag{3.2}$$

B2 Banks pay proportional transaction costs $\gamma|x|S_1$ with parameter $\gamma \in [0, 1]$, for a buy or sell trade of size $|x|$.

B3 No borrowing or short selling is permitted, no interest is earned on cash, and deposits do not change over the period. The external market is deep enough to absorb the total supply of illiquid assets, that is:

$$S_{\min} := S_0 - \lambda \sum_{i=1}^{N} q_0^i \geq 0. \tag{3.3}$$

B4 Each bank must satisfy its regulatory constraint $H^i = H^i(B_1^i) \geq 0$ at $t = 1$, if possible. A bank that at time $t = 0$ knows it cannot satisfy its regulatory constraint at $t = 1$ must liquidate all its assets in that period, so that $q_1^i = 0$.

Based on these Assumptions A and B, when the banking system trades the amounts $x = (x^i)_{i\in[N]}$, the final asset price will be $S_1 = S_0 + \lambda X$ with $X = \sum_i x^i$ and the balance sheet at time $t = 1$ of any bank i will be

$$\begin{aligned} A_1^i &= (q_0^i + x^i)S_1, \quad C_1^i = C_0^i - x^i S_1 - \gamma|x^i|S_1, && (3.4)\\ D_1^{s,i} &= D_0^{s,i}, \quad D_1^{l,i} = D_0^{l,i}, \quad E_1^i = E_0^i + q_0^i(S_1 - S_0) - \gamma|x^i|S_1 && (3.5) \end{aligned}$$

Suppose the bank i believes that the other banks will trade x^{-i} units in the period. Notice from (3.4) and (3.5) that this information determines $B_1^i = B_1^i(x^i, x^{-i})$ as a function of the trades (x^i, x^{-i}) through the final price $S_1 = S_1(x^i, x^{-i}) = S_0 + \lambda(x^i + \sum_{j\neq i} x^j)$. The regulatory functions H^i can thus be rewritten as a function of $(x^i, x^{-i}) : H^i = H^i(x^i, x^{-i})$. Under this condition, we summarize the above assumptions by proposing that each bank chooses to trade an amount x^i determined by solving the *single period optimization problem*:

$$\begin{aligned} &\textbf{if} \qquad \mathscr{A}^i(x^{-i}) := \{x_i \geq -q_0^i, C_1^i(x_i, x^{-i}) \geq 0, H^i(x_i, x^{-i}) \geq 0\} \neq \emptyset \\ &\textbf{then} \quad x^i \in \arg\max_{x^i\in\mathscr{A}^i(x^{-i})} E^i(x^i, x^{-i}) \\ &\textbf{else} \qquad x^i = -q_0^i \end{aligned} \tag{3.6}$$

This is a *rational strategy* for each bank i in that it maximizes its equity conditioned on knowing the other banks' actions, as long as it can fulfil its regulatory constraints. In the case where bank i cannot fulfil its regulatory constraint, it then has to liquidate all illiquid assets and declare bankruptcy. The scheme can be seen to be equivalent to a standard game theory problem by requiring that player i must trade $x^i \geq -q_0^i$ and introducing an infinite penalty term for violating its constraints. That is, each player i trades $x^i \geq -q_0^i$ from which it receives a singular payoff:

$$\overline{E}^i := E_1^i - \infty \cdot \mathbf{1}\left(x^i > -q_0^i, \min\left(C_1^i, H^i\right) < 0\right) \ . \tag{3.7}$$

The rational strategy of the bank i, x^i now is defined by

$$x^i \in \arg\max_{x_i \geq -q_0^i} \overline{E}^i(x^i, x^{-i}) \tag{3.8}$$

A player that knows x^{-i} and tries to maximize the payoff (3.7) will always buy the amount determined by (3.6). This game-theoretic representation of the problem justifies adopting the concept of Nash Equilibrium:

Definition 9. *A Nash Equilibrium strategy (NE) $x^* = (x^{i*})_{i\in[N]}$ is any vector such that for each i, x^{i*} solves the single period optimization problem* (3.6) *for bank i when the remaining banks trade the amounts x^{-i*}.*

This paper will focus on stylized regulatory constraints that distill essential aspects of the Basel III framework. To formulate them, it is helpful to introduce two additional *buffer variables*: $\Delta^{i,s} := C^i - D^{i,s}$ and $\Delta^i := C^i - D^i$.

1. **Liquidity coverage ratio (LCR) requirement:**

   $$\text{cash} \geq \text{short term debt} .$$

   This means the bank is compliant at time $t = 1$ if $\Delta_1^{i,s} \geq 0$, that is

   $$H^{i,L}(x^i, x^{-i}) := \Delta_0^{i,s} - (x^i + \gamma|x^i|)S_1 \geq 0 . \tag{3.9}$$

   Recall that the price S_1 is also a function of trading strategies: $S_1 = S_1(x^i, x^{-i}) = S_0 + \lambda x^i + \lambda \sum_{j\neq i} x^j$. In this context, *short term debt* has the meaning of debt that might mature or run during a 30-day crisis scenario.

2. **Capital Adequacy Ratio (CAR) requirement:**

   $$\text{Capital} \geq r \times \text{Risk Weighted Assets , for some fraction } r \in [0, 1] .$$

   In Basel III accord, the constant r is 8%. In this paper, capital is identical with equity. As for risk-weighted assets, we suppose that the weight of cash is zero, while the weight of risky asset is one. The CAR requirement thus becomes the leverage requirement.

$$\text{Equity} \geq r \times \text{Fixed Assets , for some fraction } r \in [0,1] \ .$$

Note that cash is not included on the right hand side means that having a positive amount of cash should not be penalized by the regulator because it can always be used to pay down debt without changing equity. Thus the bank is compliant if $E^i \geq rA^i$, or equivalently

$$H^{i,K}(x^i, x^{-i}) := [(1-r)q_0^i - rx^i - \gamma|x^i|]S_1 + \Delta_0^i \geq 0. \tag{3.10}$$

3. **Mixed liquidity and capital (MLC) requirements:** The mixed condition

$$\text{cash} \geq \text{short term debt} \ ; \quad \text{equity} \geq r \times \text{fixed assets}$$

means the bank is compliant if

$$H^{i,M} = \min\{H^{i,L}, H^{i,K}\} \geq 0 \ . \tag{3.11}$$

For each of the regulatory constraints LCR, CAR and MLC, we have similar results:

Theorem 6. *Suppose that the regulatory functions take one of three forms (3.9), (3.10) or (3.11). Then, there exists a vector-valued function $F = (F_i) : \mathbb{R}^N \to \mathbb{R}^N$, where F_i depends on x^{-i} only, i.e. $F_i(x^i, x^{-i}) = F_i(x^{-i})$ such that x^* is a Nash Equilibrium solution of problem (3.7) if and only if x^* solves the fixed-point equation $x^* = F(x^*)$.*

Proof of Theorem 6. We fix i and x^{-i} and prove that the optimization problem of bank i has a unique solution $x^i = F^i(x^{-i})$. The proof works for a more general class of regulatory functions $H(x) := H^i(x, x^{-i})$ that satisfy the condition:

Condition on H: For each i, the set $\mathscr{H}_+ := \{x \geq 0 : \ H(x) \geq 0\}$ is a connected interval.

It is straightforward to verify that the Liquidity Coverage Ratio (3.9), Leverage Ratio (3.10) and a mixture of both (3.11) all satisfy this condition. We also note that $\mathscr{C} := \{x \geq -q_0^i : \ C_1^i(x, x^{-i}) \geq 0\}$ is a connected interval.

Now note that the equity of bank i is a piecewise quadratic in x:

$$E(x) := E_1^i(x, x^{-i}) = \begin{cases} \lambda\gamma x^2 + (\lambda q_0^i + \gamma S_-)x + q_0^i S_- + \Delta_i. & \text{if } x \leq 0 \\ -\lambda\gamma x^2 + (\lambda q_0^i - \gamma S_-)x + q_0^i S_- + \Delta_i & \text{if } x \geq 0 \end{cases} . \tag{3.12}$$

where $S_- := S_{-i} := S_0 + \lambda \sum_{j \neq i} x^j$.

Now suppose that $\mathscr{H} \cap \mathscr{C} \cap [0, \infty) \neq \emptyset$. Then it is either suboptimal (if $0 \in \mathscr{H} \cap \mathscr{C}$) or inadmissible for the bank to sell because $E(x) < E(0)\ \forall x < 0$. Since $E(x)$ is strictly concave on $[0, \infty)$ and $\mathscr{H} \cap \mathscr{C}$ is non-empty and connected, $E(x)$ has a unique maximum on $\mathscr{H} \cap \mathscr{C}$.

Next, suppose $\mathscr{H} \cap \mathscr{C} \cap [0, \infty) = \emptyset$. Notice that, with $\gamma \in [0, 1]$ the derivative of the equity function for $x \in [-q_0^i, 0]$ is positive because

$$E'(x) = 2\gamma\lambda x + \lambda q_0^i + \gamma S_- \geq \gamma\lambda(x + q_0^i) + \gamma S_{\min} > 0.$$

Therefore, $E(x)$ is increasing on $[-q_0^i, 0]$. Thus the unique maximum is the supremum of the set $\mathscr{H} \cap \mathscr{C} \cap [-q_0, 0)$.

Finally, if $\mathscr{H} \cap \mathscr{C}$ is empty, then the unique solution is given by $x^i = -q_0^i$. □

We shall postpone computation of explicit strategies to the next section, where it is shown that the equilibrium points of the game can be fully determined in terms of their corresponding equilibrium prices.

An interesting feature of the Nash equilibrium solution is *Pareto-optimality*. This concept is defined as follows

Definition 10. *Let $\mathscr{D}$ be a subset of $\Pi_{i=1}^N [-q_0^i, \infty)$. A trading strategy of the system $\hat{x} \in \mathscr{D}$ is called Pareto-optimal if there exists no other strategy $x \in \mathscr{D}$ that is better for all banks. More precisely, let $\theta_i(x) \in \{0, 1\}$ be the survival indicator of the bank i associated to the strategy x, i.e. $\theta_i = 0$ if i defaults and $\theta_i = 1$ otherwise. The solution $\hat{x} := (\hat{x}^i)_{i=1,\ldots,N}$ is Pareto-optimal if there exists no other solution $x := (x^i)_{i=1,\ldots,N}$ such that $E_i(x) \geq E_i(\hat{x}), \theta_i(x) \geq \hat{\theta}_i(\hat{x}) \forall i$ and strict inequalities happen for some i.*

In the sequel, the set $\mathscr{D}$ of interest will belong to one of the two cases:

1. Case 1: $\mathscr{D} = \mathscr{D}^{NE}$ is the set of all possible NE strategies.

2. Case 2: $\mathscr{D} = \mathscr{D}^{ad}$ is the set of all admissible trading strategies, i.e.

$$\mathscr{D}^{ad} = \{x \in \Pi_{i=1}^{N}[-q_0^i, \infty) : \overline{E}^i(x) > -\infty\}. \tag{3.13}$$

3.3 The Aggregate Response Function

Given a constraint H that is one of (3.9), (3.10), or (3.11), assume that there exists a Nash Equilibrium strategy $x^* = \{x^{i*}\}_{i=1,\dots,N}$ that satisfies (3.6). We then define

$$S^* := S_0 + \lambda \sum_{i=1}^{N} x^{i*} \tag{3.14}$$

to be the *equilibrium price* associated with the strategy vector x^*. The proof of Theorem 6 shows that for each i, $x^{i*} = F_i(x^{-i*})$ depends on x^{-i*} only through $S^*_{-i} := S_0 + \lambda \sum_{j \neq i} x^{j*} = S^* - \lambda x^{i*}$. Thus, $x^{i*} = \overline{F}_i(S^*_{-i}))$ for some scalar function $\overline{F}_i$, and hence the optimal strategy x^{i*} for bank i is linked to the equilibrium price S^* via the equation $x^{i*} = \overline{F}_i(S^* - \lambda x^{i*})$. This is an implicit *relation* for x^{i*} in terms of S^* that may have multiple solutions. The following lemma, easy to prove, provides a necessary and sufficient condition under which such a relation can be solved *uniquely* by $x^{i*} = G_i(S^*)$:

Lemma 7. *The equation $x^{i*} = \overline{F}_i(S^* - \lambda x^{i*})$ has unique solution $x^{i*} = G_i(S^*)$ if and only if the function $s \to s + \lambda \overline{F}_i(s)$ is one-to one. Furthermore, G_i is defined by*

$$G_i(S^*) = \overline{F}_i((Id + \lambda \overline{F}_i)^{-1})(S^*). \tag{3.15}$$

This implies that if an equilibrium price S^* is known to all banks for some reason (for example, if banks can rationally predict this price after taking into account trading strategies of all other banks), then each bank is able to come up with an optimal trading strategy that depends only on S^*. This realization is sufficient to prove the following theorem that reduces the study of Nash equilibriums to the study of fixed points of a scalar equation.

Theorem 8. *Suppose that the regulatory functions take one of three forms (3.9), (3.10) or (3.11). There exist N response functions $(G_i)_{i=1\dots N}$, given explicitly by* (3.20), (3.21), *and* (3.22), *that satisfy the following properties.*

1. *If $(x^i)_{i=1\ldots N}$ is a vector of trading strategies such that bank i solves its own optimization problem (3.6), then $x^i = G_i(S)$, where $S = S_0 + \lambda \sum_{i=1}^N x^i$.*

2. *The Nash Equilibrium strategy x^{i*} of bank i is related to the equilibrium price S^* by $x^{i*} = G_i(S^*)$. Consequently, there is a one to one mapping between the set of Nash Equilibrium points (x^{i*}) and the set of equilibrium prices S^*.*

3. *The Nash Equilibrium price S^* is a solution of the fixed point equation*

$$S = S_0 + \lambda G(S) \tag{3.16}$$

where $G(S) := \sum_{i=1}^N G_i(S)$ is called the aggregate response function*.*

The existence and properties of the individual response functions $G_i(s)$ follow from Lemma 7, we will need explicit formulas for the different constraint cases. We present the argument for the LR case (3.10), and leave the other two similar cases for the reader.

For i fixed, and the final price $S_1 = S$, we determine the optimal response $x = G(S)$ of bank i. The essential point in the computation is that when the constraint $H \geq 0$ can be achieved, the function $E(x)$ must always be maximized by varying x while keeping $S_- := S - \lambda x$ fixed.

First, note that $x = G(S) \geq 0$ if and only if $S \geq -\frac{\Delta}{(1-r)q}$: If $S \geq -\frac{\Delta}{(1-r)q}$ then the bank will be compliant with $x = 0$ and this is a better strategy than any sell strategy $x < 0$ because $E(x)$ is increasing on $[-q, 0]$ (selling reduces equity due to transaction costs); conversely, if $S < -\frac{\Delta}{(1-r)q}$ then one can check that only sell strategies have the possibility to yield $H(x) \geq 0$.

Next, we assume $S \geq -\frac{\Delta}{(1-r)q}$ and $x = G_i(S) \geq 0$. Maximizing the quadratic $E^i(x)$ given by (3.12) over $[0, \infty)$ leads to

$$x_E := \left(\frac{q}{2\gamma} - \frac{S_-}{2\lambda}\right)^+ = \left(\frac{q}{\gamma} - \frac{S}{\lambda}\right)^+ , \tag{3.17}$$

which is an upper bound on $G_i(S)$. Since the bank i is not allowed to borrow money, $C \geq 0$, which implies

$$G_i(S) \leq x_C := \frac{C}{(1+\gamma)S}. \tag{3.18}$$

Finally, the bank needs $H^K \geq 0$, hence

$$G(S) \leq x_H := \frac{(1-r)q}{(r+\gamma)} + \frac{\Delta}{(r+\gamma)S}. \tag{3.19}$$

From these three inequalities, we conclude that $G(S) = \min\{x_E, x_C, x_H\}$ when $S \geq -\frac{\Delta}{(1-r)q}$.

Finally, we assume the complement: $S < -\frac{\Delta}{(1-r)q}$ and $x = G_i(S) < 0$. Then, the inequality $H^K \geq 0$ becomes

$$(\gamma - r)xS + (1-r)qS + \Delta \geq 0.$$

If $r \leq \gamma$, there is no $x < 0$ and S that satisfies this requirement, therefore $x = G(S) = -q$. On the other hand, if $r > \gamma$, the bank sells the minimum amount to achieve compliance $H^K = 0$, or if it cannot comply, sells $-q$, which leads to

$$G(S) = \max\left\{-q, \frac{(1-r)q}{(r-\gamma)} + \frac{\Delta}{(r-\gamma)S}\right\} \quad \text{when } r > \gamma \ .$$

This argument for the LR case and analogous arguments for the LCR and MLC cases yield the desired formulas for the response functions $G_i(S)$:

1. **Liquidity coverage ratio:**

$$G_i(S) = \begin{cases} \min\{(\frac{q_0^i}{\gamma} - \frac{S}{\lambda})^+, \frac{\Delta_0^{i,s}}{(1+\gamma)S}\} & \text{if } \Delta_0^{i,s} \geq 0 \\ \min\{-q_0^i, \frac{\Delta_0^{i,s}}{(1-\gamma)S}\} & \text{if } \Delta_0^{i,s} < 0 \end{cases} \tag{3.20}$$

2. **Capital Adequacy Ratio:**

$$G_i(S) = \begin{cases} \min\{(\frac{q_0^i}{\gamma} - \frac{S}{\lambda})^+, \frac{C_0^i}{(1+\gamma)S}, \frac{(1-r)q_0^i}{(r+\gamma)} + \frac{\Delta_0^i}{(r+\gamma)S}\} & \text{if } S \geq -\frac{\Delta_0^i}{(1-r)q_0^i} \\ \max\{-q_0^i, \frac{(1-r)q_0^i}{(r-\gamma)} + \frac{\Delta_0^i}{(r-\gamma)S}\} & \text{if } S < -\frac{\Delta_0^i}{(1-r)q_0^i}, r > \gamma \\ -q_0^i & \text{if } S < -\frac{\Delta_0^i}{(1-r)q_0^i}, r \leq \gamma \end{cases} \tag{3.21}$$

3. **Mixed liquidity and capital ratio:**

$$G_i(S) = \begin{cases} \min\{(\frac{q_0^i}{\gamma} - \frac{S}{\lambda})^+, \frac{\Delta_0^{i,s}}{(1+\gamma)S}, \frac{(1-r)q_0^i S + \Delta_0^i}{(r+\gamma)S}\} & \text{Case A} \\ \max\{-q_0^i, \frac{\Delta_0^{i,s}}{(1-\gamma)S}\} & \text{Case B1} \\ \max\{-q_0^i, \frac{\Delta_0^{i,s}}{(1-\gamma)S}, \frac{rq_0^i S + \Delta_0^i}{(r-\gamma)S}\} & \text{Case B2} \\ \max\{-q_0^i, \min\{\frac{rq_0^i S + \Delta_0^i}{(r-\gamma)S}, \frac{\Delta_0^{i,s}}{(1-\gamma)S}\}\} & \text{Case C1} \\ -q_0^i & \text{Case C2} \end{cases} \tag{3.22}$$

Here the cases are determined by:

$$\begin{aligned} &\text{Case A:} \quad S \geq -\frac{\Delta_0^i}{(1-r)q_0^i} \text{ and } \Delta_0^{i,s} \geq 0 \\ &\text{Case B1:} \quad S \geq -\frac{\Delta_0^i}{(1-r)q_0^i} \text{ and } \Delta_0^{i,s} \leq 0 \text{ and } r > \gamma \\ &\text{Case B2:} \quad S \geq -\frac{\Delta_0^i}{(1-r)q_0^i} \text{ and } \Delta_0^{i,s} \leq 0 \text{ and } r \leq \gamma \\ &\text{Case C1:} \quad S \leq -\frac{\Delta_0^i}{(1-r)q_0^i} \text{ and } \Delta_0^{i,s} \leq 0 \text{ and } r > \gamma \\ &\text{Case C2:} \quad S \leq -\frac{\Delta_0^i}{(1-r)q_0^i} \text{ and } \Delta_0^{i,s} \leq 0 \text{ and } r \leq \gamma \end{aligned} \tag{3.23}$$

We see that the response function for the mixed constraint case is similar in form to the leverage ratio only scenarios, and both these cases are somewhat more complicated than the LCR case.

3.4 Existence and uniqueness of Nash Equilibria

Using the results from the previous section, we know that all equilibrium prices S^* are solutions of the equation (3.16). However, the opposite statement is not necessarily true. The following theorem shows under what conditions the problem has at least one Nash equilibrium solution. Moreover, it also shows that when the problem has multiple NE solution, the higher the NE price is, the better the system will become.

Theorem 9. *Suppose that*

- *In the LCR case:* $S^-_{\min} \geq \lambda \max_{i=1\dots N}\{q^i_0\}$, *where* $S^-_{\min} := S_0 - \lambda \sum_{\Delta^{i,s}_0<0} q^i_0$.
- *In the CAR case:* $r > \gamma$ *and* $S_{\min} \geq [K+1]\lambda \max_{i=1\dots N}\{q^i_0\}$, *where* $S_{\min} := S_0 - \lambda \sum_{i=1}^N q^i_0$ *and* $K := \frac{1-r}{r+\gamma}$.

Then, the following statements are true

1. *There exists at least one NE solution to the problem (3.6).*
2. *The NE solution that corresponds to the largest NE price is Pareto-optimal amongst all NE strategies in* $\mathscr{D}^{NE}$.

Note that the technical condition for the LCR case is very easily satisfied, in particular if $S^-_{\min} \geq \lambda \sum_{\Delta^{i,s}_0<0} q^i_0$ and $S^-_{\min} \sum_{\Delta^{i,s}_0<0} q^i_0(1-\gamma) \geq \Delta^-$. The first requirement $S^-_{\min} \geq \lambda \sum_{\Delta^{i,s}_0<0} q^i_0$ means that the market is large enough to completely absorb all liquidity put on the market by distressed banks even if their leverage is 100% leverage. The second requirement implies that if the distressed banks are merged as a single aggregate bank that liquidates all its assets, it still remains able to meet the liquidity constraint. These conditions are clearly very realistic requirements.

However, the technical condition for the CAR case is strong one. Under this condition, a bank that is initially incompliant (knowing trading strategies of other banks) is not able to meet the constraint again if it tries to buy the risky asset to rise the price up. The only way to comply, if possible, is to sell part of its risky asset to deleverage its portfolio.

The next theorem studies the uniqueness of solutions of the problem.

Theorem 10. *Suppose that*

- *In the LCR case:* $(S^-_{\min})^2 > -\frac{\lambda\Delta^-}{1-\gamma}$, *where* $\Delta^- := \sum_{\Delta^{i,s}_0<0} \Delta^{i,s}_0$.
- *In the CAR case:* $(S^-_{\min})^2 > -\frac{\lambda\Delta}{r-\gamma}$, *where* $\Delta := \sum_{i=1}^N [c^i_0 - d^i_0]$.

Then the Nash equilibrium solution, if exists, is unique.

