Essays on Asymmetric Information and Trading Constraints

by

György Venter

Thesis submitted for the degree of

Doctor of Philosophy in Finance at

the London School of Economics and Political Science

September 2011

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others.

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without the prior written consent of the author.

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

Abstract

This thesis contains three essays exploring the asset pricing implications of asymmetric information and trading constraints.

Chapter 1 studies how short-sale constraints affect the informational efficiency of market prices and the link between prices and economic activity. I show that under short-sale constraints security prices contain less information. However, short-sale constraints increase the informativeness of prices to some agents who learn about the quality of an investment opportunity from market prices and have additional private information. This, in turn, can lead to higher allocative efficiency in the real economy. My result thus implies that the decrease in average informativeness due to short-sale constraints can be more than compensated by an increase in informativeness to some agents.

In Chapter 2, I develop an equilibrium model of strategic arbitrage under wealth constraints. Arbitrageurs optimally invest into a fundamentally riskless arbitrage opportunity, but if their capital does not fully cover losses, they are forced to close their positions. Strategic arbitrageurs with price impact take this constraint into account and try to induce the fire sales of others by manipulating prices. I show that if traders have similar proportions of their capital invested in the arbitrage opportunity, they behave cooperatively. However, if the proportions are very different, the arbitrageur who is less invested predates on the other. The presence of other traders thus creates predatory risk, and arbitrageurs might be reluctant to take large positions in the arbitrage opportunity in the first place, leading to an initially slow convergence of prices.

Chapter 3 (joint with Dömötör Pálvölgyi) studies the uniqueness of equilibrium in a textbook noisy rational expectations economy model a la Grossman and Stiglitz (1980). We provide a very simple proof to show that the unique linear equilibrium of their model is the

unique equilibrium when allowing for any continuous price function, linear or not. We also provide an algorithm to create a (non-continuous) equilibrium price that is different from the Grossman-Stiglitz price.

Acknowledgements

First of all, I would like to thank my advisors, Dimitri Vayanos and Kathy Yuan, for all the help I received during my years at LSE.

I am especially indebted to Dimitri, without whose guidance and encouragement this thesis could have never been completed. I appreciate all his detailed comments and insightful feedbacks with which he patiently conducted my research efforts. I only wish I would have followed his advice more promptly.

Kathy, being a connoisseur in most of my areas of interest, provided invaluable advice starting from the time she arrived to LSE, beating into me how to focus my loose ideas and present them more clearly.

After feeling intimidated by his sharp intellect for a long-long time, I was relieved to learn Péter Kondor thought my work was worth a dime, and encouraged me throughout the job market year. He deserves special thanks for his discussions, numerous suggestions, and mental support.

Many more academics have offered feedback and encouragement throughout the years. I am particularly grateful for being in the right place at the right time to meet Antoine Faure-Grimaud, Denis Gromb and Philippe Mueller among them.

I acknowledge financial support from the Department of Finance, ESRC, and the Paul Woolley Centre for the Study of Capital Market Dysfunctionality at LSE.

My time in London was made bearable thanks to the friendship and support of some close friends. I would like to thank Zsófia Bárány and Andrea Vedolin for joining me many times to appreciate some fine Monmouth coffee or a couple of glasses of Pimm's, while listening to my stupid research ideas, random moans, or pointlessly long stories. It was great to see Dávid Kaposi when Saturday came. Balázs Keszegh and Dömötör Pálvölgyi, who is also a coauthor of Chapter 3, provided support from the far-off Little Corner.

Without having to bear the risk of singling out one and offending the others, I would like to thank Aytek Malkhozov for being around for many years. One cannot ask for a better fellow delinquent than him. Our endless discussions have resulted in many exciting projects, including late-night diet coke & whiteboard sessions, flying the kite on the St Clemens rooftop, or playing football in the FMG corridor. Hopefully many conjoint papers will enrich the list of achievements soon.

Finally, I would like to thank my parents, Magdolna and György, and my brother, Lóránt. Without their continuous love and support I doubt I would have made it.

Contents

List of Figures

Chapter 1

Short-sale constraints and credit runs

1.1 Introduction

According to the view of many academics and regulators, short-sale constraints compromise market liquidity and reduce the informativeness of market prices, while preventing value-destroying price manipulation and hence severe economic inefficiencies. A press release of the Securities and Exchange Commission (SEC), issued on the 19th of September, 2008, clearly illustrates this point. They state that "under normal market conditions, short selling contributes to price efficiency and adds liquidity to the markets", but argue in favour of an emergency order that bans short selling, as shorting, observed e.g. after the collapse of Lehman Brothers, can lead to sudden price declines unrelated to true value. Since financial institutions "depend on the confidence of their trading counterparties in the conduct of their core business", if prices can influence how these institutions are perceived by counterparties and clients, low prices can have damaging effects on the value of institutions as well. Thus, providing a floor to asset prices can be beneficial.[1]