Exchangeable Banks: If the conditions given in the statement of Theorem 10 do not hold, then there may be many equilibria with different sets of defaulted banks. One can provide fully detailed results in the case of exchangeable banks, which means they have the same balance sheets $\mathscr{B}^i = \mathscr{B}$ and regulatory parameters. In the case of initially LCR-

compliant banks, the monotonicity of the response function given by (3.20) easily implies there always exists a unique Nash Equilibrium which is buy-only. The following proposition sums up results on the existence and uniqueness of solutions in all possible cases where the banks are initially noncompliant, $\Delta_0^s = \Delta_0^{i,s} < 0$.

Proposition 11. *Suppose that N exchangeable banks have $\Delta_0^s < 0$. Define*

$$S_{\min} := S_0 - N\lambda q \text{ and } S_\pm := \frac{S_0 \pm \sqrt{S_0^2 + \frac{4N\lambda\Delta_0^s}{1-\gamma}}}{2}$$

and suppose that $S_{\min} \geq \lambda q$.

1. *If $S_{\min} \geq -\frac{\Delta_0^s}{(1-\gamma)q}$ then there exists a unique NE price $S^* = S_+$ in which all banks survive.*
2. *If $S_{\min} < -\frac{\Delta_0^s}{(1-\gamma)q}$ and*
   (a) *$S_0^2 + \frac{4N\lambda\Delta_0^s}{1-\gamma} > 0$ and $S_+ > NS_-$ then there exist unique NE price $S^* = S_+$. All banks survive in this state.*
   (b) *$S_0^2 + \frac{4N\lambda\Delta_0^s}{1-\gamma} \geq 0$ and $S_+ < NS_-$ then there exist two NE prices $S^* \in \{S_+, S_-\}$. All banks survive in these states.*
   (c) *$S_0^2 + \frac{4N\lambda\Delta}{1-\gamma} < 0$ then there exists a unique NE price $S^* = S_{\min}$ in which all banks default.*

The following proposition sums up information for the CAR contraint case. Let us denote $K = \frac{1-r}{r-\gamma}, \tilde{S} = -\frac{\Delta}{(1-r)q}, \Delta = C_0 - D_0 < 0$ and suppose that $S_{\min} \geq (K+1)\lambda q$ and $r > \gamma$. As in the LCR case, we will consider the case where all banks are initially incomplicant only, since the opposite case gives unique NE solution. Recall that the fixed point equation now becomes

$$\frac{S - S_0}{N\lambda} = \sum_{i=1}^{N} \max\{-q, x_{H-}^i(S)\}\mathbf{1}(S < \tilde{S}) + \sum_{i=1}^{N} \min\{x_{EC}(S), x_{H+}^i(S)\}\mathbf{1}(S \geq \tilde{S}),$$

where $x_{H\pm}^i := \frac{(1-r)q_0}{(r\pm\gamma)} + \frac{\Delta_0}{(r\pm\gamma)S}$ and $x_{EC}(S) := \min\{(\frac{q_0}{\gamma} - \frac{S}{\lambda})^+, \frac{C_0}{(1+\gamma)S}\}$.

Proposition 12. Case 1: Selling to survive. *If the equation $S = S_0 + N\lambda x_{H+}^i(S)$ has no real solutions and*

1. *If* $x^i_{H-}(S_{\min}) > \frac{S_{\min}-S_0}{N\lambda}$ *then there exists unique NE corresponding to the larger real solution* S_2 *of the equation* $x^i_{H-}(S) = \frac{S-S_0}{N\lambda}$.

2. *Else*

   (a) *If If* $x^i_{H-}(S) = \frac{S-S_0}{N\lambda}$ *has two real solutions* $S_1 \leq S_2$*, Then* S_1, S_2 *are both NE prices when* $S_2 < NS_1$ *and* S_2 *is the only NE price otherwise.*

   (b) *Else the system defaults at the unique NE state* $x = -q$.

Case 2: Buying to survive. *If the equation* $S = S_0 + N\lambda x^i_{H+}(S)$ *has two real solutions* $S_3 \leq S_4$ *and*

1. *If* $x_{EC}(S_4) \geq \frac{S_4-S_0}{N\lambda}$ *: Then* S_3, S_4 *are both NE prices when* $S_4 < NS_3$ *and* S_4 *is the only NE price otherwise.*

2. *If* $x_{EC}(S_2) < \frac{S_2-S_0}{N\lambda}$ *and* $x_{EC}(S_1) > \frac{S_1-S_0}{N\lambda}$ *: Then* S_3 *is the only NE price.*

3. *If* $x_{EC}(S_1) < \frac{S_1-S_0}{N\lambda}$ *: Then the system default, i.e.* $x = -q$*. This is an example of Prisoner's Dilemma when banks fail to cooperate even when buying strategies are available to save the system.*

The proof of Proposition 11 is given in the appendix. The case of CAR constraint is left to interested readers.

3.5 System under distress

In this section, we consider the situation when all banks in the network are under distress, i.e. they are all initially incompliant. This is an appropriate simulation of reality when the financial system is in crisis. In the first subsection, we show that in such a difficult situation, banks tend to cooperate to reach a Pareto optimal state altogether. We then provide in the second subsection a monotonic algorithm to find all Nash equilibrium states.

3.5.1 A cooperative game

Consider a system where all banks initially violate the regulatory constraints, i.e. $H^i(0) \leq 0$ or all i. To ensurer the existence of Nash equilibria, we will assume that the technical conditions given in Theorem 9 are satisfied. Under such conditions, we have the following result.

Theorem 13. *The following statements are true*

1. *There exists unique Pareto-optimal strategy for the system amongst all admissible strategies in $\mathscr{D}^{ad}$. This strategy is a Nash equilibrium.*
2. *The NE price corresponding to this strategy is the largest solution of the fixed point equation $S = S_0 + \lambda G(S)$, where $G(S) := \sum_{i=1}^{N} G_i(S)$ denotes the aggregate response function.*

   The proof of this theorem is given in the appendix.

3.5.2 A monotonic algorithm to find NE solution

In this subsection we will provide an algorithm to find the NE prices of the constrained Nash Equilibrium game. In general, the most straightforward algorithm is the following.

- Step 1: Solve the fixed point equation $S = S_0 + \lambda G(S)$. Note that this equation can be reduced to piece-wise quadratic equations defined on consecutive intervals, therefore it can be solved explicitly.
- Step 2: For each NE price S^* that solves the above equation, we define $x^{i*} = G_i(S^*)$ and check if the vector x^* satisfies the equation $x = F(x^*)$.

The following algorithm provides a simpler method to find all NE solutions. It is based on the idea that if the system is in distress, then each admissible strategy is dominated by a NE strategy (from Theorem 13). Therefore, if starting with an admissible strategy, we can improve this strategy step by step until we reach a NE one.

Notice that when the regulatory constraint takes one of three forms LCR, CAR or MLC and when the market is under distress, the fixed point equation for S now becomes

$$S = S_0 + \lambda \sum_{i=1}^{N} G_i(S), \quad G_i(S) = \max(-q_0^i, x_H^i(S)), \tag{3.24}$$

where $x_H^i(S)$ denotes the unique feedback solution of the equation $H(x^i, x^{-i}) = 0$. Recall that the functions G_i are non decreasing when H is LCR, CAR or MLC. For ease of notation, for each subset $A \subseteq [N]$ we will denote $A^c := [N] \backslash A$ and define

$$f^A(S) := S - S_0 - \lambda \sum_{i \in A} x_H^i(S) + \lambda \sum_{i \in A^c} q_0^i. \tag{3.25}$$

Notice that this equation reduces to a quadratic one when H is LCR, CAR or MLC.

Algorithm

1. Step 1. Start with a guess $A_1 \subseteq [N]$ as the set of hypothetical survival banks. This first guess is found in such a way that equation $f^{A_1}(S) = 0$ has solution (the most trivial guess is $A_1 = \emptyset$). Denote S_1 the largest solution of this equation.

2. With $S = S_1$, define $A_2 = \{i \in [N] : G_i(S_1) > -q_0^i\}$. Then solve the equation $f^{A_2}(S) = 0$. Denote S_2 the largest solution of this equation. Later we will prove that $S_1 \leq S_2$ or equivalently, $A_1 \subseteq A_2$. Furthermore, S_2 corresponds to an admissible strategy $x = G(S_2)$ for the system.

3. Repeat Step 2 we will obtain a monotonic sequence $A_2 \subseteq \cdots \subseteq A_n$. Step 3 stops when $A_n = A_{n+1}$. Denote $A_n = X_1$. This is the set of survival banks in a Nash Equilibrium state. Moreover, it is the least favorable state if we start with $A_1 = \emptyset$.

4. Step 4. Add one more hypothetical survival bank from X_1^c to X_1 if exists. More precisely, we search i over X_1^c and stop at the first i when the equation $f^{B_1}(S) = 0$ has solution, where $B_1 := X_1 \cup \{i\}$. Now we repeat step 3 until we find the second NE state, and keep going until we find all NE solutions.

The above algorithm is based on the following lemma.

Lemma 14. *Let $A \subseteq [N]$ be such that the equation $f^A(S) = 0$ has at least one solution. Let S^A be the largest solution of this equation. The following statements are true.*

1. *Define* $B := \{i : G_i(S^A) > -q_0^i\}$. *Then the equation* $f^B(S) = 0$ *has at least one solution. Moreover,* $S^B \geq S^A$ *and* S^B *corresponds to an admissible strategy for the system. Consequently, there exists a NE state whose set of survival banks* X *contains* B.

2. *Conversely, if* X *is the set of survival banks in a NE state and* $B \subseteq X$, *then the equation* $f^B(S) = 0$ *has at least one solution and* $S^B \in [S_{\min}, S_0]$. *Moreover, the mapping*

$$2^X \ni B \to S^B$$

is a non decreasing mapping.

3.6 A Game with Tipping Points

Since the assumption that all banks know the full balance sheets and trading strategies of all other banks is counterfactual, it is important to investigate the robustness of the system as a whole when banks implement strategies based on wrong information. The following very simple example highlights the general statement that multiagent games such as ours exhibit dramatic *tipping points*, that is, circumstances where a very small adjustment leads to huge consequences.

We compare two scenarios involving N banks. The benchmark scenario, Scenario A, has exchangeable banks that are marginally compliant with the LR requirement. In the alternative, Scenario B, the only adjustment is that one bank transfers equal cash amounts of K to the $N-1$ remaining banks, leaving the total capital of the system unchanged. We compare the outcomes using a systemic risk measure M that penalizes banks that are sufficiently noncompliant:

$$M_\epsilon = \sum_{i=1}^{N} \mathbf{1}(H^{i,K} < -\epsilon) \tag{3.26}$$

In Scenario A, we assume:

C1 All banks exercise their Pareto strategy for the LR requirement, under the assumption all other banks are identical to themselves.

C2-A Banks have identical balance sheets $\mathscr{B}$ which are borderline compliant, i.e. $H_0^K = (1-r)q_0S_0 + \Delta_0 = 0$, and furthermore, $S_0 \geq \frac{N\lambda(1-r)q}{(r+\gamma)}$.

In Scenario B, we assume [C1] and replace [C2-A] by:

C2-B The balance sheets of Scenario A are adjusted by transferring equal cash amounts of K from one bank, now called non-compliant (NC), to the $N-1$ remaining banks, now called the compliant (C).

In Scenario A, the condition $S_0 \geq \frac{N(1-r)q}{(r+\gamma)}$ ensures that each bank follows (3.21) and does nothing, $x = 0$. This results in an unchanged equilibrium price $S_1 = S_0$ and the resultant systemic risk measure is $M_\epsilon = 0$.

In the alternative, Scenario B, the initially non-compliant (NC) bank assumes all other banks are in a similar state of non-compliance, and will sell assets. The amount sold, $-x_{NC} > 0$, is computed by assuming the final price $S_{NC} = S_0 + \lambda N x_{NC} < S_0$ that arises when all N banks sell $-x_{NC}$ computed using (3.10) with Δ replaced by $\Delta - (N-1)K$:

$$[(1-r)q - (r-\gamma)x_{NC}]S_{NC} + \Delta - (N-1)K = 0 \tag{3.27}$$

Neglecting terms of $O(K^2)$, one finds

$$x_{NC} \sim \frac{-(N-1)K}{\lambda N(1-r)q - (r-\gamma)S_0} + O(K^2)$$

On the other hand, the initially compliant (C) banks have excess cash. They assume the final price will be $S_C = S_0 + \lambda N x_C > S_0$ that arises when all N banks buy x_C computed using (3.10) with Δ replaced by $\Delta + K$:

$$[(1-r)q - (r+\gamma)x_C]S_C + \Delta + K = 0 \tag{3.28}$$

Again, neglecting terms of $O(K^2)$, one finds

$$x_C \sim \frac{K}{\lambda N(1-r)q - (r+\gamma)S_0} + O(K^2)$$

The actual final price is therefore $S_1 = S_0 + \lambda[(N-1)x_C + x_N]$, which we note is strictly

between S_{NC} and S_C. The C-banks end up *non-compliant* with

$$H_C^K = [(1-r)q - (r+\gamma)x_C]S_1 + \Delta + K = -\lambda[(1-r)q - (r+\gamma)x_C][x_C - x_N] < 0 \ ; \quad (3.29)$$

while the NC-bank ends up *compliant* with

$$H_{NC}^K = [(1-r)q - (r-\gamma)x_{NC}]S_1 + \Delta - (N-1)K = \lambda[(1-r)q - (r-\gamma)x_{NC}][x_C - x_N] > 0 \ . \quad (3.30)$$

In fact, if

$$\epsilon \lesssim K\lambda(1-r)q \left[\frac{1}{\lambda N(1-r)q - (r+\gamma)S_0} + \frac{(N-1)}{\lambda N(1-r)q - (r-\gamma)S_0} \right] + O(K^2) \ ,$$

all $N-1$ initially compliant banks end up non-compliant with $H^K < -\epsilon$, and the resultant risk measure will be the extreme value $M_\epsilon = N-1$ in Scenario B. In other words, if the safety buffer ϵ is small relative to the shock K, the dramatic result is that the initially non-compliant bank becomes compliant, while all the remaining banks end up non-compliant, due to the price impact caused by the selling of one bank!

One might object to this conclusion on the grounds that the non-compliant bank is excessively pessimistic in assuming all other banks are also non-compliant, or that the compliant banks are unjustifiably complacent. However, one can argue that non-compliant banks, aware of their deficient information, will make pessimistic estimates in order to be certain to comply. Moreover, the initially compliant banks, knowing they have an implied ϵ buffer, seem to be at least partially justified in making such a complacent "business as usual" assumption.

The overall conclusion is that when banks act rationally but with wrong information, surprisingly dramatic tipping points can occur.

3.7 Systemic Risk Scenarios

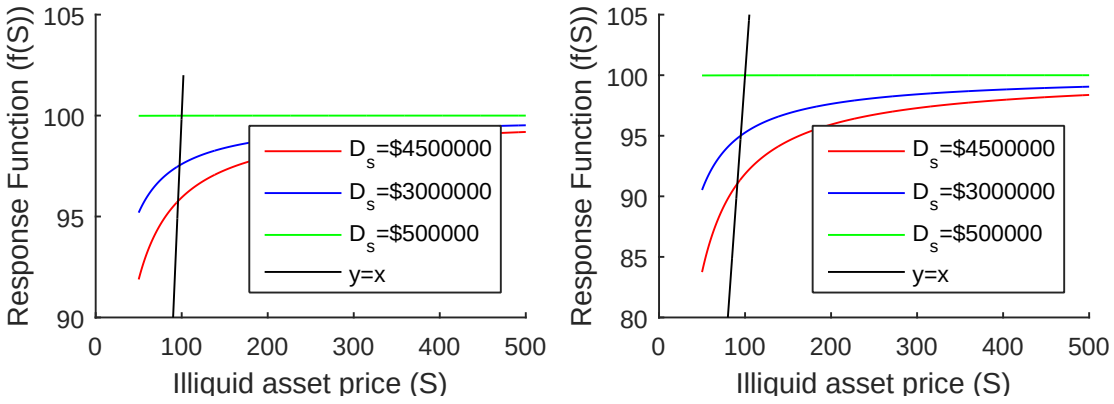
In this section we study the effect of regulations on a market with a fixed number of participants, and a single shared illiquid asset. The usual assumption of perfect information is still kept, as well as the transaction costs and permanent price impact factors are shared

by all market participants. Unless otherwise stated, all agents conform to the parameter values given in the table below, with all dollar values perturbed within 10% of their mean values under a uniform distribution.

Table 3.2: Sample Parameters

Parameter	Parameter Value
Initial Stock Price	$S_0 = \$100$
Temporary Market Impact	$\Gamma = 1\%$
Permanent Market Impact	$\Lambda = \$10^{-7}$/unit
Initial Holdings	$q_0 = 200000$
Initial Cash	$C_0 = \$1000000$
Initial Equity	$E_0 = 5\%$ of total asset
Number of agents	$N = 100$

3.7.1 Liquidity Requirements



(a) $N = 100$ (b) $N = 200$

Figure 3.1: The change in response function from varying level of liquidity requirement as well as number of market participants

The result of having only the Liquidity Requirement enforced upon the banks is straightforward. As seen in figure 3.1, the more liquid asset required to be held, the less illiquid asset can be invested into. Furthermore it is relatively straightforward, when an agent does not have enough liquid assets to satisfy this requirement, the agent will simply sell until he can satisfy his requirement or liquidate himself while trying. All of this results in a lower equilibrium price for the illiquid asset, and demonstrating that the more liquidity

hoarded, the lower the marked to market price of the illiquid asset. We can also see that by doubling the number of participants in the market, we further enhance the firesale effect due to more agents needing to unload the illiquid asset in order to meet their liquidity requirements. This is a stark comparison to the point of view of a single agent's balance sheet, where 10-20% of shares could have been maintained if only one bank was participating in the market.

3.7.2 Stylized Experiments

For this experiment we start with a stylized system of exchangable banks which must satisfy both the liquidity constraint as well as the leverage ratio constraint. The parameters given are:

1. total fixed assets: $NqS_0 = \$10Tr$
2. number of banks: $N = 10,100$
3. $S_0 = 100$, $\gamma = 0.02, r = 0.05$
4. $D_0^\ell = (1-r)qS_0$, $D_0^s = 0.05D^\ell$
5. C_0 uniformly random over range $[(1-\eta_-)D_0^s, (1+\eta_+)D_0^s]$.

Furthermore the parameter λ is set so that $S_0 - \lambda * Nq = \frac{1}{2}S_0$. As shown in the previous section, it is natural that the resulting illiquid asset price is lower than the initial price if the banks enters into a sell strategy. On the other hand, by limiting the initial cash accounts of the banks to exactly matching their short term debt requirements ($\eta = 0$), banks become unable to execute buy strategies. The following figure demonstrates the above two points and compares the possible strategies if the banks were required raise equity during the trading period.

The above figure shows that when all agents are indeed on the boundary of the liquidity constraint, buy strategies are not available even if the leverage constraint is relaxed. More drastically, if regulators were to impose a harsher leverage constraint onto a system of banks such as this one, it could result in the Nash Equilibrium strategy becoming limited to a total system-wide liquidation event.

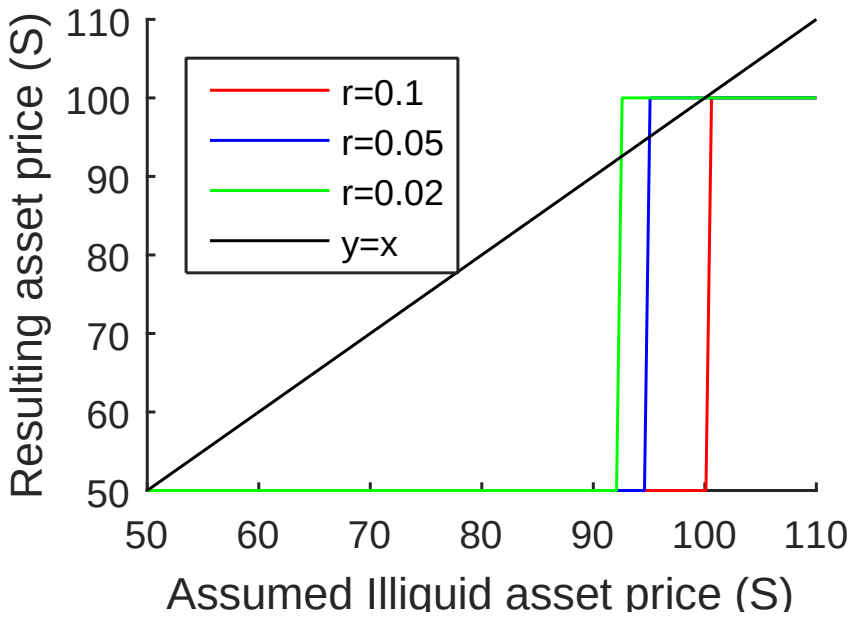
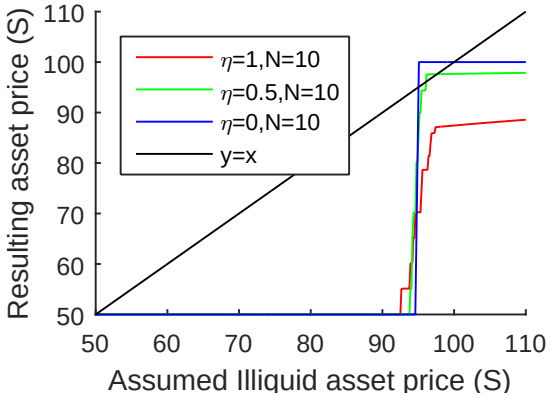


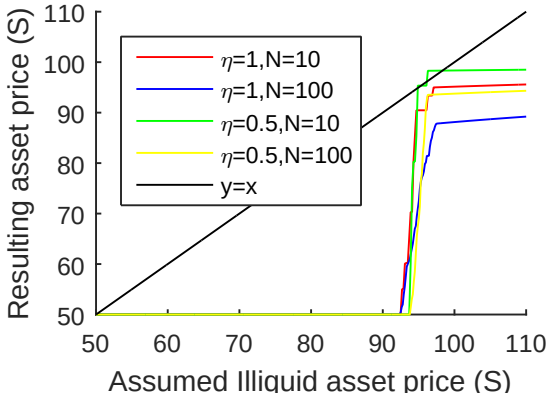
Figure 3.2: The system of exchangable banks starts with initial $r = 0.05$, but at the end of the trading period they must satisfy a new leverage requirement. All banks are fully exchangable hench $\eta = 0$

The following two figures study the effects of heterogeneity on the banking system. In particular each bank has a different amount of initial cash, while the total cash in the system remained the same and as a result individual equity levels are changed accordingly as well. By giving varying amount of initial cash account to each bank, all banks now must execute different strategies, and even more so, banks are no longer at their constrained optimal asset allocation initially. This results in a net loss to the system through the transaction cost γ regardless of the resulting illiquid asset price value. In particular, when banks assume that all other banks have the same balance sheet as itself, this leads to banks which must sell to aim for a even lower price, and banks who simply want to spend all excess cash to purchase less shares. As seen in Figure 3.3 (a), the red line is much lower than its corresponding response function in Figure 3.3 (b), due to the above effect.

Overall, based on the previous figures of aggregate response function, there appears to be a large asymmetry between the upward movement and the downward movement of the price of the illiquid asset. In the all of the previous figures, it seems even with extra



(a) Banks assume all other banks are exchangable



(b) Banks have full information

Figure 3.3: The heterogiety level of the agents with in the system is varied.

cash on hand to invest, the banks cannot push the illiquid asset price up high enough to make it worthwile. Since the permanent market impact symmetrically impacts the illiquid asset price, it falls to the proportional temporary impact factor to cause this effect. In fact, when we study the equity maximization portion of the individual response functions, we can see both that γ appears on the denominator of the first term, hence a minor increase in γ will greatly reduce the optimal illiquid asset holding level to while maximizing equity. A closer inspection also show that the second term $\frac{S}{\lambda} \approx O(N)$, hence also decreases the optimal illiquid asset holding levels. This provides sound evidence of that having more competition in the market reduces price manipulation behaviours.

3.8 Conclusions and Remarks

We studied the effects of regulatory constraints on a system of agents, and provide a framework to further study multiple types of constraints simultaneously as well as obtain the Pareto Nash Equilibrium explicitly in linear time. More specifically, we focused on two typical regulatory constraints faced by a system of financial institution. The first one being a constraint on the short term liquid asset an institution must maintain, the second one being the leverage ratio. Both of these constraints have been implemented, and strictly monitored in the real world setting by various regulatory agencies, providing the necessity

as well as the interest to study these to regulatory constraints. In the following paragraphs, we will provide a small summary of the effects of each regulatory constraint, as well as some analysis into future work that could be built upon this paper.

For the liquidity requirement, if implemented by itself in this simple agent-based model framework, the cost to the system is easy to measure and there are some qualitative analysis of the Nash Equilbrium strategy can be done without finding the NE explicitly. Based on our analysis from the previous sections, the only strategy capable of allowing a bank to reach its minimum cash holding level is to sell the illiquid asset regardless of its price. This creates to a downward pressure on the price of the illiquid asset, and in turn damages the mark-to-market balance sheets of all agents in the system. This leads to the case, where the more liquid asset institutions are required to hold, the lower the illiquid asset price will become. This result is straightforward, and expected, and is one of the reasons we chose to analyse liquidity requirement first.

The leverage ratio requirement can be viewed as a more harsh requirement than the liquidity requirement. From our numerical studies, we can see that under the parameters that we chose, there is always a lower equilibrium price point corresponding to a total liquidation strategy of the system. The transaction cost and the negative pressure on the illiquid asset, both impacts the equity level of the agents especially when the agents are trying to convert illiquid asset into liquid assets which compounds the problem by costing them liquid asset through the proportional transaction cost. This analysis shows the pro-cyclicality of using leverage ratio as a regulatory constraint, where in times of crisis, the simple drop of the illiquid asset price will breach the leverage threshold, and cause further drops in asset prices.

In reality, banks are bound by vast amounts of regulatory constraints, even many regulations themselves have competing interests. In this paper we show a proof of concept for this agent-based simulation framework by analysing the effects of having two simultaneous regulatory constraints. Future work include implementating this agent-based model framework directly into a network systemic risk model such as proposed by Eisenberg and Noe (2001). The interbank network and the initial shock can be modelled by the short term debt holdings of the agents. Also, a generalization of the linear demand function can be used with minimal effects on the algorithm to determine the equilibrium price.

Appendix 3.A Proofs of Main Results

Proof of Theorem 9

Proof. **LCR case.** We first prove the existence of solutions. To do so, let us get back to the optimization problems of individual banks. If $\Delta_0^{i,s} \geq 0$, the optimal trading strategy of bank i, knowing S_{-i}^* is given by

$$x^{i*} = F_i(x^{-i*}) := (\frac{q_0^i}{2\gamma} - \frac{S_{-i}^*}{2\lambda})^+ \wedge \frac{-S_{-i}^* + \sqrt{(S_{-i}^*)^2 + \frac{4\Delta_i^{i,s}}{1+\gamma}}}{2\lambda}.$$

Otherwise, if $\Delta_0^{i,s} < 0$, by using the condition in the statement of the theorem, we can write the optimal strategy for the bank i as follows

$$x^{i*} = F_i(x^{-i*}) := -q_0^i \vee \frac{-S_{-i}^* + \sqrt{[(S_{-i}^*)^2 + \frac{4\Delta_i^{i,s}}{1-\gamma}]^+}}{2\lambda}.$$

In this formula, the fractional term corresponds to the larger solution of the constraint equation $H^{i,K} = 0$ (if it exists), or the maximizer of $H^{i,K}$ in case $H^{i,K} = 0$ has no solution. In the former case, we chose the larger solution to minimize the selling amount and therefore optimize the equity; while the latter case yields $x^{i*} = -q_0^i$ because $S_{-i}^* \geq 2\lambda q_0^i$. Writing the optimal trading strategy this way, we can see that F_i is continuous.

Let us define $F := (F_i)_{i=1\dots N}$ and consider the fixed point equation $x = F(x)$. Each solution of this equation corresponds to a NE point of the game. It is easy to see that the function F is bounded and continuous, therefore the fixed point equation has at least one solution, which ensures the existence of NE points.

Now, consider two NE points x^1, x^2 associated to NE prices S_1, S_2, respectively. Suppose that $S_1 < S_2$. We will prove that S_2 is a better NE price than S_1. Indeed, for initially compliant banks, the functions G_i are non increasing in S, therefore $x_i^1 \geq x_i^2$. Note that $E_i(x) = \Delta_0^i + (q_0^i - \gamma x)S$, we can deduce that $E_i(x^1) \leq E_i(x^2)$. Therefore, S_2 is better for the bank i than S_1 (note that they are both alive). Likewise, for the initially distressed banks we have $x_i^1 \leq x_i^2$ and $E_i(x^1) \leq E_i(x^2)$. The former implies that $\theta_i(x^1) \leq \theta_i(x^2)$. We then come to a conclusion that in any case, S_2 is always better for the system than S_1.

Note that the number of NE points cannot be larger than the number of solutions to the equilibrium equation (3.16), which is finite. Therefore, the Pareto-optimal NE point of the game exists and is uniquely determined by the largest NE price.

CAR case. We now prove the existence of solutions. If $H(0, x^{-i*}) \geq 0$ then the optimal strategy of the bank i is to buy. The computation of the optimal strategy of the bank i in this case is quite straightforward

$$x^{i*} = \min\{x^i_E(S^{-i*}), x^i_C(S^{-i*}), x^i_{H+}(S^{-i*})\},$$

where $x^i_E(S^{-i*}), x^i_C(S^{-i*})$and $x^i_{H+}(S^{-i*})$ denote the values at which $E^i(x, x^{-i*})$ attains its maximum, $C^i(x, x^{-i*}) = 0$ and $H^i(x, x^{-i*}) = 0$, respectively. These functions are well defined and bounded, continuous.

If $H(0, x^{-i*}) \geq 0$ then the optimal strategy of the bank i is to sell. We use a similar argument as in the LCR case to obtain

$$x^{i*} = \max\{-q^i_0, x^i_{H-}(S^{-i*})\},$$

where

$$x^i_{H-}(S^{-i*}) = -\frac{S^{-i*} - K\lambda q^i_0}{2\lambda} + \frac{1}{2\lambda}\sqrt{\left((S^{-i*} - K\lambda q^i_0)^2 + \frac{4\lambda H^i(0, x^{-i*})}{r+\gamma}\right)^+}.$$

The response function $F(x)$ is thus continuous and bounded. Consequently, the fixed point equation $x = F(x)$ has at least one solution. This proves the existence of a NE solution of the problem.

We now show that if S_1 and S_2 are two NE prices such that $S_1 \leq S_2$ then S_2 is better than S_1 for the system. The technique of the proof is slightly different from the LCR case because the response functions G_i are not necessarily monotonic. It suffices to prove that $E_i(S_1) \leq E_i(S_2)$ for the case $G_i(S_1), G_i(S_2) \geq 0$. Indeed, let $\hat{S}^i$ be the unique solution of the equation

$$x^i_{EC}(S) := \min\{x^i_E(S), x^i_C(S)\} = x^i_{H-}(S).$$

If $S_1 \geq \hat{S}^i$ then $G_i(S_1) = x^i_{EC}(S_1)$ and $G_i(S_2) = x^i_{EC}(S_2)$. Since x^i_{EC} is non increasing and

$S_1 \leq S_2$ we can deduce that $x_1 \geq x_2$. Thus $E_i(S_1) \leq E_i(S_2)$.

Otherwise if $S_1 \leq \hat{S}$ then $G_i(S_1) = x_{H-}(S_1)$ and $G_i(S_2) \leq x_{H-}(S_2)$. Note that $x_{H-}(S) = \frac{1-r}{r-\gamma} q_0^i + \frac{\Delta_0^i}{(r-\gamma)S}$ and $E_i(S) = \Delta_0^i + [q_0^i - \gamma G_i(S)]S$ We have

$$E_i(S_1) = \frac{1+r-\gamma}{r-\gamma}\Delta_0^i + \frac{r(1-\gamma)}{r+\gamma} q_0^i S_1 \leq \frac{1+r-\gamma}{r-\gamma}\Delta_0^i + \frac{r(1-\gamma)}{r+\gamma} q_0^i S_2 = E_i(S_2).$$

□

Proof of Theorem 10

Proof. We consider the LCR case first. Rewrite the equation (3.16) in the form $S - S_0 - \lambda \sum_{i \in I^+} G_i(S) = \lambda \sum_{i \in I^-} G_i(S)$. Let us denote by $p(S)$ and $q(S)$ the functions on the left and the right hand sides of this equation, respectively. We will show that the mapping G is a contraction one with given conditions. It is easy to show that $q([0, M)) \subseteq p([0, M))$ for $M > 0$ large enough. Moreover, p, q has derivatives almost everywhere except at a finite points and $p'(S) \geq 1$ a.e. Therefore, the fixed point equation can be rewrite in the form $S = p^{-1}(q(S))$. We will show that $p^{-1}(q)$ is a contraction mapping on $[0, \infty)$. Indeed, at a point S where both p, q have derivatives, by using the explicit form of p, q and the hypothesis given in the statement of the theorem we have $q'(S) \leq \frac{\lambda \Delta^-}{(1-\gamma)S^2} < 1$. Therefore $(p^{-1}(q(S)))' = \frac{q'(S)}{p'(p^{-1}(q(S)))} < 1$.

We now consider the CAR case. Denote $\tilde{S}^i := -\frac{\Delta_0^i}{(1-r)q_0^i}$ and let $\hat{S}^i$ be the unique solution of the equation

$$x_{EC}^i(S) := \min\{x_E^i(S), x_C^i(S)\} = x_{H-}^i(S).$$

We can rewrite the response function G_i as follows

$$G_i(S) = [-q_0^i \vee x_{H-}(S)]\mathbf{1}(S < \tilde{S}^i) + x_{H+}(S)\mathbf{1}(\hat{S}^i \geq S \geq \tilde{S}^i) + x_{EC}(S)\mathbf{1}(S > \max\{\hat{S}^i, \tilde{S}^i\}).$$

Note that the functions $[-q_0^i \vee x_{H-}]$ and x_{H+} are non decreasing, whereas the function x_{EC} is non increasing. We can then apply the same technique as in the LCR case to show that under the conditions given in the statement of the theorem, te fixed point equation (3.16)

can be reduced to a new one with contraction mapping, hence uniqueness of solutions.

□

Proof of Proposition 11

Proof. When all banks have the same set of parameters, the fixed point equation for S^* is given by (denote $\Delta := -\Delta_i$ and $q_i = q$.)

$$\frac{S - S_0}{\lambda} = -N[q \wedge \frac{\Delta}{(1-\gamma)S}]. \tag{3.31}$$

The fixed point equation for x^* is given by

$$x^i = -\left[q \wedge \frac{S_{-i} - \sqrt{[(S_{-i})^2 - \frac{4\Delta}{1-\gamma}]^+}}{2\lambda}\right]. \tag{3.32}$$

Notice that $S_{\min}$ is solution of the equation

$$\frac{S - S_0}{\lambda} = -Nq,$$

and the condition $S_{\min} < -\frac{\Delta}{(1-\gamma)q}$ is to ensure that $S_{\min}$ solves (3.31). Notice that once $S_{\min}$ solves (3.31), the condition $S_{\min} \geq \lambda q$ will ensure that $x^i = -q$ does not solve (3.32), so $x^i = -q$ is never a Nash equilibrium solution. Furthermore, under the condition $S_{\min} < -\frac{\Delta}{(1-\gamma)q}$, if $S_\pm$ are two real solutions of the equation

$$\frac{S - S_0}{\lambda} = -\frac{N\Delta}{(1-\gamma)S} \leftrightarrow S^2 - S_0 S + \frac{\lambda N \Delta}{1-\gamma} = 0, \tag{3.33}$$

then we can show that $S_\pm$ also solve (3.31). Consequently, S_+ is always the Pareto-optimal Nash equilibrium solution of the problem.

We now just have to determine under what condition, S_- is also another price equi-

librium. Notice that given S_-, we can define x^i via the equation $x^i = G_i(S_-)$, or

$$x^i = \frac{-S_{-i} \pm \sqrt{(S_{-i})^2 - \frac{4\Delta}{1-\gamma}}}{2\lambda},$$

where $S_{-i} = S_- - \lambda x_i$. We can see that x_i corresponds to a true Nash equilibrium if and only if

$$x_i < -\frac{S_{-i}}{2\lambda} = -\frac{S_- - \lambda x_i}{2\lambda} \leftrightarrow S_- > -\lambda x_i.$$

However, from (3.33) we have $S_+ S_- = \frac{\lambda N \Delta}{1-\gamma}$. Hence

$$\lambda x_i = \lambda G(S_-) = -\frac{\lambda \Delta}{(1-\gamma)S_-} = -\frac{S_+}{N}.$$

Substitute this to the inequality $S_- > -\lambda x_i$ we arrive at

$$S_- > \frac{S_+}{N} \leftrightarrow S_+ < N S_-.$$

This is the condition under which S_- becomes a true Nash equilibrium price. □

Proof of Theorem 13

Proof. We will prove this theorem for the LCR case only. The CAR case is similar (with some slight modification).

1. Following Theorem 9, between two NE prices, the larger one is better for the system. Then, it suffices to prove that any admissible trading strategy is dominated by a NE one. We will prove a stronger statement: For any admissible strategy $x \in \mathscr{D}^{ad}$, there exists a NE strategy x^* such that $x^* \geq x$ componentwisely.

We proceed the proof by induction. For $N = 1$ the statement is obvious because the optimal trading stratey of an incompliant bank is to sell a minimal amount of risky asset until the liquidity requirement is met. We suppose that this holds true until $N = n$. We will show that the statement is also true for $N = n + 1$.

Indeed, suppose that $x = (x_1^0, x_{-1}^0)$ is an admissible strategy. From the case $N = 1$, when knowing x_{-1}^0 the bank 1 can find an optimal strategy $x_1^1 \geq x_1^0$. Consider the system in

the new state (x_1^1, x_{-1}^0). Since $x_1^1 \geq x_1^0$, this new strategy is still an admissible one because it is better than the old one x for the system. By induction hypothesis, when knowing x_1^1 the banks $\{2, \ldots, n+1\}$ can find a NE strategy $x_{-1}^1 \geq x_{-1}^0$ which is a better strategy for the system than x_{-1}^0. We repeat this process infinite times to obtain an increasing sequence of admissible trading strategies (x_1^k, x_{-1}^k) such that

$$x_1^{k+1} = F_1(x_{-1}^k), x_{-1}^{k+1} = F_{-1}(x_1^k).$$

Since this sequence is increasing and bounded by zero, it converges to some limit x^*. Moreover, by the continuity of the functions F_i, we can pass by limit and obtain the following equations

$$x_1^* = F_1(x_{-1}^*), x_{-1}^* = F_{-1}(x_1^*).$$

This means that x^* is a NE strategy that dominates the initial strategy x.

2. It remains now to show that the largest NE price coincides with the largest solution of the equation (3.16). Indeed, let $S_{\max}^*$ be the largest solution of (3.16). It is obvious that $S_{\max}^*$ is larger than any NE price. By contradiction, we suppose that $S_{\max}^*$ is not yet a NE price and will show that there exists a larger NE price. Indeed, let denote $x^{i*} := G_i(S_{\max}^*)$ and define $S_{-i}^* := S_{\max}^* - \lambda x^{i*}$. Following the proof of Theorem 9, we have

$$x^{i*} = \tilde{F}_i(x^{-i*}) := -q_0^i \vee \frac{-S_{-i}^* \pm \sqrt{[(S_{-i}^*)^2 + \frac{4\Delta_i^{i,s}}{1-\gamma}]^+}}{2\lambda}.$$

Here, the function $\tilde{F}_i$ might take one of two forms depending on the constraint equation $x^{i*}(S_{-i}^* + \lambda x^{i*})(1-\gamma) = \Delta_i^{i,s}$. In general we have $\tilde{F}_i \leq F_i$. Moreover, the contradictory hypothesis implies that there exists i such that $\tilde{F}_i < F_i$, (i.e. this bank acts irrationally). Let us define

$$B := \Pi_{i=1}^N [x^{i*}, 0]. \tag{3.34}$$

Since G_i is a non decreasing function for initially distressed banks and non increasing otherwise, it is obvious that $F(B) \subseteq B$. By the Brouwer's fixed point theorem, there exists $z \in B, z \neq x^*$ such that $F(z) = z$. Define $S^z := S_0 + \lambda \sum_{i=1}^N z_i.$, then S^z is a NE price and $S^z > S_{\max}^*$, which is a contradiction. □

Proof of Lemma 14

Proof. 1. By comparing the functions f^B and f^A, it is easy to check that $f^B(S^A) \leq 0$, which implies that $S^A \leq S^B$.

For all $i \in B$ we have $G_i(S^B) \geq G_i(S^A) > -q_0^i$. Therefore, S^B corresponds to an admissible strategy for the system. By Theorem 13, any admissible strategy is dominated by a NE strategy.

2. By a similar method as in the previous part, we can show that $f^A(S^X) \geq 0$. Moreover, by definition of f^A, we have $f^A(S_{\min}) \leq 0$. So, there exists a maximum number $S^A \in [S_{\min}, S^X]$ such that $f^A(S^A) = 0$.

Now if $C \subseteq D \subseteq X$, then we can check that $f^D(S^C) \leq 0$, thus $S^C \leq S^D$. □

Chapter 4

Double Cascade Model of Financial Crises

"Double Cascade Model of Financial Crises" has been published in International Journal of Theoretical and Applied Finance and is a joint work with Prof. Thomas Hurd, Dr. Davide Cellai, and Dr. Sergey Melnik. The contributions of the author include: 1. Computer programming, numerical analysis. 2. Equal share of technical derivation of this paper, equal share of finalization of this paper, equal share of drafting. Modifications from the original paper include formatting and merging of the bibliography with thesis bibliography. In section 4.6, the author has made the following additions to the original paper, 1. Asset Firesale Mechanism, 2. Systemic Risk Measures, 3. Numerical analysis of the impact of asset firesales using Monte Carlo simulation based on the EU network.

4.1 Introduction

Since the banking crisis of 2007-2008, the types of shocks transmitted through interbank networks during a crisis now thought to be important include not only shocks arising from defaulting banks, but a number of additional phenomena, most notably asset shocks originating with forced sales by banks of illiquid assets and funding liquidity shocks as illiquid banks recall loans from other banks. A comprehensive treatment of these contagion

effects and how they can be computed in stylized financial network specifications can be found in Hurd (2016).

A well-developed strand of literature on bank default cascade models, starting with Eisenberg and Noe (2001) and reviewed in Upper (2011), is based on a network of banks wherein insolvency of a given bank, defined as a bank whose net worth becomes non-positive, will generate shocks to the asset side of the balance sheet of each of its creditor banks. Under some circumstances, such "downstream" shocks can cause further insolvencies that may build up to create a global insolvency cascade. One such contribution is by Nier et al. (2007) that uses Monte Carlo simulation to determine how the key network parameters for their model of defaults in a stylized random network of 25 banks can influence the total number of defaults in a nonlinear, indeed sometimes nonmonotonic, fashion. The paper of Gai and Kapadia (2010a) and its extension by Gleeson et al. (2012), adapt the Watts (2002) network model of information cascades to the context of financial systems, deriving both analytical and Monte Carlo results for the dependence of the default cascade on different structural parameters. May and Arinaminpathy (2010) present approximate analytical formulas for the Nier et al model and Gai-Kapadia model that can explain some of the main properties of the simulation results found in those papers. Amini et al. (2012) and Amini et al. (2013) develop a simple but general analytical criterion for resilience to default contagion, based on an asymptotic analysis of default cascades in heterogeneous networks.

More recently, after remarking on the observed "freezing" of interbank lending around the time of the Lehman collapse, papers by Gai and Kapadia (2010b), Gai et al. (2011) and Lee (2013) adopt variations of an idea that funding illiquidity[1] of a bank can also be transmitted contagiously through interbank exposures. They argue that a bank whose liquid assets are insufficient to cover demands on its liabilities will be illiquid or "stressed", and must reduce its interbank lending, thereby creating shocks to the liability side of its debtor banks' balance sheets. Such "upstream" shocks can cause illiquidity stress in other banks that in some circumstances may build up to create a global illiquidity cascade.

A third channel, sometimes called market illiquidity or the asset fire sale effect, is identified by Cifuentes et al. (2005) who consider a network of banks that hold a common

[1]Funding illiquidity, being the insufficiency of liquid assets to cover a run on liabilities, is distinct from market illiquidity, where assets become difficult to sell due to an oversupply in the market. See Tirole (2011) and Brunnermeier and Pedersen (2009) for a detailed analysis of these concepts.

asset. As stress in the system builds up, some banks are forced to sell large positions in the illiquid asset, leading to a downward price spiral coupled to worsening bank balance sheets across the network. This mechanism has been extended to the sharing of multiple assets in a paper by Caccioli et al. (2014).

These papers simplify by omitting economic details that obscure their focus on a single channel of systemic risk. They share the view that the end result of a crisis is a new equilibrium that is reached through a sequence of mechanistic steps where banks update and modify their balance sheets in response to price changes and shocks transmitted from other banks. Usually, properties of the system at step n of the cascade can be computed in terms of the properties at step $n-1$, yielding a set of equations that can be called a *cascade mapping*. In some models, an analytic cascade mapping can be found that yields a detailed picture of systemic risk without the need for Monte Carlo simulation based computations. Even in such cases however, large scale benchmarking of the analytic approximation method against Monte Carlo simulations is undertaken to validate its use.

An important question now arises whether one can successfully integrate two or more such cascade mechanisms, incorporate "spillover" effects in a realistic, analytically tractable way, and discern new features of systemic risk. The purpose of the present paper is to do exactly this, by introducing a network model of a double illiquidity-insolvency cascade, and deriving an analytic cascade mapping that describes it. This double cascade model, shown schematically in Figure 4.1, can resolve issues not addressable within a single cascade model, for example, to quantify the effect of a bank's behavioural response to liquidity stress on the level of eventual defaults in the entire system. In particular, one can show that a bank that reacts to stress at a time of crisis by shrinking its own interbank lending, thereby inflicting liquidity shocks to its debtor banks, will also protect itself from default due to interbank default shocks.

In the literature on how contagion channels "spill over" into each other, Diamond and Rajan (2005) explain the liquidity, solvency and behaviour of banks through the role they play in intermediating between lenders and borrowers. As discussed in Tirole (2011), the holding of very illiquid assets can cause solvent banks (those with positive equity) to default due to their inability to raise enough liquidity to meet short-term needs. Battiston et al. (2012) model the robustness dynamics of a network of banks linked by their interbank exposures, highlighting the feedback effects of a "financial accelerator" mechanism on the

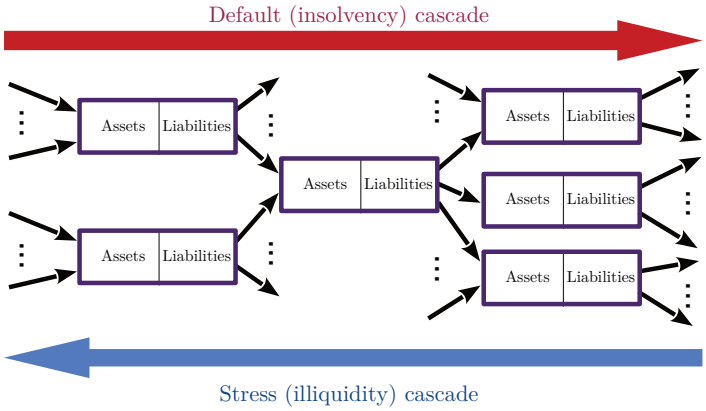


Figure 4.1: A schematic diagram of part of a financial network, showing banks and their balance sheets connected by directed edges representing interbank exposures. Default shocks are transmitted in the forward "downstream" direction, while liquidity or stress shocks are transmitted in the upstream direction.

level of systemic risk. Roukny et al. (2013) further develop this idea, investigating the systemic impact of network topologies and different levels of market illiquidity and applying this methodology to the Italian interbank money market from 1999 to 2011.

It should be made clear that in order to focus on pure contagion effects, the double cascade model and the above models simplify by explicitly ruling out certain other systemic mechanisms that have been explored in the economics and finance literature. In the original paper, the most important economic effect that is ruled out is the asset "fire sale", which means that for the duration of the crisis, changing values of fixed assets are ignored. This is now further investigated following the numerical studies section. Some of the above references discuss the impact of financial cascades on the non-financial economy and the consequent feedback into the financial markets. They find that fire sales of assets amplify any cascade that takes hold in the network. Using readily available data from the European Banking Authority (EBA), we were able to measure the addition of the fire sale effect and determine whether out main conclusions are robust to such an extension.

Section 2 of the paper sets out the network framework and assumptions underlying the balance sheet structure of banks, the timing of the crisis and bank behaviour. These assumptions lead to stock flow consistent rules for the transmission of shocks through the double stress/default cascade, including the conditions for banks to become stressed or

defaulted. Section 3 develops our main technical contribution, which yields an explicit analytic cascade mapping for default and stress probabilities valid in a large heterogeneous network of banks with random balance sheets and interbank exposures and connectivity given by a random"skeleton". Section 4 provides a parallel development of default and stress probabilities for cascades on networks where it is assumed that the skeleton of interbank connections is known explicitly, but balance sheets and exposure sizes are random. Several representative financial experiments are reported in Section 5. First we summarize experiments that validate the accuracy of the cascade mapping by direct comparison of large network analytics to Monte Carlo simulation results. Secondly, we investigate the relationship between the level of stress and default, verifying our assertion that average default probability decreases as average stress probability increases. A final experiment demonstrates that our analysis is still useful when pushed to a highly heterogeneous specification of the model that is consistent with known heuristics of financial networks, such as fat-tailed degree and exposure distributions, and the 2011 stress testing data on 90 large banks in the EU system. We observed that the stylized EU network is very resilient, and only extremely large shocks to the average default buffer size will trigger a cascade of defaulted and stressed banks. Piggy backing on the numerical analysis of this section, we were able to include the firesale effect as part of the simulation process to demonstrate the interactions between these three channels of contagion.

One conclusion of this paper is that our analytic asymptotic results on default and stress probabilities are broadly consistent with results from Monte Carlo simulations on random networks. A second conclusion is that under the assumption of no asset fire sales, stress and default are inversely related: as banks respond to stress more vigorously, creating more network stress, they protect the network from default. Finally, based on numerical studies, we conclude that firesale effect does amplify the systemic risk of the financial network, and provides another link between the stress contagion and the default contagion.

4.2 Cascade Mechanisms

The assumptions we now make lead to a stylized network model admitting double cascades of illiquidity and insolvency that can be computed efficiently under a wide range of initial conditions. The picture to adopt is that of a system hit "out of the blue" by a large

shock that triggers a systemic crisis that proceeds stepwise in time. Many realistic effects that would significantly complicate our model are left out. For example, banks respond mechanistically according to simple stylized rules rather than dynamically optimizing their behaviour. We take a static view, and ignore external cash flows, interest payments and price changes for the duration of the cascade. We make an assumption of zero-recovery on interbank assets on the default of a counterparty during the cascade. Our assumptions imply that insolvency (that is, when equity is nonpositive) is equivalent to default (that is, the bank can no longer pay some liability). This means that even when stressed, banks always pay recalled interbank liabilities in full, until the moment they become insolvent. We make a final market price assumption, that the balance see price of a unit of the fixed asset is the equilibrium price determined by an inverse demand function. This assumption implies that the balance sheet is updated due to changes in the equilibrium price of the fixed asset, and insolvence could occur due to it. It should be noted for the analytical analysis sections we rule out other channels of systemic risk, in particular asset fire sales.

The network consists of a collection of N banks, labelled by numbers from the set $\mathscr{N} = \{1, 2, \ldots, N\}$, each structured with a balance sheet as shown in Figure 4.2.

The subset of debtor banks of a bank $v \in \mathscr{N}$ is called thein-neighbourhood of v, denoted by $\mathscr{N}_v^-$, and the *in-degree* is the number of debtors $j_v = |\mathscr{N}_v^-|$. Similarly the bank's creditor banks form a subset called the out-neighbourhood $\mathscr{N}_v^+$ of v whose size $k_v = |\mathscr{N}_v^+|$ is called the *out-degree*. The web of interbank counterparties is called the "skeleton", and is identifiable as a directed graph, i.e. a collection of directed arrows, called edges, between pairs of nodes. Each debtor-creditor pair v, w with $w \in \mathscr{N}_v^+$ is denoted by an edge $\ell = (vw)$ pointing from v to w.[2] The *type* (j, k) of a bank v is its in-degree $j = j_v$ and out-degree $k = k_v$, and we write $v \in \mathscr{N}_{jk}$. The *type* $(k, j) = (k_\ell, j_\ell)$ of an edge $\ell = (vw)$ is the out-degree $k_\ell = k_v$ of the debtor bank v and the in-degree $j_\ell = j_w$ of the creditor bank w, and we write $\ell \in \mathscr{E}_{kj}$.

The bank equity or net worth $\Delta_v = \max(A_v^F + A_v^{IB} + A_v^L - L_v^D - L_v^{IB}, 0)$ of bank v incorporates *limited liability*, which assumes that the bank must default and be forced to liquidate at the first moment it becomes insolvent, that is, its liabilities exceed its assets.

[2]The convention that arrows point from debtors to creditors means that default shocks propagate in the downstream direction. Confusingly, much of the systemic risk literature uses the reverse convention that arrows point from creditors to debtors.

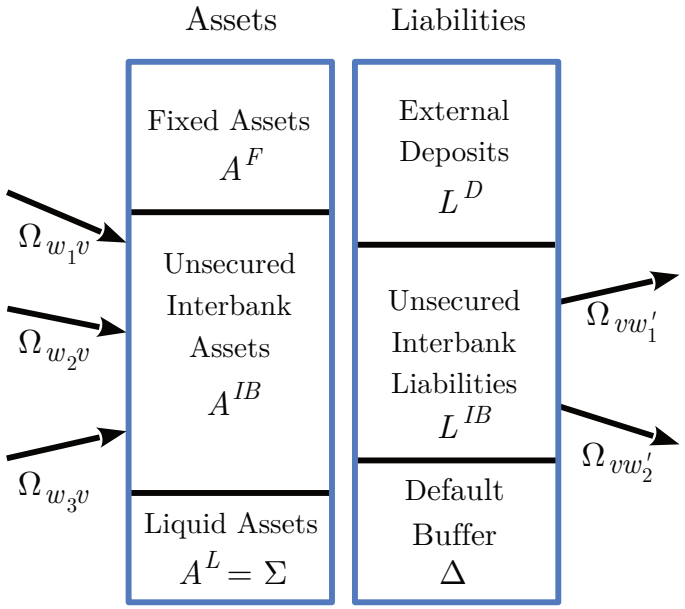


Figure 4.2: The stylized balance sheet of a bank v with in-degree $j_v = 3$ and out-degree $k_v = 2$. Banks w_1, w_2, w_3 are debtors of v while w_1', w_2' are its creditors. The total exposure of v to w_1 is denoted with $\Omega_{w_1 v}$, and so on. The default buffer Δ of bank v is the difference between assets and liabilities, and the "stress buffer" Σ is the preferred asset class from which it pays creditors' demands.

Thus Δ_v can be interpreted as the bank's *default buffer* that must always be kept positive to avoid default.

Banks are also concerned about the possibility of runs on their liabilities, and try to keep a positive buffer of liquid assets A^L (such as cash and government bonds), henceforth called the *stress buffer* Σ, which are the preferred asset class from which to pay creditors' demands. A bank v with Σ_v fully depleted to zero will be called *stressed*. This does not mean the bank is*illiquid* in the sense of being unable to meet the creditors' demands. Rather it means that the bank is experiencing a significant degree of stress in meeting these demands and must turn to other assets, first interbank assets, then fixed assets, to realize the needed cash. The no fire sale assumption means that as long as they are solvent, stressed banks always meet their payment obligations.

The interbank liabilities L^{IB} and assets A^{IB} decompose into bilateral interbank exposures. For any bank v and one of its creditors $w \in \mathscr{N}_v^+$, we denote by Ω_{vw} the total exposure of w to v. Then we have the constraints

$$A_v^{IB} = \sum_{w \in \mathscr{N}_v^-} \Omega_{wv}; \quad L_v^{IB} = \sum_{w \in \mathscr{N}_v^+} \Omega_{vw} \ . \tag{4.1}$$

Prior to the onset of the crisis, all banks are assumed to be in the normal state. Then, on day $n = 0$, a collection of banks, possibly all banks, are assumed to experience initial shocks. Two kinds of initial shocks are possible. First, an asset shock causes a drop in the mark-to-market value of the fixed asset portfolio, and reduces the default buffer. If the downward asset shock leads to $\Delta \leq 0$, the bank must default. The second kind of shock is a demand shock by the external depositors that reduces the stress buffer. If the demand shock leads to $\Sigma \leq 0$, the bank becomes stressed.

Timing and Bank Behaviour Assumptions:[3]

1. Prior to the crisis, all banks are in the normal state, neither stressed nor insolvent. The crisis commences on day 0 after initial shocks trigger the default or stress of one or more banks;

2. Balance sheets are recomputed daily on avmark-to-market basis and banks respond

[3]Treating cascade steps as daily is simply an aid to understanding.

daily on the basis of their newly computed balance sheets. All external cash flows, interest payments, and asset and liability price changes are ignored throughout the crisis. Note that if asset firesale effects are not incorporated, the fixed asset portion of the balance sheet will remain constant through out the crisis.

3. An insolvent bank, characterized by $\Delta = 0$, is forced into liquidation by the regulator. At this moment, each of its creditor banks are obliged to write down its defaulted exposures to zero thereby experiencing a *solvency shock* that reduces its default buffer Δ.

4. A stressed bank, defined to be a non-defaulted bank with $\Sigma = 0$, reacts one time only, at the moment it becomes stressed, by reducing its interbank assets A^{IB} to $(1-\lambda)A^{IB}$ where λ taken to be a fixed constant across all banks during the crisis. It does so by recalling a fraction λ of its interbank loans, thereby transmitting a *stress shock* to the liabilities each of its debtor banks. Since nondefaulted banks are able to pay all of their liabilities, such recalled loans are repaid in full.

5. A newly defaulted bank also triggers maximal stress shocks (i.e. with $\lambda = 1$) to each of its debtor banks as its bankruptcy trustees recall all its interbank loans, reducing A^{IB} to 0;

6. Stress shocks reduce any bank's stress buffer Σ and stressed banks remain stressed until the end of the cascade or until they default.

Remark 1. *The stress parameter λ was introduced by Gai et al. (2011) to simplify banks' response to liquidity stress and to capture the effect of liquidity hoarding. We suppose that banks react sufficiently strongly to stress by choosing $\lambda \in [0.4, 1]$, to capture the picture that they hoard liquidity to preempt the need for further response later in the crisis. Unlike default, a bank is free to set its response to liquidity: both λ and the size of the stress buffer Σ are its own policy decisions. In a more realistic and complex model, λ would be endogenously determined for each bank, reflecting the cumulative demands on its liabilities.*

The dynamics of the cascade that follow from these assumptions is stock flow consistent. This means that as long as both counterparties are solvent, every transaction involves four equal and opposite balance sheet adjustments: to both the assets and liabilities of both counterparties. The dynamics also has the property that it is completely determined by a reduced set of balance sheet data that consists of the collection of buffers

Δ, Σ and interbank exposures Ω.

We shall now apply the cascade rules to a financial network found on day 0 of the crisis in an initial state described by the following elements: the interbank links which form a directed graph $\mathscr{E}$ on the set of banks $v \in \mathscr{N}$, called the *skeleton*; the buffers Δ_v, Σ_v for nodes $v \in \mathscr{N}$; and the exposures Ω_{vw} for edges $\ell = (vw) \in \mathscr{E}$. After n cascade steps, we identify $\mathscr{D}_n$, the set of defaulted banks, $\mathscr{S}_n$, the set of undefaulted banks that are under stress after n steps, and $\mathscr{D}_n^c \cap \mathscr{S}_n^c$ which contains the remaining undefaulted, unstressed banks.[4] In our model, banks do not recover from either default or stress during the crisis, so the sequences $\{\mathscr{D}_n\}_{n\in\mathbb{N}}$ and $\{\mathscr{D}_n \cup \mathscr{S}_n\}_{n\in\mathbb{N}}$ are non-decreasing.

To say that a bank v is defaulted at step n means that default shocks to step $n-1$ exceed its default buffer:

$$\mathscr{D}_n := \begin{cases} \{\Delta_v = 0\} & \text{for } n = 0 \\ \\ \left\{ v \;\middle|\; \sum_{w\in\mathscr{N}_v^-} \Omega_{wv}\xi_{wv}^{(n-1)} \geq \Delta_v \right\} & \text{for } n \geq 1 \ , \end{cases} \tag{4.2}$$

where the variables ξ indicate the fractional sizes of the various default shocks impacting v. Similarly, for any upstream neighbour $w \in \mathscr{N}_v^-$, we say that v is defaulted at step n *without regarding* w and write $v \in \mathscr{D}_n \not\!\!\circledR\, w$ under the conditions

$$\mathscr{D}_n \not\!\!\circledR\, w := \begin{cases} \{\Delta_v = 0\} \cap \mathscr{N}_w^+ & \text{for } n = 0 \\ \\ \left\{ v \in \mathscr{N}_w^+ \;\middle|\; \sum_{v'\in\mathscr{N}_v^-\setminus w} \Omega_{v'v}\xi_{v'v}^{(n-1)} \geq \Delta_v \right\} & \text{for } n \geq 1 \ . \end{cases} \tag{4.3}$$

Finally, to say v is stressed at step n means both that it is not yet defaulted and the stress shocks to step $n-1$ exceed the stress buffer, i.e. $\mathscr{S}_n = \mathscr{D}_n^c \cap \hat{\mathscr{S}}_n$ where

$$\hat{\mathscr{S}}_n := \begin{cases} \{\Sigma_v = 0\} & \text{for } n = 0 \\ \\ \left\{ v \;\middle|\; \sum_{w\in\mathscr{N}_v^+} \Omega_{vw}\zeta_{vw}^{(n-1)} \geq \Sigma_v \right\} & \text{for } n \geq 1 \ , \end{cases} \tag{4.4}$$

[4]For any set $\mathscr{B}$, $\mathscr{B}^c$ denotes its complement. For an *event* defined by some condition P, for example $v \in \mathscr{D}_n$, the indicator function for that event is written $\mathbb{1}(P)$.

where the variables ζ indicate the fractional sizes of the stress shocks impacting v.

Accounting for the fact that at the moment when v becomes stressed it reduces its interbank exposures, one can see that for $n \geq 1$ the fractions $\xi^{(n)}$ are given recursively by

$$\xi_{wv}^{(n)} := \begin{cases} \xi_{wv}^{(n-1)} & \text{when } w \in \mathscr{D}_{n-1} \cup \mathscr{D}_n^c \\ 1 & \text{when } w \in \mathscr{D}_n \setminus \mathscr{D}_{n-1} \text{ and } v \in \hat{\mathscr{S}}_{n-1}^c \\ 1-\lambda & \text{when } w \in \mathscr{D}_n \setminus \mathscr{D}_{n-1} \text{ and } v \in \hat{\mathscr{S}}_{n-1} \ , \end{cases} \tag{4.5}$$

with the initial values $\xi_{wv}^{(0)} = \mathbb{1}(w \in \mathscr{D}_0)$. Similarly, accounting for the assumption that defaulted creditors transmit a maximal stress impact, whereas stressed creditors only transmit a stress shock of a fraction λ of the interbank exposure, one has for $n \geq 0$

$$\zeta_{vw}^{(n)} := \begin{cases} 0 & \text{when } w \in \hat{\mathscr{S}}_n^c \cap (\mathscr{D}_n^c \not\rightarrow v) \\ \lambda & \text{when } w \in \hat{\mathscr{S}}_n \cap (\mathscr{D}_n^c \not\rightarrow v) \\ 1 & \text{when } w \in \mathscr{D}_n \not\rightarrow v \ . \end{cases} \tag{4.6}$$

The reader will be curious to see the $\not\rightarrow$ condition in (4.6). The rationale is that the $\not\rightarrow$ condition explicitly eliminates an apparent feedback where the stress shock to v from w at step n seems to depend on the default state of v at step $n-2$. This feedback is apparent, not real, because any stress shock to a defaulted bank is inconsequential.

The above prescription completely characterizes the cascade mapping of the system to itself. This mapping arrives at the fixed point that represents the end of the crisis in at most $2N$ steps when acting on a network of N banks. The next two sections are devoted to the derivation of two probabilistic versions of this double cascade mapping: first for the case of networks with random skeletons, and second for the case of networks with random balance sheets on a fixed skeleton.

4.3 Networks with Random Skeletons

Our model of a random financial network has three layers of structure: the skeleton (a random directed graph $(\mathscr{N}, \mathscr{E})$); the buffer random variables Δ_v, Σ_v, $v \in \mathscr{N}$ and the random

exposures $\Omega_\ell, \ell \in \mathscr{E}$. We give a complete specification of the distributional properties of these random variables.

The skeleton $(\mathscr{N}, \mathscr{E})$ is a random directed assortative configuration graph (ACG) of the type studied by Hurd (2015), generalizing the well-known undirected configuration random graph model introduced in Bender and Canfield (1978) and Bollobás (1980). The model is parametrized by N and the node and edge type distributions

$$P_{jk} = \mathbb{P}[v \in \mathscr{N}_{jk}], \quad Q_{kj} = \mathbb{P}[\ell \in \mathscr{E}_{kj}], \quad j,k \leq K \tag{4.7}$$

where for simplicity in the following we assume in and out-degrees j,k are bounded by a constant K. P and Q can be considered as bivariate distributions which have marginals:

$$P_k^+ := \sum_j P_{jk}, \quad P_j^- := \sum_k P_{jk}, \quad Q_k^+ := \sum_j Q_{kj}, \quad Q_j^- := \sum_k Q_{kj} \ . \tag{4.8}$$

Assortativity is defined to be the Pearson correlation of Q, considered as a bivariate probability distribution. Several studies of real financial networks, notably Bech and Atalay (2010), highlight the fact and relevance of their observed negative assortativity.

Definition 11 (The ACG Construction)**.** *The node and edge type probability laws P, Q are* consistent *if:*

$$\begin{aligned} z &:= \sum_j jP_j^- = \sum_k kP_k^+ \ , \\ Q_k^+ &= kP_k^+/z, \quad Q_j^- = jP_j^-/z, \ \textit{for all } j,k \leq K. \end{aligned} \tag{4.9}$$

Given consistent data N, P, Q, a random graph $(\mathscr{N}, \mathscr{E})$ with N nodes is sampled as follows:

1. *Draw a sequence of N node-type pairs $X = ((j_1,k_1), \ldots, (j_N,k_N))$ independently from P, and accept the draw if and only if it is feasible, i.e. $\sum_{v\in\mathscr{N}} j_v = \sum_{v\in\mathscr{N}} k_v$, and this defines the number of edges E that will result. Label the vth node with k_v out-stubs (each out-stub is a half-edge with an out-arrow, labelled by its degree k_v) and j_v in-stubs, labelled by their degree j_v. Define the partial sums*

$$u_j^- = \sum_v \mathbb{1}(j_v = j), \quad u_k^+ = \sum_v \mathbb{1}(k_v = k), \quad u_{jk} = \sum_v \mathbb{1}(j_v = j, k_v = k), \tag{4.10}$$

the number $e_k^+ = ku_k^+$ of k-stubs (out-stubs of degree k) and the number of j-stubs (in-stubs of degree j), $e_j^- = ju_j^-$.

2. *Conditioned on X, the result of Step 1, choose an arbitrary ordering ℓ^- and ℓ^+ of the E in-stubs and E out-stubs. The matching sequence, or "wiring", W of edges is selected by choosing a pair of permutations $\sigma, \tilde{\sigma} \in S(E)$ of the set $\mathscr{E}$. This determines the edge sequence $\ell = (\ell^- = \sigma(\ell), \ell^+ = \tilde{\sigma}(\ell))$ labelled by $\ell \in \mathscr{E}$, to which is attached a probability weighting factor*

$$\prod_{\ell \in \mathscr{E}} Q_{k_{\sigma(\ell)} j_{\tilde{\sigma}(\ell)}} \ . \tag{4.11}$$

The ACG simulation algorithm defines the required class of random graphs, but is infeasible because the acceptance condition in Step 1 is achieved only rarely, and drawing random permutations as in Step 2 is impractical. The paper Hurd (2015) proposes an approximate simulation algorithm that appears to work well in practice.

The non-negative default buffer random variables $\Delta_v, v \in \mathscr{N}$ have point masses at $x = 0$ that represent the bank initial default probability p_v^0. We assume that the distribution functions of Δ_v depend only on the type (j,k), and have the following form:

$$D_{jk}(x) = \mathbb{P}[\Delta_v \leq x \mid v \in \mathscr{N}_{jk}] \ ; \frac{d}{dx} D_{jk}(x) := p_{jk}^0 \delta_0(x) + d_{jk}(x) \ . \tag{4.12}$$

where $d_{jk}(x) \geq 0$ is a specified function with $\int_0^\infty d_{jk}(x)dx = 1 - p_{jk}^0$. Similarly, the stress buffer Σ_v has a point mass at $x = 0$ that represents this bank's initial stress probability q_v^0 and a distribution function that depends only on its node type (j,k). Thus the stress buffer distribution functions of nodes $v \in \mathscr{N}_{jk}$ have the following form:

$$S_{jk}(x) = \mathbb{P}[\Sigma_v \leq x, v \notin \mathscr{D}_0 \mid v \in \mathscr{N}_{jk}] \ ; \frac{d}{dx} S_{jk}(x) := q_{jk}^0 \delta_0(x) + s_{jk}(x) \ . \tag{4.13}$$

where $s_{jk}(x) \geq 0$ is a specified function with $\int_0^\infty s_{jk}(x)dx = 1 - p_{jk}^0 - q_{jk}^0$. The exposure random variables $\Omega_\ell, \ell \in \mathscr{E}$ are positive (i.e. there is zero probability to have a zero weight) and have distributions that depend only on the edge type (k,j). These can be specified by the distribution functions

$$W_{kj}(x) = \mathbb{P}[\Omega_\ell \leq x \mid \ell \in \mathscr{E}_{kj}]; \ \frac{d}{dx} W_{kj}(x) = w_{kj}(x) \ . \tag{4.14}$$

Finally, conditional on the random skeleton $(\mathscr{N}, \mathscr{E})$, the collection of random variables $\{\Delta_v, \Sigma_v, \Omega_\ell\}$ is assumed to be mutually independent.

It was proven in Hurd (2016) that the probability measure on random networks just defined has a property calledlocally tree-like independence (LTI) extending the well-known locally tree-like property of configuration graphs that cycles of any fixed finite length occur with an asymptotically zero density as the number of nodes N goes to infinity for fixed P, Q. The probabilistic analysis to follow rests on this extended type of independence that holds as $N \to \infty$:

The locally tree-like independence (LTI) property: Consider the double cascade model defined by the collection of random variables $(\mathscr{N}, \mathscr{E}, \Delta, \Sigma, \Omega)$. Let $\mathscr{N}_1, \mathscr{N}_2 \subset \mathscr{N}$ be any two subsets that share exactly one node $\mathscr{N}_1 \cap \mathscr{N}_2 = \{v\}$ and let X_1, X_2 be any pair of random variables where for each $i = 1, 2$, X_i is determined by the information on $\mathscr{N}_i$. Then X_1 and X_2 are conditionally independent, conditioned on the information $\Delta_v, \Sigma_v, j_v, k_v$ located at the node v.

The cascade mapping we now propose is based on the observation that a probability such as $\mathbb{P}[v \in \mathscr{D}^n]$ that depends on the nth order neighbourhood of the node v can be approximated iteratively, in terms of the probabilities of the states of the first order neighbouring nodes and edges of v at step $n-1$. The accuracy of such a scheme depends on the degree of dependence between the states of these first order neighbours of v. In ideal situations, there is no dependence at all, and the cascade mapping is exact. In our present model, there are two sources of residual dependence amongst these neighbours, and our approximation amounts to neglecting this residual dependence.

Let us define

$$\begin{aligned} p_{jk}^{(n)} &= \mathbb{P}\left[v \in \mathscr{D}_n \mid v \in \mathscr{N}_{jk}\right] \ , \\ q_{jk}^{(n)} &= \mathbb{P}\left[v \in \mathscr{S}_n \mid v \in \mathscr{N}_{jk}\right] \ , \\ \widetilde{p}_{jk}^{(n)} &= \mathbb{P}\left[v \in \mathscr{D}_n \not\Join w \mid v \in \mathscr{N}_{jk} \cap \mathscr{N}_w^+\right] \ , \\ \widehat{q}_{jk}^{(n)} &= \mathbb{P}\left[v \in \hat{\mathscr{S}}_n \mid v \in \mathscr{N}_{jk}\right] \ . \end{aligned} \tag{4.15}$$

where $p_{jk}^{(n)}$ is the probability that a (j,k)-node has defaulted by time n, $q_{jk}^{(n)}$ is the probability

that a (j,k)-node is stressed at time n, $\widetilde{p}_{jk}^{(n)}$ is the probability that a (j,k)-node has defaulted by time n, without regarding of one of its in-neighbours, and $\widehat{q}_{jk}^{(n)}$ is the probability that a (j,k)-node satisfies the stress condition (4.4) at time n. Our aim is to compute $p_{jk}^{(n)}, \widehat{q}_{jk}^{(n)}, \widetilde{p}_{jk}^{(n)}$, plus an auxiliary quantity

$$t_{kj}^{(n)} = \mathbb{P}\left[\xi_{wv}^{(n)} = 1 \mid (w,v) \in \mathscr{E}_{kj}\right], \tag{4.16}$$

recursively over n. It is also convenient to define

$$\begin{aligned} p_k^{(n)} &= \mathbb{P}[v \in \mathscr{D}_n \mid k_v = k] = \sum_j p_{jk}^{(n)}\ P_{j|k}\ , \\ \widetilde{p}_j^{(n)} &= \mathbb{P}[v \in \mathscr{D}_n \circledR\ w \mid j_v = j, v \in \mathscr{N}_w^+] = \sum_k \widetilde{p}_{jk}^{(n)}\ P_{k|j}\ , \\ \widehat{q}_j^{(n)} &= \mathbb{P}[v \in \hat{\mathscr{S}}_n \mid j_v = j] = \sum_k \widehat{q}_{jk}^{(n)}\ P_{k|j}\ , \end{aligned} \tag{4.17}$$

where $P_{j|k} := \frac{P_{jk}}{P_k^+}, P_{k|j} := \frac{P_{jk}}{P_j^-}$.

Our computations will rely on two facts. The first is that if X, Y are two independent random variables with probability density functions (PDFs) $f_X(x) = F_X'(x), f_Y(y) = F_Y'(y)$, then

$$\begin{aligned} \mathbb{P}\left[X \geq Y\right] = \mathbb{E}\left[\mathbb{1}(X \geq Y)\right] &= \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{1}(X \geq Y) f_X(x) f_Y(y)\,\mathrm{d}x\mathrm{d}y \\ &= \int_{\mathbb{R}} F_Y(x) f_X(x)\,\mathrm{d}x = \left\langle F_Y, f_X \right\rangle . \end{aligned} \tag{4.18}$$

In general, the Hermitian inner product on $\mathbb{R}$ is defined as $\langle f, g\rangle = \int_{-\infty}^{\infty} \bar{f}(x) g(x)\,\mathrm{d}x$, but here, both operands are real functions and the conjugate operator disappears. A second fact is that if $X_1, X_2, \cdots, X_n$ are n independent random variables with PDFs f_{X_i}, then the PDF of the sum $X = X_1 + X_2 + \cdots + X_n$ is the convolution

$$f_X = f_{X_1} \circledast f_{X_2} \circledast \cdots \circledast f_{X_n} = \circledast_{k=1}^n f_{X_k}\ , \tag{4.19}$$

where the convolution product of two functions is the function defined by $(f \circledast g)(x) = \int_{\mathbb{R}} f(y) g(x-y) dy$. For convolution powers, we write $\circledast_{k=1}^n f_X = f_X^{\circledast n}$.

The following cascade mapping provides a closed set of recursive formulas for $n \geq 1$ starting with $p_{jk}^{(0)} = \widetilde{p}_{jk}^{(0)}, \widehat{q}_{jk}^{(0)}, t_{kj}^{(0)} = p_k^{(0)}$ determined by the initial shock and stress probabilities.

Cascade Mapping: For any $n \geq 1$, suppose $p_{jk}^{(n-1)}, \widehat{q}_{jk}^{(n-1)}, \widetilde{p}_{jk}^{(n-1)}, t_{kj}^{(n-1)}$ are known. The cascade mapping gives the quantities $p_{jk}^{(n)}, \widehat{q}_{jk}^{(n)}, \widetilde{p}_{jk}^{(n)}, t_{jk}^{(n)}$ recursively by

$$
\begin{aligned}
p_{jk}^{(n)} &= \left\langle D_{jk}, \left(g_j^{(n-1)}\right)^{\circledast j} \right\rangle , && (4.20)\\
\widetilde{p}_{jk}^{(n)} &= \left\langle D_{jk}, \left(g_j^{(n-1)}\right)^{\circledast j-1} \right\rangle , && (4.21)\\
\widehat{q}_{jk}^{(n)} &= \left\langle S_{jk}, \left(h_k^{(n-1)}\right)^{\circledast k} \right\rangle , && (4.22)\\
t_{kj}^{(n)} &= t_{kj}^{(n-1)} + (p_k^{(n)} - p_k^{(n-1)})(1 - \widehat{q}_j^{(n-1)}) , && (4.23)
\end{aligned}
$$

where D_{jk} and S_{jk} are defined in (4.12) and (4.13), respectively, and we use (4.17) to compute $p_{k'}^{(n-1)}, \widehat{q}_{j'}^{(n-1)}, \widetilde{p}_{j'}^{(n-1)}$. Moreover, the stress probabilities $q_{jk}^{(n)}$ are determined by

$$
\begin{aligned}
1 - q_{jk}^{(n)} - p_{jk}^{(n)} &= \mathbb{P}[v \in \hat{\mathscr{S}}_n^c \cap \mathscr{D}_n^c \mid v \in \mathscr{N}_{jk}] \\
&= \left(1 - \widehat{q}_{jk}^{(n)}\right) \left(1 - \left\langle D_{jk}, \left(\tilde{g}_j^{(n-1)}\right)^{\circledast j} \right\rangle\right) . && (4.24)
\end{aligned}
$$

The probability distribution functions in these formulas are also computed recursively:

$$
\begin{aligned}
g_j^{(n-1)}(x) &= \sum_{k'} \Big[(1 - p_{k'}^{(n-1)})\delta_0(x) + t_{k'j}^{(n-1)} w_{k'j}(x) \\
&\quad + (p_{k'}^{(n-1)} - t_{k'j}^{(n-1)}) \cdot \frac{1}{1-\lambda} w_{k'j}(x/(1-\lambda)) \Big] \cdot Q_{k'|j} , && (4.25)\\
h_k^{(n-1)}(x) &= \sum_{j'} \Big[(1 - \widehat{q}_{j'}^{(n-1)})(1 - \widetilde{p}_{j'}^{(n-1)}))\delta_0(x) + \widetilde{p}_{j'}^{(n-1)} w_{kj'}(x) \\
&\quad + \widehat{q}_{j'}^{(n-1)}(1 - \widetilde{p}_{j'}^{(n-1)}) \cdot \frac{1}{\lambda} w_{kj'}(x/\lambda) \Big] \cdot Q_{j'|k} , && (4.26)\\
\tilde{g}_j^{(n-1)}(x) &= \sum_{k'} \Big[(1 - p_{k'}^{(n-1)})\delta_0(x) + p_{k'}^{(n-1)} w_{k'j}(x) \Big] \cdot Q_{k'|j} && (4.27)
\end{aligned}
$$

with $Q_{k|j} = \frac{Q_{kj}}{Q_j^-}, Q_{j|k} = \frac{Q_{kj}}{Q_k^+}$.

Justification: To justify (4.20) we need to suppose that the collection of random variables

$\Omega_{wv}\xi_{wv}^{(n-1)}$ for different $w \in \mathscr{N}_v^-$ are mutually conditionally independent. However, there are two sources of dependence: the first is the usual breaking of the LTI property for finite N due to cycles in the skeleton, the second, specific to this model, is the dependence on whether v is in $\mathscr{S}_{n-1}$. Our approximation is to neglect both sources of dependence, allowing the use of (4.18) to give

$$\begin{aligned} p_{jk}^{(n)} &= \mathbb{P}[\Delta_v \le \sum_{w \in \mathscr{N}_v^-} \Omega_{wv}\xi_{wv}^{(n-1)} \mid v \in \mathscr{N}_{jk}] \\ &= \left\langle D_{jk}, \left(g_j^{(n-1)}\right)^{\circledast j} \right\rangle \ . \end{aligned} \tag{4.28}$$

where

$$g_j^{(n-1)}(x) = \sum_{k'} Q_{k'|j} \frac{d}{dx}\mathbb{P}[\Omega_{wv}\xi_{wv}^{(n-1)} \le x \mid v \in \mathscr{N}_{jk}, w \in \mathscr{N}_v^-, k_w = k'] \ . \tag{4.29}$$

Under the conditions $v \in \mathscr{N}_{jk}, w \in \mathscr{N}_v^-, k_w = k'$, the events
$\{\xi_{wv}^{(n-1)} = 0\},\ \{\xi_{wv}^{(n-1)} = 1\}, \{\xi_{wv}^{(n-1)} = 1-\lambda\}$ have conditional probabilities
$1 - p_{k'}^{(n-1)}$, $t_{k'j}^{(n-1)}$, $(p_{k'}^{(n-1)} - t_{k'j}^{(n-1)})$ respectively and the three events are conditionally independent of Ω_{wv} assuming the LTI property. These facts lead to (4.25).

To verify (4.21), we use (4.3) instead of (4.2) and follow these same steps. To verify (4.22), we use (4.4), (4.15) and (4.18) to give the formula

$$\widehat{q}_{jk}^{(n)} = \mathbb{P}[\Sigma_v \le \sum_{w \in \mathscr{N}_v^+} \Omega_{vw}\hat{\zeta}_{vw}^{(n)} \mid v \in \mathscr{N}_{jk}] = \left\langle S_{jk}, \left(h_k^{(n-1)}(x)\right)^{\circledast k} \right\rangle \tag{4.30}$$

where

$$h_k^{(n-1)}(x) = \sum_{j'} \frac{d}{dx}\mathbb{P}[\Omega_{vw}\hat{\zeta}_{vw}^{(n-1)} \le x \mid v \in \mathscr{N}_{jk}, w \in \mathscr{N}_v^+, j_w = j'] \ Q_{j'|k} \ . \tag{4.31}$$

To verify (4.26), note that under the conditions $k_v = k, w \in \mathscr{N}_v^+, j_w = j'$, the events $\{\hat{\zeta}_{vw}^{(n-1)} = 0\}, \{\hat{\zeta}_{vw}^{(n-1)} = 1\}, \{\hat{\zeta}_{vw}^{(n-1)} = \lambda\}$ are equivalent to the events
$\{w \in \hat{\mathscr{S}}_{n-1}^c \cap \mathscr{D}_{n-1}^c \not\circledR v\},\ \{w \in \mathscr{D}_{n-1} \not\circledR v\}, \{w \in \mathscr{S}_{n-1}\}$ and hence have conditional probabilities $(1-\widehat{q}_{j'}^{(n-1)})(1-\widetilde{p}_{j'}^{(n-1)})$, $\widetilde{p}_{j'}^{(n-1)}, \widehat{q}_{j'}^{(n-1)}(1-\widetilde{p}_{j'}^{(n-1)})$ respectively. To verify (4.23), we apply the LTI property again to compute the conditional probability of the disjoint

union $\{\xi_{wv}^{(n)} = 1\} = \{\xi_{wv}^{(n-1)} = 1\} \cup \{w \in \mathscr{D}_n \cap \mathscr{D}_{n-1}^c, v \in \hat{\mathscr{S}}_{n-1}^c\}$ defined by (4.5).

Finally, to verify (4.24), we note that $\hat{\mathscr{S}}_n^c \cap \mathscr{D}_n^c = \hat{\mathscr{S}}_n^c \cap \{\Delta_v > \sum_{w \in \mathscr{N}_v^-} \Omega_{wv} \mathbf{1}_{\{w \in \mathscr{D}_{n-1}\}}\}$. By LTI, these last two events are independent conditioned on $v \in \mathscr{N}_{jk}$, and the required formula results by following the steps taken to prove (4.20).

Remark 2. *We see from this argument that the cascade mapping has two sources of error. The first, stemming from cycles in the skeleton that cause finite size corrections to the LTI property, is familiar, and known to behave like $O(N^{-1})$ on configuration graphs as the network size N grows to infinity, and to be zero on finite tree graphs. The second source of error is less familiar and stems from dependence due to the assumption that the default shocks that impact v are diminished by a factor $1-\lambda$ when v becomes stressed, that is, they depend on whether v is stressed or not.*

The iterates of this cascade mapping converge as $n \to \infty$ to a fixed set of probabilities that represent the eventual state of the system at the end of the cascade. These final probabilities can be used to measure the overall impact of the crisis. For example, the expected number of eventually defaulted and stressed banks are

$$\text{Expected number of defaulted banks} \;=\; N \sum_{jk} P_{jk}\; p_{jk}^{(\infty)} = N \sum_{k} P_k^+\; p_k^{(\infty)} \tag{4.32}$$

$$\text{Expected number of stressed banks} \;=\; N \sum_{jk} P_{jk}\; q_{jk}^{(\infty)}\ . \tag{4.33}$$

4.4 Networks with Fixed Skeletons

The goal of the present section is to derive approximate probabilistic formulas describing the double cascade on a network where the skeleton is actually known (deterministic) and finite, while the buffers and weights are random. This analysis will allow us to address systemic risk in tractable models of real observed financial networks, without the need for Monte Carlo simulations.

Let $A = A_{vv'}, v, v' \in \mathscr{N}$ be the nonsymmetric adjacency matrix of the fixed directed graph $(\mathscr{N}, \mathscr{E})$. We number the nodes in $\mathscr{N}$ by $v = 1, 2, \ldots, N$ and the links by $\ell = 1, 2, \ldots, E$ where $E = \sum_{1 \le v, v' \le N} A_{vv'}$. The buffer random variables Δ_v, Σ_v at each node are assumed to have a mass $p_v^{(0)}, q_v^{(0)}$ at 0 (representing the initial default and stress

probabilities) and continuous support with density functions $d(x), s_v(x)$ on the positive reals. The edge weights $\Omega_\ell, \ell \in \mathscr{E}$ have continuous support with densities $w_\ell(x)$ on the positive reals but no mass at 0. The random variables $\{\Delta_v, \Sigma_v, \Omega_\ell\}, v \in \mathscr{N}, \ell \in \mathscr{E}$ are assumed to be an independent collection.

The aim of this section is to use the LTI property as an approximation to derive formulas for the marginal likelihoods $p_v^{(\infty)}, q_v^{(\infty)}$ for the eventual default and stress of all individual nodes, as well as the possibility to compute formulas for more detailed systemic quantities. This approximation is not exact for the same two reasons as before: there is dependence between the default shocks hitting a bank v due to the stress response, and when there are cycles in the skeleton. In general, when the skeleton is a single random realization from a configuration graph ensemble, we expect the LTI approximation to get better with increasing N. We now present an approximate analysis, paralleling the previous section, of the sequence of probabilities

$$\begin{aligned} p_v^{(n)} &= \mathbb{P}[v \in \mathscr{D}_n] \ , \\ q_v^{(n)} &= \mathbb{P}[v \in \mathscr{S}_n] \ , \\ \widetilde{p}_{wv}^{(n)} &= \mathbb{P}[v \in \mathscr{D}_n \not\ni w] \ , \\ \widehat{q}_v^{(n)} &= \mathbb{P}[v \in \hat{\mathscr{S}}_n] \ , \end{aligned} \tag{4.34}$$

for each node v or edge wv. For the same reason as before we need in addition to track

$$t_{wv}^{(n)} = \mathbb{P}\left[\xi_{wv}^{(n)} = 1 \mid v \in \mathscr{N}_w^+\right] \ . \tag{4.35}$$

Inductively, we have

$$p_v^{(n)} = \left\langle D_v, \circledast_{v' \in \mathscr{N}_v^-} \left(g_{v'v}^{(n-1)} \right) \right\rangle \ , \tag{4.36}$$

$$\widetilde{p}_{wv}^{(n)} = \left\langle D_v, \circledast_{v' \in \mathscr{N}_v^- \setminus w} \left(g_{v'v}^{(n-1)} \right) \right\rangle \ , \tag{4.37}$$

$$\widehat{q}_v^{(n)} = \left\langle S_v, \circledast_{v' \in \mathscr{N}_v^+} \left(h_{vv'}^{(n-1)} \right) \right\rangle \ , \tag{4.38}$$

$$t_{wv}^{(n)} = t_{wv}^{(n-1)} + (p_w^{(n)} - p_w^{(n-1)})(1 - \widehat{q}_v^{(n-1)}) \ , \tag{4.39}$$

$$q_v^{(n)} = 1 - p_v^{(n)} - \left(1 - \widehat{q}_v^{(n)}\right) \left\langle D_v, \circledast_{v' \in \mathscr{N}_v^-} \left(\tilde{g}_{v'v}^{(n-1)} \right) \right\rangle \ . \tag{4.40}$$

The PDFs can be computed as before under the LTI approximation:

$$\begin{aligned} g_{wv}^{(n-1)}(x) &= (1-p_w^{(n-1)})\delta_0(x) + t_{wv}^{(n-1)} w_{wv}(x) \\ &\quad + (p_w^{(n-1)} - t_{wv}^{(n-1)}) \cdot \frac{1}{1-\lambda} w_{wv}(x/1-\lambda) \ , \qquad (4.41) \\ h_{vw}^{(n-1)}(x) &= (1-\widehat{q}_w^{(n-1)})(1-\widetilde{p}_{vw}^{(n-1)})\delta_0(x) + \widetilde{p}_{vw}^{(n-1)} w_{vw}(x) \\ &\quad + \widehat{q}_w^{(n-1)}(1-\widetilde{p}_{vw}^{(n-1)}) \cdot \frac{1}{\lambda} w_{vw}(x/\lambda) \ , \qquad (4.42) \\ \tilde{g}_{wv}^{(n-1)}(x) &= (1-p_w^{(n-1)})\delta_0(x) + p_{wv}^{(n-1)} w_{wv}(x) \ . \qquad (4.43) \end{aligned}$$

4.5 Numerical Experiments

In this section, we report briefly on numerical experiments that illustrate the methods developed in this paper. Firstly, we aim to convince the reader that the LTI method correctly computes the fixed point of the double cascade mapping in networks with large values of N by cross validating it using an independently coded Monte Carlo (MC) implementation. For efficiency, the LTI implementation uses a Fast Fourier Transform approach to compute the convolutions in (4.20)-(4.23). This technique was developed in Hurd and Gleeson (2013) and is sketched in Appendix A. Secondly, we will show how the LTI method can lead to answers to questions about the nature of systemic risk, particularly the intertwining of stress and default. Thirdly, we shall show how the method performs in a challenging stylized network specified to reflect the complex characteristics of a 2011 dataset on the network of 90 most systemically important banks in the European Union.

4.5.1 Experiment 1: Verifying the LTI Method

This experiment aims to verify that the LTI method performs as expected when applied to a stylized financial network whose specification is similar to that given in Gai and Kapadia (2010a). It consists of a random directed Poisson skeleton with $N = 20000$ nodes and mean degree $z = 10$, where each node v can be viewed as a bank with a default buffer $\Delta_v = 0.04$ and stress buffer $\Sigma_v = 0.035$. Unlike the deterministic interbank exposures used in Gai and Kapadia (2010a), the weight Ω_ℓ of an edge ℓ is taken from a log-normal distribution with mean $\mu_\ell = 0.2 j_\ell^{-1}$, and standard deviation $0.383\mu_\ell$. Note that this specification makes the

exposure size dependent on the lending bank. An initial shock is applied to the network that causes each bank to default with a 1% probability.

We compare the final fractions of defaulted and stressed banks as computed using MC simulation with 1000 realizations and the LTI analytic formulas in (4.32) and (4.33). It is simple to generate a directed Poisson random graph of size N with mean degree $z > 0$: one simply selects directed edges independently from all $N(N-1)$ potential edges, each with probability $p = z/(N-1)$. The resultant bi-degree distribution is a product of binomials, $P_{jk} = \mathbb{P}[v \in \mathscr{N}_{jk}] = \text{Bin}(N-1, p, j) \times \text{Bin}(N-1, p, k)$, which for large N nodes is approximately a product of Poisson(z) distributions.

Figure 4.3 plots the results as functions of the stress response parameter λ, with error bars that represent the 10^{th} and 90^{th} percentile of the MC result. It shows the expected agreement between MC and LTI analytics, with discrepancies that can be attributed to finite N effects present in the MC simulations.

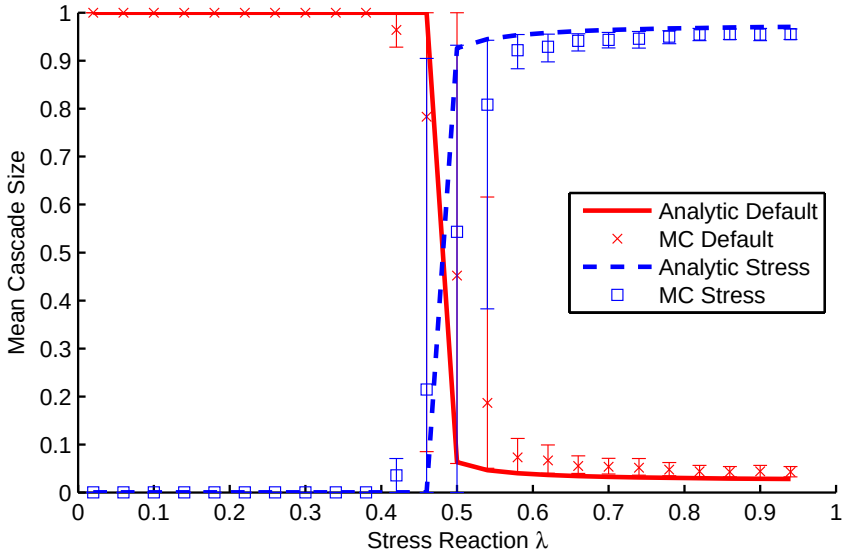


Figure 4.3: Experiment 1. The mean default and stress cascade sizes computed by MC simulations (symbols) and LTI analytics (lines). The skeleton is a Poisson directed network with $N = 20000$ nodes and mean degree $z = 10$. All banks have default buffers $\Delta_v = 0.04$ and stress buffers $\Sigma_v = 0.035$. An edge weight Ω_ℓ of an edge ℓ is taken from a log-normal distribution with mean $\mu_\ell = 0.2 j_\ell^{-1}$ and standard deviation $0.383\mu_\ell$. Error bars indicate the 10^{th} and 90^{th} percentiles of the MC result.

One fundamental property of our model is clearly shown in this experiment: stress and default are negatively correlated. This fact can be explained by the stress response which enables banks to react to liquidity shocks before they default, by reducing their interbank exposures. This response creates yet more stress, but leads to a more resilient network. The "knife-edge" property of default cascades is also clearly shown: In the model parametrization we chose, a very small increase in λ dramatically alters the stability of the network. We also note that MC error bars are very large near the knife-edge.

4.5.2 Experiment 2: A Stylized Poisson Network

The next experiment focuses again on Poisson networks, with the aim to better understand the effects of various parameters on network resilience. In general, we continue to find confirmation that the LTI results accurately reflect observations from MC simulations.

Experiment 2A: Effects of Default and Stress Buffers

We consider how the parametrized financial network of Experiment 1 in a default-susceptible state with $\lambda = 0.5$ can be made resilient to random shocks by either increasing the default buffers or decreasing the stress buffers. Such changes, for example, can be prompted by financial regulators.

In Fig. 4.4(a) we illustrate how the change in default buffers affects the default cascade size. We observe a very fast transition from 100% default cascade to almost no default as Δ increases over the interval $[0.04, 0.045]$. This knife-edge property is observable in both the LTI analytics and in the MC simulations. Note again that the MC error bars, representing the 10^{th} and 90^{th} percentiles, become very large near the knife-edge.

In Fig. 4.4(b), we examine the influence of the stress buffer on the default cascade size. If stress buffers are reduced, banks start to react to stress shocks more quickly, which can in turn dramatically reduce default cascade risk in the network. Figures 4.4(a) and (b) show that there may be several approaches to improving network resilience.

Experiment 2B: Effects of Graph Connectivity and Stress Response

Aside from mandating changes to the behaviour of banks during or prior to a crisis by imposing constraints on stress and default buffers, regulators can also influence the shape of the financial network as a whole. In Experiment 2B, we demonstrate how systemic risk is influenced by the skeleton itself. To this end, we calculate the sizes of default and stress cascades in a directed Poisson network as a function of the connectivity parameter z and the stress response λ. In our simple model specification, the mean degree z is the only parameter that controls the shape of the skeleton, whereas in a more realistic modelling approach the skeleton may have many more parameters.

In this experiment, we increased the model complexity by assuming each bank has a random default buffer taken from a log-normal distribution with mean 0.18 and standard deviation 0.18, and a stress buffer from an independent log-normal distribution with mean 0.12 and standard deviation 0.12. The edge weights Ω_ℓ come from a log-normal distribution with mean and standard deviation proportional to $(j_\ell k_\ell)^{-0.5}$, with the average edge weight on the entire network equal to 1. Once again we apply an initial shock so that each bank has 1% chance of defaulting initially.

In Figs. 4.5(a) and (b), we respectively show the mean sizes of default and stress

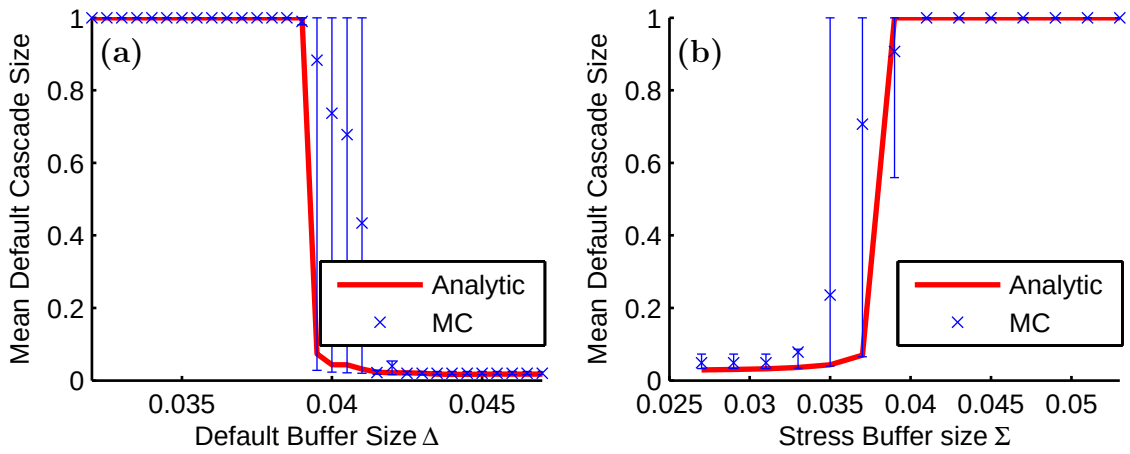


Figure 4.4: Experiment 2A. The mean default cascade size as a function of (a) default buffer Δ and (b) stress buffer Σ. The LTI analytic approximation (lines) correctly predicts MC simulations results (symbols). Here $\lambda = 0.5$ and other parameters are chosen as in Experiment 1 (Fig. 4.3).

cascades as functions of network mean degree z and the stress response λ. For clarity of the graphics, we do not show the MC simulations, as they agree, in a similar fashion as earlier, with the LTI analytics . Again, in these plots we notice the strong anti-correlation between stress and default probabilities as we vary z and λ. It is also interesting to observe that the final level of stress is not monotonic in the connectivity parameter z.

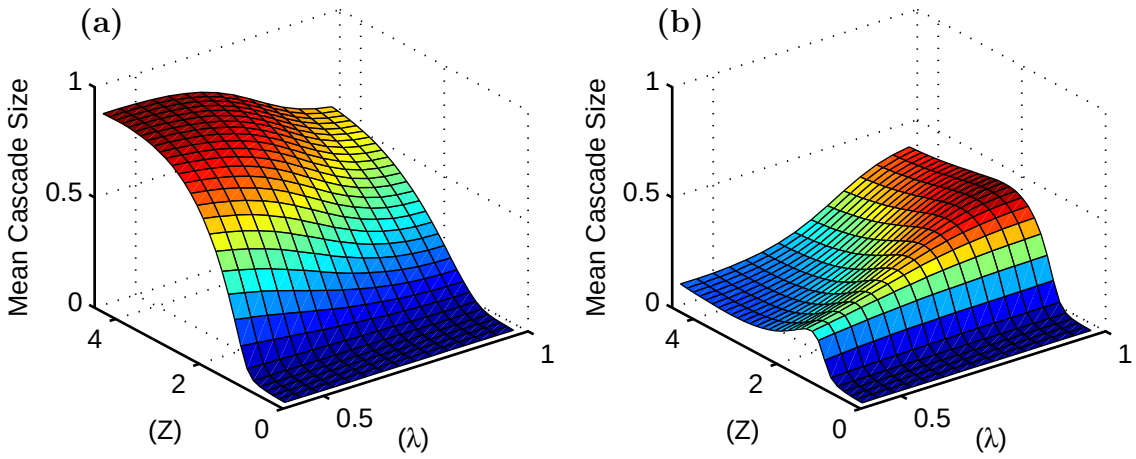


Figure 4.5: Experiment 2B. The sizes of the mean default cascade (a) and the mean stress cascade (b) on a directed Poisson network as functions of the network mean degree z and stress response parameter λ. Here the values of edge weights (interbank exposures), default and stress buffers are taken from log-normal distributions as specified in the text. Other parameters are the same as in Fig. 4.3.

4.5.3 Experiment 3: An EU-Inspired Network with 90 Nodes

It is well known that Poisson random graphs are an inadequate description of real economic networks, so we are interested in having a rough picture of the systemic risk of actual financial networks. In Experiment 3, we consider a single realization of a 90 node graph that aims to capture stylized features of the interbank network of the European Union. We computed cascade dynamics on this network using both the MC simulations and the LTI analytic method of Section 4.4 for networks with a fixed skeleton. As a preliminary step (not reported), we validated the consistency of the LTI analysis by verifying that the two methods agree as expected on a number of tree networks.

Numerous studies of real-world financial networks, notably Bech and Atalay (2010) and Cont et al. (2010), have observed their highly heterogeneous structure and concluded that in and out-degrees have fat tailed distributions, as do exposure sizes, and presumably, buffers. Our schematic model of 90 EU banks was designed to capture these basic statistical features, and as well to fit aggregated statistics from data published on the 2011 ECB stress testing of systemically important banks in the European Union. We show the skeleton of our stylized network in Figure 4.6. The details of its construction and the specifications for buffer and exposure distributions are given in Appendix B.

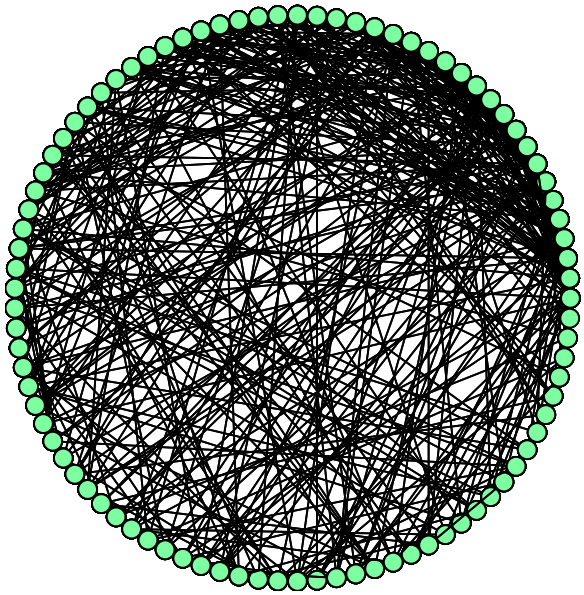


Figure 4.6: A representation of the skeleton of the 90 bank network of Experiment 3. We do not show the edge directions here to avoid cluttering the figure. The nodes are plotted with total degree increasing in the clockwise direction, with the minimally connected bank being the rightmost node.

In the first part of the experiment, we calculate the mean sizes of stress and default cascades that start from the default of one random bank in the EU network. Figure 4.7(a) shows the mean default and stress cascade sizes as functions of the stress response parameter λ. This graph shows no evidence of cascading (beyond the initially defaulted bank and the stress it causes to its immediate neighbors), demonstrating that the MC and analytic computations agree that the EU network in June 2011 was resilient to such a shock.

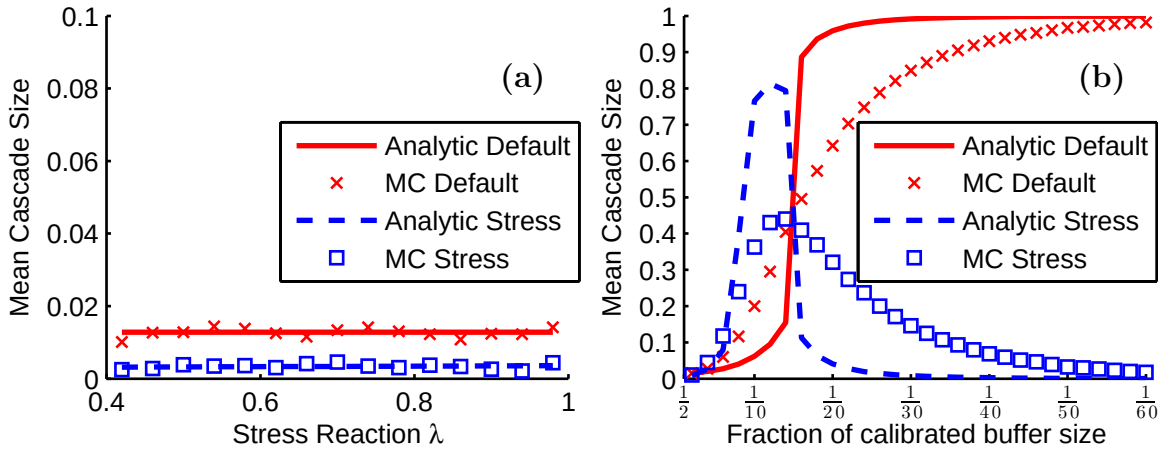


Figure 4.7: Experiment 3: sizes of default and stress cascades on a stylized EU interbank network starting from the default of a single randomly chosen bank. (a) Cascade sizes for various values of the stress response λ. The EU financial system at the time of the 2011 stress testing exercise appears to be resilient to single-bank shocks. (b) The same system as in (a) where a dire pre-shock crisis has reduced the bank default buffers to 10% of their original value, and stress buffers to a fraction of their value indicated on the x-axis.

To move the EU network to a knife-edge situation where a large scale double cascade can be triggered by the default of a single bank, we found it was necessary to imagine a dire crisis where prior to the default shock, the Core Tier 1 Capital of all institutions (i.e. their default buffers) has been decimated to 10% of their initial amount. Furthermore, each bank's stress buffer is also reduced to a certain fraction of its original value. In Fig. 4.7(b) we take $\lambda = 0.7$ and show the cascade sizes versus the remaining fraction of stress buffers.

The MC simulations show that for smaller shocks the size of a stress cascade first dominate default cascades: the banking system becomes highly illiquid, but most banks are able to protect themselves from default. Beyond a certain critical value of the global shock, however, default cascades take over stress cascades, as banks are hit by a shock so high that they default immediately, without the possibility to protect themselves as they would if they got first in the stressed state. Our analytical approach is able to capture the essence of these mechanisms, identifying correctly the critical point where default cascades overtake stress cascades. The quantitative size of the cascades away from this critical point, however, is less well matched by the analytical model, probably because of the presence of short cycles in the skeleton. As cycles tend to slow down the cascades, the transitions are

smoother in the MC. It is quite interesting that, in spite of our two approximations, our model is able to capture the general mechanisms: a peak in the number of stressed banks is a signal that can indicate quite precisely the point where systemic default takes over in a banking network.

4.6 Asset Firesale

One drawback of the base model provided in the previous section is the lack of a direct linkage between the stress (funding illiquidity) and solvency. As a result, an institution which has fully withdrawn from the financial network is now immune to the default cascade. Realistically, there are many additional pathways of contagion that do not involve transmission from immediate neighbours. One such mechanism is the firesale effect.

In this section, we build on the base model by using a firesale contagion based on the equilibrium price model of Cifuentes et al. (2005), and study the additional systemic risk posed on the EU network by this new channel of contagion. The foundation of this firesale effect is the macro-economic impact on the price of the fixed assets of each institution in the system. In general, when a financial institution defaults, its fixed assets will be liquidated to the market at some point, in which case additional supply is injected into the market. moreover, when a financial institution chooses to hoard cash, it decreases the demand in the market for these fixed assets, and further reduces their marked to market price.

4.6.1 Model Setup

We restrict our setup in the following sections to the finite network constructed in the previous section based on the major financial institutions in the EU. Let us once again define $A = A_{vv'}, v, v' \in \mathscr{N}$ to be the nonsymmetric adjacency matrix of the fixed directed graph $(\mathscr{N}, \mathscr{E})$, which represents the EU network. We number the nodes in $\mathscr{N}$ by $v = 1, 2, \ldots, N$ and the links by $\ell = 1, 2, \ldots, E$ where $E = \sum_{1 \leq v, v' \leq N} A_{vv'}$. The buffer random variables Δ_v, Σ_v at each node are assumed to have a mass $p_v^{(0)}, q_v^{(0)}$ at 0 (representing the initial default and stress probabilities) and continuous support with density functions $d_v(x), s_v(x)$ on the positive reals. The edge weights $\Omega_\ell, \ell \in \mathscr{E}$ have continuous support with densities $w_\ell(x)$ on

the positive reals but no mass at 0. The random variables $\{\Delta_v, \Sigma_v, \Omega_\ell\}, v \in \mathscr{N}, \ell \in \mathscr{E}$ are assumed to be a conditionally independent collection.

In addition to the previously defined default and stress cascade, we implement the asset firesale cascade mechanism in the following way. The fixed asset portion A_v^F of all banks in the network is assumed to consist of a single asset type. The aggregate supply of this illiquid asset across the network as a function of price (p) is denoted by $s(p) = \alpha \sum_{v|\Delta_v \leq 0} A_v^F + s_0$, where α is a positive constant and s_0 is the precrisis supply level. The aggregate demand of this illiquid asset across the network is denoted $d(p) = d_0 - \beta \sum_{v|\Sigma_v \leq 0} \lambda A_v^{IB}$, where β is a positive constant and d_0 is the precrisis demand level. After the onset of the crisis, banks will only liquidate their fixed assets if bankruptcy occurs, resulting in additional supply of this illiquid asset to the market. Similarly, if interbank lendings are withdrawn in the network during the crisis, this removes the possibility of using these funds to purchase illiquid assets and resulting in reduced demand of the illiquid asset. Combining the above two equations, we denote the excess supply function of the network during the crisis as $\tilde{s} = s - s_0 + d_0 - d$. Finally, there exist external traders trading in the illiquid asset outside of the network, who are purely price takers and are assumed to clear the market with the following inverse demand function during the crisis:

$$p = e^{-\tilde{s}} = e^{-\alpha \sum_{v|\Delta_v \leq 0} A_v^F - \beta \sum_{v|\Sigma_v \leq 0}}. \tag{4.44}$$

We make the assumption that the price throughout the crisis cannot exceed the starting price of 1. α and β are chosen so that, if all FIs within the network were to put their fixed assets into liquidation, a price drop of 15% would occur, and similarly if all interbank loans were withdrawn, then a price drop of 10% would occur.

In addition, we will modify the timing and behaviour assumptions and cascade mechanism in order to precisely define our asset firesale cascade as part of the general cascade mechanism.

Timing and Bank Behaviour Assumptions:

Assumption 2 is replaced with $2'$.

$2'$. Total amount of fixed asset belonging to defaulted banks and the total value of interbank loans recalled by stressed banks are updated. Balance sheets are recomputed daily

on a mark-to-market basis and banks respond daily on the basis of their newly computed balance sheets. All external cash flows, interest payments, and liability price changes are ignored throughout the crisis. The fixed asset price follows the price given by the inverse demand function subject to the total excess supply of the network as of the start of the current day.

Cascade Mechanism:

An additional equation to describe the updates to the default buffer is added and equation (4.2) is replaced with the following, where the default buffer Δ is replaced with the default buffer at cascade step n, $\Delta^{(n)}$ as defined below.

$$\Delta_v^{(n)} = \Delta_v - A_v^F(1 - e^{-\alpha \sum_{v \in \mathscr{D}_n} A_v^F - \beta \sum_{v \in \mathscr{S}_n} \lambda A_v^{IB}}) \tag{4.45}$$

$$\mathscr{D}_n := \begin{cases} \{\Delta_v = 0\} & \text{for } n = 0 \\ \\ \left\{ v \;\middle|\; \sum_{w \in \mathscr{N}_v^-} \Omega_{wv} \xi_{wv}^{(n-1)} \geq \Delta_v^{(n-1)} \right\} & \text{for } n \geq 1 \ , \end{cases} \tag{4.46}$$

It should be noted that based on the above equations and behaviour assumptions, our asset firesale mechanism solely impacts the default buffer of banks, and any additional defaults which may occur are implemented through the existing default cascade mechanism. The asset firesale mechanism also links the impact from the stress cascade to the default cascade through the term representing the aggregate demand of the network in the fixed asset price inverse demand function.

4.6.2 Systemic Risk Measures

The safety of the financial system cannot be judged on the expected default size alone. We must also measure the rare events, since the study of risk is the study of tail events. We will now focus on the precisely defining measures for the tail systemic risk of this financial system system.

Value at Risk (VaR) is a measure of the riskiness of an investment portfolio. It measures the amount in dollar values at risk limited to a specified time frame T. The loss

which is incurred on the portfolio wealth X must exceed α probability of occurring within the time frame. Mathematically it is defined as:

Definition 12. Value at Risk (VaR) *over the time horizon* $[0, T]$ *with confidence level* α *is,*

$$\mathrm{VaR}_\alpha(X_T) = \inf\{x \in \mathbb{R} : \mathbb{P}(X_T + x < 0) \leq 1 - \alpha\}. \tag{4.47}$$

Furthermore, let $\mathscr{G}$ be a set of random financial networks satisfying a set of parameters θ. Let $\{\mathscr{D}_n\}^\theta_{n\in\mathscr{N}}$ and $\{\mathscr{S}_n\}^\theta_{n\in\mathscr{N}}$ be a sequence of sets of defaulted and stressed banks generated by the cascade mechanism described previously on $\mathscr{G}$. Let A^{IB}_θ be the total amount of interbank assets lost due to default by the end of a cascade process. Then we define the following systemic risk measures for our particular cascade mechanism for a random financial network:

Definition 13. *The* α *percentile* Systemic Value at Risk (SVaR) *at the end of the cascade is:*

$$\mathrm{SVar}_\alpha(\mathscr{G}) = \frac{\mathrm{VaR}_\alpha(A^{IB}_\theta)}{\sum_v A^{IB}_v} \tag{4.48}$$

In other words, Systemic Value at Risk can be viewed as the proportion of interbank asset in the financial system that was lost due to defaults as a result of the cascade mechanism.

The *Systemic Expected Shortfall(SES)* modeled after Expected Shortfall of a portfolio is a risk measure, that measures the expected loss of a portfolio conditioned on a loss that exceeds VaR_α has occurred.

Definition 14. Systemic Expected Shortfall *of a Random Financial Network with* α *confidence level at the end of the cascade is,*

$$\mathrm{SES}_\alpha(\mathscr{G}) = \frac{1}{\alpha}\int_0^\alpha \mathrm{SVaR}_s(\mathscr{G})ds \tag{4.49}$$

where SVaR *is the Systemic Value at Risk.*

Acharya et al. (2010) defined Marginal Expected Shortfall of a financial firm as its short-run expected loss conditional on the market taking a loss greater than its Value at

Risk at α level. In relation to cascade models on FRN, we will define Systemic Marginal Expected Shortfall (SMES) as the probability that a specific bank is defaulted conditioned on the Systemic Value at Risk of the financial network has been breached.

Definition 15. Systemic Marginal Expected Shortfall (SMES) *for a bank v at the end of the cascade with α percentile is,*

$$\mathrm{SMES}_\alpha(v) = \mathbb{E}[\mathbb{1}_{v\in\mathscr{D}_\infty^\theta} | A_\theta^{IB} \geq \mathrm{SVaR}_\alpha(\mathscr{G}) \sum_v A_v^{IB}]. \tag{4.50}$$

Systemic Marginal Expected Shortfall can also be thought of as the probability that a certain bank will contribute towards the systemic expected shortfall at a probability level α percentile.

We will use the systemic risk measures defined above as part of the additional numerical exploration into the impact of the asset firesale mechanism in the next section.

4.6.3 Numerical Experiments

We explore the effects of the asset firesale mechanism on the EU network from the section 4.5.3, using 5000 Monte Carlo simulations for each experiment. In addition, we utilize the systemic risk measures defined previously to explore the tail effects of the cascade mechanisms on the EU network which was previously deemed safe from contagion.

Figure 4.8 demonstrates the additional impact of the asset firesale on even the simplest systemic risk measures, the expected default rate and expected stress rate. As seen in the figures, the effect is definately visible, in general 5 additional institutions default on average even when the firesale effect is added. Furthermore, the effect of the stress reaction is similarly invisible even with the firesale effect. This figure also shows that the firesale effect increases the number of defaulted banks and its effect is monotonically increasing with the stress response parameter. This is in contrast to the previous results where the effect of the default cascade decreases monotonically with respect to the stress response parameter λ.

Figure 4.9 paints a starker picture of this banking network when the firesale mechanism is applied. The measure here is essentially a Value at Risk for the number of defaulted

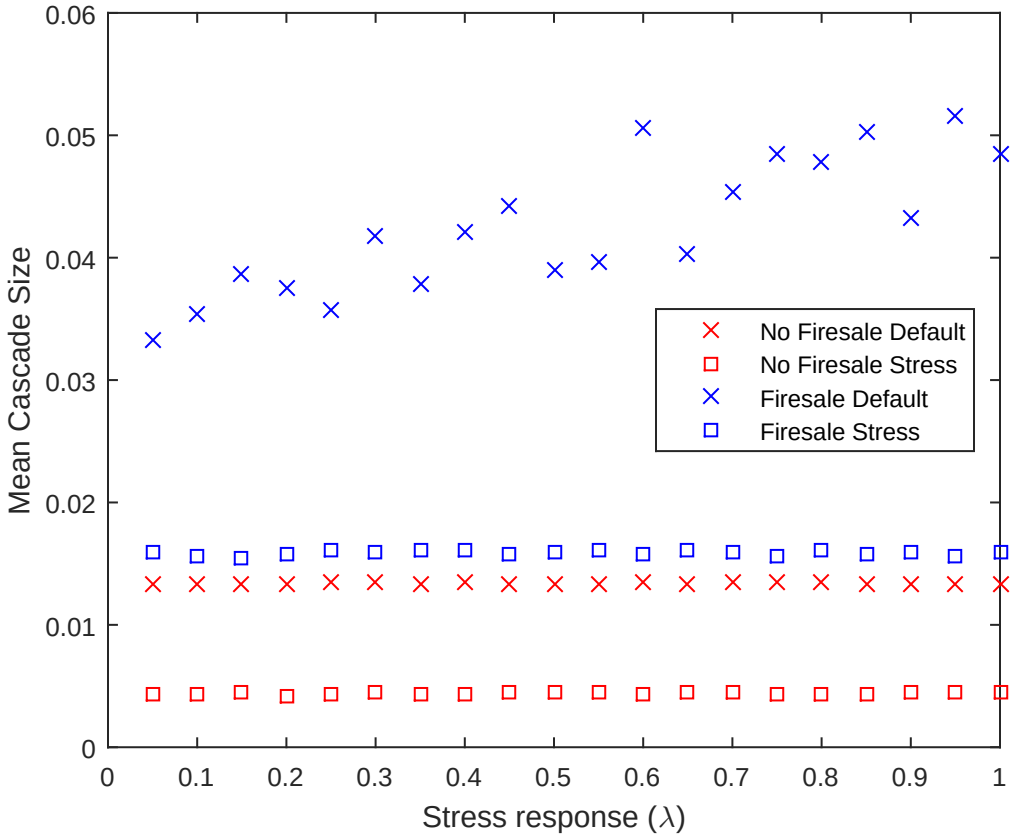
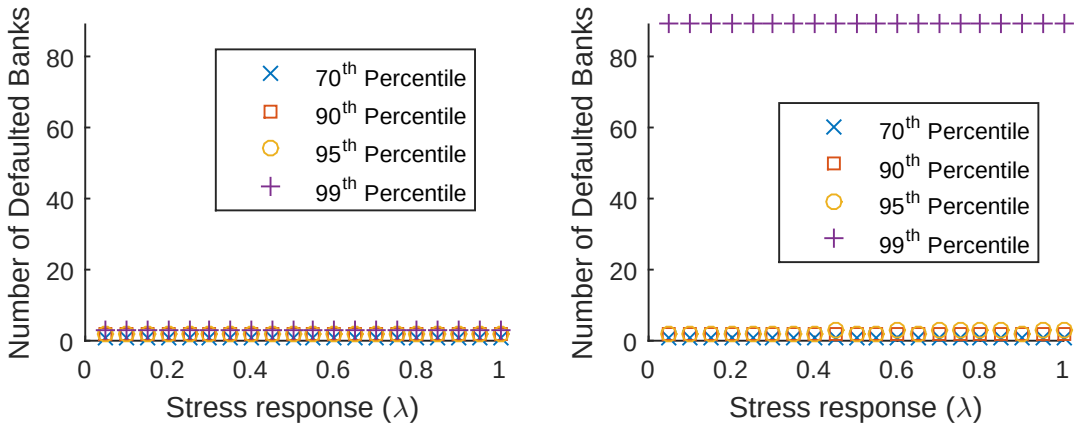


Figure 4.8: Effect of firesales on default and stress contagion in the European interbank network. A comparison of 5000 Monte Carlo simulations based on the original double cascade parametrization of the EU network with and without the firesale mechanism.



(a) without firesale mechanism

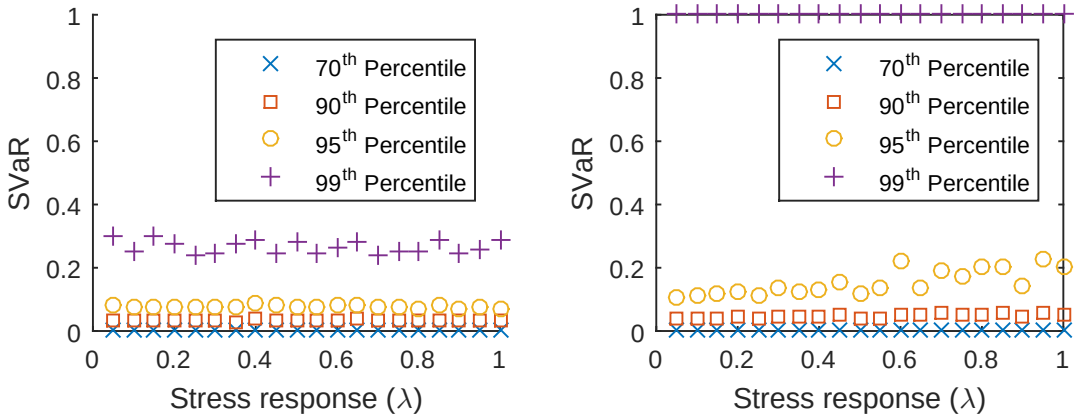
(b) with firesale mechanism

Figure 4.9: The number of defaulted banks at various percentiles is computed with and without firesale mechanism as a function of the stress response parameter λ.

banks. In this case, at the 70th percentile, we see no difference between the expected default size, but at anything above the 99th percentile, we see cases of total system wide default contagion. When the firesale mechanism is not applied, we can see that the result is actually quite similar to the mean cascade size with only 3 banks defaulting at the 99^{th} percentile. This gives us a first look at the impact of the asset firesale mechanism on a tail measure.

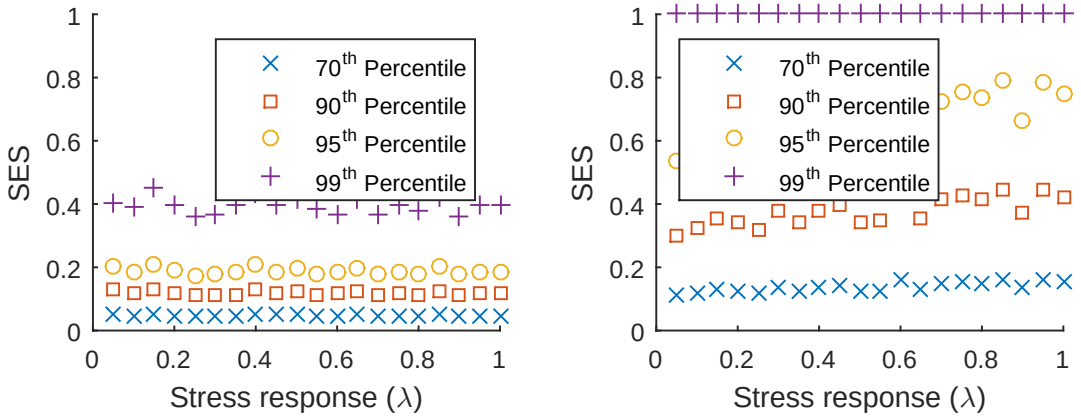
We further show the dangers of tail events to the European banking network through Figure 4.10 and Figure 4.11, which shows the Systemic Value at Risk and Systemic Expected Shortfall at various percentiles. This demonstrates the tremendous damage which will occur to the European interbank network in these rarer events. We can also see pronounced effect of the firesale price impact from the stress response parameter λ. Hoarding additional cash reserves will lead to additional technical defaults due to reduced marked to market A^F prices for the banks in this network. Finally at the 99th percentile, we see that the contagion will spread to the whole system regardless of the stress response. This is additional indication of the importance of tail events in this financial network.

In Figure 4.12, we plot the Marginal Expected Shortfall of two banks with very different number of counterparties. The two figures tell two distinct stories. The highly connected node is able to shock the entire network by first directly shocking enough neigh-



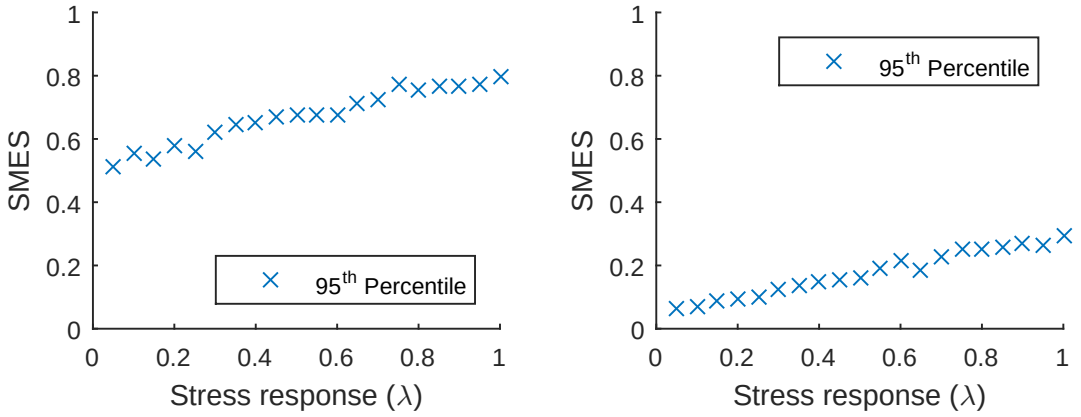
(a) without firesale mechanism (b) with firesale mechanism

Figure 4.10: The Systemic Value at Risk is computed for the RFN calibrated based on the EBA data. The figures show the resulting SVaR at various percentiles and with and without firesale mechanism. The plots are constructed by varying the stress response parameter λ.



(a) without firesale mechanism (b) with firesale mechanism

Figure 4.11: The Systemic Expected Shortfall for the FRN based on the European network. SES is computed with and without firesale mechanism and as a function of the stress response parameter λ.



(a) SMES of highly connected bank (3,5)

(b) SMES of (1,1) bank

Figure 4.12: The Systemic Marginal Expected Shortfall is computed for two different banks while varying the stress response parameter λ. This figure compares the effect of connectivity on banks and their default probability when the financial network is on the brink of collapse.

bours such that a global cascade happens or enough pricing shock occurs that causes a global cascade. It is also more likely to receive a shock from a counterparty simply due to having a higher number of counterparties. On the other hand, the less connected bank is constrained by the number of counterparties it can shock, and receive default shocks from. The SMES shows that the less connected bank's default during crisis is predominantly caused by the increased asset firesale effect driven by increases in the stress parameter λ.

4.7 Conclusions

Our double cascade model is a natural extension of a strand of systemic risk research that studies elementary models that build in either default or stress cascades, but not both. Only by combining the default and stress mechanisms into a single model can one measure such features as the intuitively obvious effect of banks using the stress response to reduce their risk of default.

Developing a feasible and reliable computation framework for a model as complex as ours poses challenges that have been met in the experiments described in Section 5. We have demonstrated how computations can be done by two complementary approaches: the

Monte Carlo (MC) method and the locally tree-like independence (LTI) analytic method. Having these two approaches, each with its pros and cons, allows cross validation, increasing confidence in the results one obtains. MC simulation, with its natural flexibility, remains the workhorse computational framework for general systemic network risk. However, where it applies, the LTI method, given its relation to similar methods in other areas of network science, adds the possibility to understand the flow of the cascade in a different and sometimes better way than one can with MC simulations alone. For example, using LTI one can determine sensitivities to changing parameters through explicit differentiation. In some situations, for example in simulating general assortative (P, Q) configuration models, the MC method is infeasible or leads to unacceptably long run times, whereas the LTI method can be computed without difficulty. Ultimately what is important is that both methods have complementary strengths and weaknesses, and when used in combination lead to robust and reliable conclusions about a wide range of network effects.

Our numerical experiments explore only a small sample of simple model specifications, leaving many promising financial networks to be investigated using our techniques. While the systemic importance of parameters such as network connectivity, mean buffer strength, and the size of the interbank sector have been studied previously, other parameters such as the stress response, the buffer and exposure variances, and graph assortativity, remain almost completely unexplored. The effect of market illiquidity, specifically causing asset firesales have been implemented and studied in the later sections of this chapter. Financial network databases, and the statistical methods for matching such data to the model, are still in an underdeveloped state, but are needed to tie down the wide range of parameters in our model. Planned future investigations of the double cascade model and its extensions will uncover and explain further interesting and unexpected systemic risk phenomena, and find uses by policy makers and regulators.

The double cascade model gave a individual response function for each financial institution to "protect" itself during crisis events via withdrawing from the interbank network. While this was more realistic behaviour at an agent level, it did not take into account of its feedback to non-network-based contagions. Through the addition of the firesale effect, done by directly impacting the marked to market price of each financial institution's fixed assets, we were able to link the individual behaviour of liquidity hoarding to the default contagion in the network.

In the numerics sections, we show two very fundamental results about the European banking network. The first is that the addition of the firesale effect increases default contagion since it can only lower the marked to market price of the fixed assets, but this particular effect is minimal to the expected default cascade size measure. The second result is much more drastic than the first. The expected default cascade measure may show that the European banking network is safe from contagion, but when we apply traditional risk measures to systemic risk, and study similar tail behaviours, we see a much starker picture. At the 99^{th} percentile of our risk measures, total financial system collapse becomes the norm. Already bad situations are compounded further by extreme firesale pricing spirals, and we witness the direct impact of additional liquidity hoarding destroying the entire system through the asset firesale mechanism.

Appendix 4.A Discrete Probability Distributions and the Fast Fourier Transform

Numerical implementation of these models follows the methods outlined in Hurd and Gleeson (2013). In this section, we analyze the case where all random variables $\{\Delta_v, \Sigma_v, \Omega_\ell\}$ take values in a specific finite discrete set $\mathscr{M} = \{0, 1, \dots, (M-1)\}$ with a large value M. It is well known that in this situation the convolutions in (4.19) are slow to compute by direct integration, but can be performed exactly and efficiently by use of the discrete Fast Fourier Transform (FFT).

Let X, Y be two independent random variables with probability mass functions (PMF) p_X, p_Y taking values on the non-negative integers $\{0, 1, 2, \dots\}$. Then the random variable $X + Y$ also takes values on this set and has the probability mass function (PMF) $p_{X+Y} = p_X * p_Y$ where the convolution of two functions f, g is defined to be

$$(f * g)(n) = \sum_{m=0}^{n} f(m)g(n-m) \ . \tag{4.51}$$

Note that p_{X+Y} will not necessarily have support on the finite set $\mathscr{M}$ if p_X, p_Y have support on $\mathscr{M}$. This discrepancy leads to the difficulty called "aliasing".

We now consider the space $\mathbb{C}^M$ of $\mathbb{C}$-valued functions on $\mathscr{M} = \{0, 1, \dots, M-1\}$. The discrete Fourier transform, or fast Fourier transform (FFT), is the linear mapping $\mathscr{F} : a = [a_0, \dots, a_{M-1}] \in \mathbb{C}^M \to \hat{a} = \mathscr{F}(a) \in \mathbb{C}^M$ defined by

$$\hat{a}_k = \sum_{l \in \mathscr{M}} \zeta_{kl} a_l \ , k \in \mathscr{M} \ . \tag{4.52}$$

where the coefficient matrix $Z = (\zeta_{kl})$ has entries $\zeta_{kl} = e^{-2\pi i kl/M}$. The "inverse FFT" (IFFT), is given by the map $a \to \tilde{a} = \mathscr{G}(a)$ where

$$\tilde{a}_k = \frac{1}{M} \sum_{l \in \mathscr{M}} \bar{\zeta}_{kl} a_l \ , k \in \mathscr{M} \ . \tag{4.53}$$

If we let $\bar{a}$ denote the complex conjugate of a, we can define the Hermitian inner product

between

$$\langle a, b\rangle := \sum_{m\in\mathscr{M}} \bar{a}_m b_m \ . \tag{4.54}$$

We also define the convolution product of two vectors:

$$(a * b)(n) = \sum_{m\in\mathscr{M}} a(m)\ b(n-m \text{ modulo } M), \quad n \in \mathscr{M}\ . \tag{4.55}$$

Note that this agrees with (4.51) if and only if the sum of the supports of a and b is in $\mathscr{M}$. Otherwise the difference is called an aliasing error: our numerical implementations reduce or eliminate aliasing errors by taking M sufficiently large.

The following identities hold for all $a, b \in \mathbb{C}^M$:

1. Inverse mappings: $a = \mathscr{G}(\mathscr{F}(a)) = \mathscr{F}(\mathscr{G}(a))$;
2. Conjugation: $\overline{\mathscr{G}(a)} = \frac{1}{M}\mathscr{F}(\bar{a})$;
3. Parseval Identity: $\langle a, b\rangle = M\langle \tilde{a}, \tilde{b}\rangle = \frac{1}{M}\langle \hat{a}, \hat{b}\rangle$;
4. Convolution Identities: $\tilde{a} \cdot\!* \tilde{b} = \widetilde{(a * b)},\ \hat{a} \cdot\!* \hat{b} = \widehat{(a * b)}$,

where $\cdot\!*$ denotes the component-wise product.

As an example to illustrate how the above formulas help, we observe that a typical formula (4.20) can be computed instead by the formula

$$p_{jk}^{(n)} = \frac{1}{M}\ \langle \mathscr{F}(D), \left(\hat{g}_j^{(n-1)}\right)^j \rangle\ , \tag{4.56}$$

where $\hat{D} = \mathscr{F}(D),\ \hat{g}_j^{(n-1)} = \mathscr{F}(g_j^{(n-1)})$ and the power is the component-wise vector multiplication. Such FFT-based formulas can be computed systematically, very efficiently, if the discrete probability distributions for Δ, Σ, Ω are initialized in terms of their Fourier transforms.

Appendix 4.B EU Network Construction

In 2011, the European Banking Authority (EBA) made public a dataset[5] on the interbank exposures of a selection of 90 medium to large European banks, as well as other information such as their Core Tier 1 Capital and Risk Weighted Assets. In this dataset, interbank exposures are aggregated by country on the liability side, which means we only know the aggregated amount each bank v has lent to all banks in each EU country c. In this Appendix we explain how we built the synthetic network in Experiment 3 that mimics stylized facts of real financial networks and uses this EBA data as a source of additional information about interbank liabilities and bank balance sheets.

Motivated by the ubiquitous relevance of networks with fat-tailed degree distributions, as reported in papers such as Cont et al. (2010); Bech and Atalay (2010), we built the skeleton based on the preferential attachment model of Bollobás et al. (2003). Given four parameters α, γ, δ_- and δ_+ and letting $\beta = 1 - \alpha - \gamma$, this model grows a random directed network from a finite initial "seed graph" using three rules:

1. With probability α, add a new vertex v together with an edge from v to an existing vertex w, where w is chosen according to $j_w + \delta_-$ (that is, with probability proportional to $j_w + \delta_-$).

2. With probability β, add an edge from an existing vertex v, chosen according to $k_v + \delta_+$, to an existing vertex w, chosen according to $j_w + \delta_-$.

3. With probability γ, add a new vertex v together with an edge from an existing vertex w to v, where w is chosen according to $k_w + \delta_+$.

This method is known to lead to fat-tailed degree distributions $P_j^- \sim j^{-\tau_-}$ and $P_k^+ \sim k^{-\tau_+}$ with Pareto exponents $\tau_- = 1 + \frac{1+\delta_-(\alpha+\gamma)}{\alpha+\beta}$, $\tau_+ = 1 + \frac{1+\delta_+(\alpha+\gamma)}{\gamma+\beta}$. The four parameters $\alpha = 0.169$, $\gamma = 0.169$, $\delta_- = \delta_+ = 4.417$ are determined by the following conditions: we assumed Pareto exponents $\tau_- = \tau_+ = 4$ (ensuring finiteness of certain moments), that $\alpha = \gamma$, and that the mean degree is $z = 10$ (which is a typical value observed in financial networks). To achieve this, we searched over a range of values of α, generating for each α 1000 samples of networks each consisting of $N = 1000$ nodes. For each realized network

[5] http://www.eba.europa.eu/risk-analysis-and-data/eu-wide-stress-testing/2011

sample, we selected the subnetwork of the 90 most connected nodes and calculated its mean degree z. Finally we selected the α value which provided a mean degree z closest to 10. This set of parameters was then used to produce the skeleton of Experiment 3 by generating a single sample of a (directed) scale-free graph with $N = 1000$ nodes, and retaining its most connected subnetwork of 90 nodes.

In order to build bank balance sheets, we assumed an LTI compatible specification of the buffer and exposure random variables as log-normally distributed conditionally on the network topology:

$$\Delta_v = (k_v j_v)^{\beta_1} \exp[a_1 + b_1 X_v] \ , \tag{4.57}$$

$$\Sigma_v = \frac{2}{3}(k_v j_v)^{\beta_1} \exp[a_1 + b_1 \tilde{X}_v] \ , \tag{4.58}$$

$$\Omega_v = (k_v j_v)^{\beta_2} \exp[a_2 + b_2 X_\ell]. \tag{4.59}$$

where the collection $\{X_v, \tilde{X}_v, X_\ell\}$ consists of independent standard normal random variables. To fix the parameter values, first we arbitrarily set $\beta_1 = 0.3$ and $\beta_2 = -0.2$, with the rationale that the default buffer should increase with bank connectivity, while a larger number of counterparties should imply lower average bilateral exposures. We used the reported Core Tier 1 Capital as a proxy for the default buffers, and thus we matched the first and second sample moments $E[\Delta_v]$ and $E[\Delta_v^2]$ using equation (4.57). Since we found no proxy in the data for the stress buffers Σ, we arbitrarily selected the same parameters as for Δ, but with a prefactor 2/3. Finally, matching equation (4.59) with sample moments $E[A_v^{IB}]$ and $E[(A_v^{IB})^2]$, from the aggregated interbank exposure data, gives us enough equations to determine the full list of parameters: $\beta_1 = 0.3$, $a_1 = 8.03$, $b_1 = 0.9$, $\beta_2 = -0.2$, $a_2 = 8.75$, $b_2 = 1.16$.

Chapter 5

Conclusion

Research on systemic risk has been growing throughout the past decade. Nonetheless, systemic risk is still stigmatized as a risk that can only be measured by regulators. Furthermore, society as a whole has become dependent on regulations to minimize the effects of systemic risk. This thesis aims to build on the foundations of systemic risk set in the literature with richer modelling of individual agent contributions and more accurate measures for these contributions towards the systemic risk of the financial system. This richer modelling allows individual agents within the financial system to better determine how their actions lead to additional systemic risk. In addition, by accurately measuring systemic risk, these agents can internalize its impact to their own welfare and have an incentive to adjust their actions to lower systemic risk. We also discuss many contributing factors to systemic risk that can only be changed by additional regulations. We hope that through this thesis, individuals within the financial system and regulators of the financial system can better understand their contribution towards systemic risk and be able to more effectively control it.

Four years ago, Professor Hurd opened the door for me to start a journey to explore and perhaps extend the research in systemic risk. We started our research focusing on network-based systemic risk models. I distinctly remember when I read Duncan Watts seminal paper on information cascades for the first time, I wondered if his model could be used for systemic risk, and could systemic risk models be useful for all the other aspects of society he mentioned? I still can't answer that question, but during these four years,

I've learned to incorporate more than just the Watts cascade model as part of studying the systemic risk of financial systems. I have been able to incorporate methods from statistical mechanics, biology, and in this thesis, market impact models as part of an effort to further understand systemic risk.

This thesis started out from the introduction chapter as two topics of research, market impact models and network-based systemic risk, and perhaps forcefully joined together by my continuous insistence that they are both required as part of systemic risk. In this section, now that the bodies of work have been explored individually, I want to conclude with an overview of the big picture, the synthesis of this thesis. In this way, I hope to convince the reader that indeed these two topics can be brought together, not just in an ad-hoc manner, and to leave a trail of possibilities that arise from the union of these two topics.

Let us first review the chapters of this thesis as a whole and the links which connect them. There exist direct relationships between consecutive chapters. Chapter 2 provided the understanding of how a single agent would optimize with market impact: without this, the multi-agent setup of Chapter 3 would not be possible. Chapter 3 showed that asset price impact can be approximated in aggregate reasonably well, which led to the addition of asset firesale contagion to Chapter 4.

Another way to look at these three chapters side by side, is to understand the differences in the assumptions in the models in these chapters. One aspect which makes the models in the three chapters different is the fundamental set of information of the agents. In chapter two, the entire universe consists of a single agent, hence the single agent is allowed to act without competition or outside influence. As a result, the agent can even create asset price manipulation strategies. In the third chapter, the agent still has perfect information, but it comes at a cost of having other agents competing with him. In this case, if all agents indeed follow the Pareto Optimal Nash Equilibrium strategy, then the outcome is expected by all the agents. Yet, even in this scenario, we can already see that the impact of dividing agents up into smaller individual agents reduces the resulting market price of the illiquid asset. Finally in chapter 4, the agents' information sets are limited to only the size of the interbank lending between immediate neighbours. In particular, a pure asset firesale contagion would produce similar results to that of a liquidity constraint game from chapter 3. This leads to the suggestion that the much more plausible assumption of

the network-based information set is equivalent to perfect information assumption for some systemic risk contagions.

A usual assumption of network-based systemic risk, and even many similar agent-based models with market impact, is that while a crisis happens, agents are constrained to a sell only strategy. Though this assumption leads to an easier system to analyse, it removes the possibility for well positioned financial institutions to take advantage of this situation. In Chapter 2 and Chapter 3, we showed that it is often in the interest of the agents to purchase risky assets if their constraints allow them, regardless of the future price of the risk asset. It is a worthwhile step in the development of systemic risk to allow for normal behaviour of agents to exist throughout a crisis and is an interesting challenge for new Ph.D. student.

Chapter two provided the foundation for work on a single agent performing portfolio optimization in a financial market with price impact. This work already allows for multiple risky assets to exist in the market, and even allows for hedging strategies to exist. This model fundamentally assumes that the preferences of the management of the financial institution are allowed to be carried out without question. In practice, financial institutions' risk seeking preferences are more bound by regulatory limits than their own appetites. Naturally, this forms additional constraints on this portfolio optimization scheme. Generalizing the mean-variance framework introduced in chapter two, these constraints can be viewed as probabilistic penalties and allow the optimal portfolio and trading strategy to be computed similarly. An interesting and open question arises from such a setup: is it the regulators that will be computing this optimization scheme to determine the riskiness of the financial institution, or is it the financial institution itself? Mathematically, the technology of chapter two provides a firm basis for the solution to both questions.

Another line of thought away from increasing the complexity of the model is to increase the number of agents. We came up with the models in chapter three utilizing this line of thought, but we made fundamentally big changes to the market impact models along the way. In chapter three, we focused and favoured the concept that agents absolutely knew when they must liquidate themselves. As such, this required illiquid assets with zero volatility. Naturally, a different model would arise if we assumed the stochastic asset price process as shown in chapter two, and allowed many agents to compete for these assets. This model would be a true multiplayer extension of the model in chapter two.

Conceptually, the Nash Equilibrium solution can be found in a similar way, with much higher dimensionality. A major technical challenge is the issue of unbounded solutions. This problem arises due to moving away from optimal liquidation strategies and allowing for a degree of freedom in the terminal portfolio. Since the terminal portfolio is variable, there now exists a possibility for unbounded strategies. This effect may be amplified with additional agents in the market, though it is hoped that the opposite occurs since another agent will be able to make a profit against a possible Ponzi scheme easily. One key benefit of such a model compared to the model in chapter three is that these Nash Equilibrium strategies are continuous time trading strategies, and the issues such as front running will occur naturally.

As the previous idea brings the work of chapter two closer to chapter three, the implementation of multiple assets to the model of chapter three would allow the two models to be compared side by side, apples to apples. The implementation of multiple assets for the model in chapter three is straightforward. The asset price processes would not differ from the single asset case and the agents' objective functions and constraints would not change either. The majority of the changes would come from the nature of the response functions, since the boundary formed by selling or buying the asset until the constraint is now a surface rather than a point.

Chapter three also showed that asset firesale mechanisms by themselves do impact financial networks significantly even when direct connections between agents do not exist. This leads to many unanswered questions as to what would happen if asset firesale mechanisms were involved in a financial network with direct connections between agents. In chapter four, we explored an ad-hoc method of stitching together the cascade mechanisms based on connectivity between agents and an asset firesale mechanism but this is simply one possible method of doing so. In the opposite direction, a few lines of thoughts also exist for the cases where we start with the multi-agent market impact model from chapter three and wished to implement a network model on top of it. One idea would be to use the market impact model to generate the shock, which triggers the network cascade model. The idea behind this type of setup is based on the fact that network-based systemic risk models require an exogenous shock to trigger the cascade. This exogenous shock is usually vague in description, but with a multi-agent market impact model, the exogenous shock can be the result of a price shock from the external market which sets up the network for

the cascade mechanisms.

Another possible method to bring these two types of models together is to look at what weaknesses the two models possess and if the other model can effectively remove these weaknesses. One of the implausible assumptions made by the multi-agent market impact model from chapter three is the assumption of perfect information amongst the agents in the system. This assumption is very common in game theory literature, but nevertheless still unjustifiable in the real world. One common implementation to restrict perfect information, is to limit perfect information locally such as on a grid or lattice, or upon direct interaction when agents come into contact. This requires a structure which defines the distance between the agents of this game. Coincidentally, on a network structure, distance between agents is well defined, and the scope of information sharing is usually limited to the edges which are directly connected to the agents. In this sense, the technology from network-based models for systemic risk is well suited for game theory with imperfect information. Taking baby steps, I believe that an agent-based model similar to chapter three, based on a network model with perfect information limited to immediate neighbours will provide a more accurate picture of asset firesales as part of systemic risk, and is feasible with techniques similar to those of existing network models.

Network-based systemic risk models like that of chapter four build upon each other and allow us to study more complex cascade mechanisms, and multiple cascade mechanisms at the same time. For example, cascade mechanisms using soft thresholds which allow for partial losses, have only been studied one at a time. I also believe that a soft threshold cascade mechanism for liquidity hording can be analysed analytically in conjunction with an asset firesale mechanism. All of these ideas allow for a pedagogical approach to studying cascade mechanisms one step at a time and provide wonderful research opportunities for the next few years. Finally, I wish to speculate on a research topic which may be much further down the line. The cascade model in chapter four, and similar network models of systemic risk assume the existence of the network structure as given. I believe that game theoretic methods similar to those of chapter three may also explain why, banks will form an interbank network as a result of one period of trading. In particular, I want to understand what drives networks to form with properties, such as fat tailed degree distributions and core-periphery structure. In many ways, this is a vast and open field of research, which is quite new and may hold many possibilities. Since the techniques developed in this thesis

cannot be directly applied to understanding the question of how these networks are formed and their mysteries, I feel this will be a research topic that lies much further down the road.

Index

Bibliography

Acharya, V. V., Pedersen, L. H., Philippon, T., and Richardson, M. P. (2010). Measuring systemic risk.

Alfonsi, A. and Schied, A. (2010). Optimal trade execution and absence of price manipulations in limit order book models. *SIAM Journal on Financial Mathematics*, 1(1):490–522.

Almgren, R. and Chriss, N. (1999). Value under liquidation. *Risk*, 12(12):61–63.

Almgren, R. and Chriss, N. (2001). Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40.

Almgren, R. and Lorenz, J. (2007). Adaptive arrival price. *Trading*, 2007(1):59–66.

Almgren, R., Thum, C., Hauptmann, E., and Li, H. (2005). Direct estimation of equity market impact. *Risk*, 18(7):58–62.

Almgren, R. F. (2003). Optimal execution with nonlinear impact functions and trading-enhanced risk. *Applied Mathematical Finance*, 10(1):1–18.

Amini, H., Cont, R., and Minca, A. (2012). Stress testing the resilience of financial networks. *International Journal of Theoretical and Applied Finance*, 15:1–20.

Amini, H., Cont, R., and Minca, A. (2013). Resilience to contagion in financial networks. *Mathematical finance*.

Anand, K. (2013). The macro-financial risk assessment framework : Model features and policy use. In *Fields Quantitative Finance Seminar*.

Battiston, S., Gatti, D. D., Gallegati, M., Greenwald, B., and Stiglitz, J. E. (2012). Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk. *Journal of Economic Dynamics and Control*, 36(8):1121–1141.

Bech, M. L. and Atalay, E. (2010). The topology of the federal funds market. *Physica A: Statistical Mechanics and its Applications*, 389(22):5223–5246.

Bender, E. A. and Canfield, E. R. (1978). The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A*, 24(3):296–307.

Bertsimas, D. and Lo, A. W. (1998). Optimal control of execution costs. *Journal of Financial Markets*, 1(1):1–50.

Bollobás, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316.

Bollobás, B., Borgs, C., Chayes, J., and Riordan, O. (2003). Directed scale-free graphs. In *Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 132–139. Society for Industrial and Applied Mathematics.

Boss, M., Elsinger, H., Summer, M., and Thurner, S. (2004). Network topology of the interbank market. *Quantitative Finance*, 4(6):677–684.

Brown, D. B., Carlin, B. I., and Lobo, M. S. (2010). Optimal portfolio liquidation with distress risk. *Management Science*, 56(11):1997–2014.

Brunnermeier, M. K. and Pedersen, L. H. (2005). Predatory trading. *The Journal of Finance*, 60(4):1825–1863.

Brunnermeier, M. K. and Pedersen, L. H. (2009). Market liquidity and funding liquidity. *Review of Financial Studies*, 22(6):2201–2238.

Caccioli, F., Shrestha, M., Moore, C., and Farmer, J. D. (2014). Stability analysis of financial contagion due to overlapping portfolios. *Journal of Banking & Finance*, 46:233–245.

Cartea, Á. and Jaimungal, S. (2015). Optimal execution with limit and market orders. *Quantitative Finance*, 15(8):1279–1291.

Cetin, U., Jarrow, R. A., and Protter, P. (2004). Liquidity risk and arbitrage pricing theory. *Finance and stochastics*, 8(3):311–341.

Chan, K. K. and Milne, A. (2014). Bank competition, fire-sales and financial stability. *The European Journal of Finance*, 20(10):874–891.

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

Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703.

Cont, R., Moussa, A., and Santos, E. B. (2010). *Network Structure and Systemic Risk in Banking Systems*, Handbook of Systemic Risk, Editors: Fouque, J.-P. and Langsam, J. pages 327–367. Cambridge University Press.

Davis, M. H. and Norman, A. R. (1990). Portfolio selection with transaction costs. *Mathematics of Operations Research*, 15(4):676–713.

Diamond, D. W. and Rajan, R. G. (2005). Liquidity shortages and banking crises. *The Journal of finance*, 60(2):615–647.

Eisenberg, L. and Noe, T. H. (2001). Systemic risk in financial systems. *Management Science*, 47(2):236–249.

Facchinei, F. and Kanzow, C. (2007). Generalized nash equilibrium problems. *4OR*, 5(3):173–210.

Gai, P., Haldane, A., and Kapadia, S. (2011). Complexity, concentration and contagion. *Journal of Monetary Economics*, 58:453–470.

Gai, P. and Kapadia, S. (2010a). Contagion in financial networks. *Proceedings of the Royal Society A*, 446(2120):2401–2423.

Gai, P. and Kapadia, S. (2010b). Liquidity hoarding, network externalities, and interbank market collapse. In *Proc. R. Soc. A*, volume 466, page 439.

Gatheral, J. and Schied, A. (2011). Optimal trade execution under geometric brownian motion in the almgren and chriss framework. *International Journal of Theoretical and Applied Finance*, 14(03):353–368.

Gatheral, J. and Schied, A. (2013). Dynamical models of market impact and algorithms for order execution. *HANDBOOK ON SYSTEMIC RISK, Jean-Pierre Fouque, Joseph A. Langsam, eds*, pages 579–599.

Gleeson, J. P., Hurd, T., Melnik, S., and Hackett, A. (2012). Systemic risk in banking networks without monte carlo simulation. pages 27–56.

Godley, W. and Lavoie, M. (2012). *Monetary economics: an integrated approach to credit, money, income, production and wealth*. Springer.

Goh, K.-I., Lee, D.-S., Kahng, B., and Kim, D. (2003). Sandpile on scale-free networks. *Phys. Rev. Lett.*, 91:148701.

Haldane, A. (2010). The $100 billion dollar question. *Bank of England*.

Harding, R. and Atkins, R. (2014). Problem of banks seen as 'too big to fail' still unsolved. IMF.

Hurd, T. (2015). The construction and properties of assortative configuration graphs. *arXiv preprint arXiv:1512.03084*.

Hurd, T. R. (2016). *Contagion! Systemic Risk in Financial Networks*. Springer.

Hurd, T. R., Cellai, D., Melnik, S., and Shao, Q. H. (2016). Double cascade model of financial crises. *International Journal of Theoretical and Applied Finance*, 19(05):1650041.

Hurd, T. R. and Gleeson, J. P. (2013). On watts contagion model with random link weights. *Journal of Complex Networks*, 1(1):25–43.

Janson, S. (2009). On percolation in random graphs with given vertex degrees. *Electronic Journal of Probability*.

Janson, S. et al. (2014). The probability that a random multigraph is simple. ii. *Journal of Applied Probability*, 51:123–137.

Lee, S. H. (2013). Systemic liquidity shortages and interbank network structures. *Journal of Financial Stability*, 9(1):1–12.

Lehalle, C.-A. (2013). Market microstructure knowledge needed for controlling an intra-day trading process. *arXiv preprint arXiv:1302.4592*.

Lorenz, J. and Almgren, R. (2011). Mean–variance optimal adaptive execution. *Applied Mathematical Finance*, 18(5):395–422.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

May, R. M. . and Arinaminpathy, N. (2010). Systemic risk: the dynamics of model banking systems. *Journal of The Royal Society Interface*, 7(46):823–838.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: the continuous–time model. *Rev. Econom. Statist.*, 51:247–257.

Mitchell, R. M. (2001). *FREE MONEY–Plan for Prosperity*. Rodger Malcolm Mitchell.

Moazeni, S., Coleman, T. F., and Li, Y. (2013). Optimal execution under jump models for uncertain price impact. *Journal of Computational Finance*, 16(4):1–44.

Motter, A. E. and Lai, Y.-C. (2002). Cascade-based attacks on complex networks. *Phys. Rev. E*, 66:065102.

Nier, E., Yang, J., Yorulmazer, T., and Alentorn, A. (2007). Network models and financial stability. *Journal of Economic Dynamics and Control*, 31(6):2033–2060.

Obizhaeva, A. A. and Wang, J. (2013). Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets*, 16(1):1–32.

Perold, A. F. (1988). The implementation shortfall: Paper versus reality. *The Journal of Portfolio Management*, 14(3):4–9.

Pham, H. and Tankov, P. (2008). A model of optimal consumption under liquidity risk with random trading times. *Mathematical Finance*, 18(4):613–627.

Roukny, T., Bersini, H., Pirotte, H., Caldarelli, G., and Battiston, S. (2013). Default cascades in complex networks: Topology and systemic risk. *Scientific reports*, 3.

Schied, A. and Schöneborn, T. (2009). Risk aversion and the dynamics of optimal liquidation strategies in illiquid markets. *Finance and Stochastics*, 13(2):181–204.

Schied, A., Schöneborn, T., and Tehranchi, M. (2010). Optimal basket liquidation for cara investors is deterministic. *Applied Mathematical Finance*, 17(6):471–489.

Schöneborn, T. (2008). *Trade execution in illiquid markets: Optimal stochastic control and multi-agent equilibria.* PhD thesis, Technischen Universität Berlin.

Schwarcz, S. L. (2008). Systemic risk. *Geo. LJ*, 97:193.

Tian, H. and Weinan, E. (2014). Fire sale in financial networks. In *Information Sciences and Systems (CISS), 2014 48th Annual Conference on*, pages 1–5. IEEE.

Tirole, J. (2011). Illiquidity and all its friends. *Journal of Economic Literature*, 49(2):287–325.

Turner, A. (2010). Too much 'too big to fail?'. *Economist's View*.

Upper, C. (2011). Simulation methods to assess the danger of contagion in interbank markets. *Journal of Financial Stability*, 7(3):111–125.

U.S. House of Representatives (2013). Who is too big to fail: Are large financial institutions immune from federal procescution? In *Hearing before the Subcommittee on Oversignt and Investigations of the Committee on Financial Services, U.S. House Of Representatives, One Hundred Thirteenth Congress, First Session*.

Watts, D. J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9):5766–5771.

Willinger, W., Govindan, R., Jamin, S., Paxson, V., and Shenker, S. (2002). Scaling phenomena in the internet: Critically examining criticality. In *Proceedings of National Acadamy of Sciences*.

Wray, L. R. (2001). The endogenous money approach.

Zazzaro, A. (2002). How heterodox is the heterodoxy of the monetary circuit theory? the nature of money and the microeconomy of the circuit," working papers 163, universita' politecnica delle marche (i), dipartimento di scienze economiche e sociali.

Zhang, H. and Moorsel, A. (2009). Fast generation of scale free networks with directed arcs. In Bradley, J., editor, *Computer Performance Engineering*, volume 5652 of *Lecture Notes in Computer Science*, pages 131–148. Springer Berlin Heidelberg.

Zhang, T. (2014). *Nash Equilibria in Market Impact Models: Differential Game, Transient Price Impact and Transaction Costs, PhD Thesis*. PhD thesis, Mannheim, Universität Mannheim, Diss., 2014.