In this chapter I examine how short-sale constraints affect both the informational efficiency

[1] See http://www.sec.gov/news/press/2008/2008-211.htm. A similar point is reached by the Financial Services Authority (FSA) discussion paper on short-selling (Financial Services Authority (2009), p.11-12). In particular, they claim that the negative impact of shorting "reduces the ability of a firm to raise equity capital or to borrow money, and makes it harder for banks to attract deposits."

of prices, and the link between prices and economic activity. I show that under short-sale constraints security prices contain less information. This is consistent with previous work, and my contribution is to derive a simple closed-form solution of a rational expectations equilibrium (REE) with short-sale constraints. My main result concerns the feedback to the real economy. I find that although prices contain less information, short-sale constraints increase the informativeness of prices to some agents who have additional private information. This, in turn, yields an equilibrium of the real economy that has higher allocative efficiency. My result thus implies that the decrease in average informativeness due to short-sale constraints can be more than compensated by an increase in informativeness to some agents.

To analyze the informational effects of short-sale constraints, I extend an asset pricing model with information spillover from the financial market to the real economy. I use a noisy rational expectations model of a financial market with asymmetric information, where noise comes from a random demand shock, as in Grossman and Stiglitz (1980), and I introduce short-selling constraints on a subset of informed traders. For the real part of the economy, I consider a distressed financial institution (e.g. investment bank) that requires outside capital from multiple lenders or short-term creditors to support its existing positions.[2] Creditors, endowed with dispersed private information about the value of the bank's assets, consider whether to supply capital to this institution. I model bank financing as game with strategic complementarities: the bank survives if the amount of capital provided by creditors is sufficiently large, and creditors' payoff is higher if the bank avoids bankruptcy.[3] Besides their private signals, creditors also observe the price of a traded security. The connection between the security market and the financing is provided by the correlation between the payoff of the security and the unknown quality of the financial institution. Therefore, the price, which gathers information in the security market, constitutes a public signal to capital providers.[4]

[2] There is ample anecdotal evidence about Bear Stearns and Northern Rock not being able to secure short-term financing and being the victims of runs by their creditors at the beginning of the 2007-2008 crisis; see, for example, Brunnermeier (2009) and Shin (2009). Ivashina and Scharfstein (2010) show that after the failure of Lehman Brothers in September 2008, there were runs by short-term lenders on financial institutions, making it hard for banks to roll over their short term debt. Moreover, runs on other financial institutions, such as investor withdrawals from hedge funds or mutual funds can be viewed as a coordination game among capital providers, see Shleifer and Vishny (1997) and the vast literature on limited arbitrage.

[3] Indirect and direct evidence of coordination motives among creditors have been shown by Asquith et al. (1994), Brunner and Krahnen (2008) and Hertzberg et al. (2010). Moreover, Chen et al. (2010) document coordination motives among investors of mutual funds.

[4] The financial asset can be interpreted, for example, as a zero-net-supply derivative on the share price of

The main observation of the model is that even though short-sale constraints decrease the information content of prices, certain creditors endowed with additional private information can learn more from asset prices with short-sale constraints than without the constraints. The idea is that when creditors combine their private signals with the market price to form beliefs about the state of the world, they also have to assess to what extent a high (low) market price reflects a high (low) fundamental value or a high (low) demand shock, i.e. whether informed traders buy or whether they (would) short-sell. In presence of short-sale constraints, a high demand shock increases the price, and informed investors would like to short, but cannot. It leads to a decrease in the aggregate order flow, which is dominated by noise trading. High price realizations are hence more noisy and less informative about the true state of the world, as negative information about fundamentals is less incorporated into prices. Put differently, a given price realization means lower payoff if one believes the constraint binds in the market.

To see the intuition for how prices can provide more information in presence of trading constraints, consider a creditor who receives a private signal realization higher than the price she sees. A high signal means that according to her private information, states when the payoff is lower than the price have low probabilities as they are tail events. She knows that if informed traders (those who can) are shorting, the same price realization corresponds to lower fundamentals compared to the case without short-sale constraints. But lower fundamentals have smaller probabilities according to her private belief. Combining these two observations, she assigns a smaller probability to informed traders shorting the asset than without short-sale constraints, and thus thinks the asset payoff is higher. This reinforces her private signal, and implies that her posterior can be more precise than without short-sale constraints.

Then I study the effect of short-sale constraints in the security market on the bank financing. I show that the presence of short-sale constraints introduces multiple equilibria in the coordination game, even when private information is arbitrarily precise. This result stems from the observation that creditors with high private signal learn more from the market price in presence of short-sale constraints. Indeed, when their posterior variance is smaller, creditors have more precise assessments about both the bank's fundamental and about the information of other creditors. For every level of private noise precision, when short-sale

the bank, as an industry index that includes the bank, or the price of a security that the bank has on the asset side of the balance sheet.

constraints are sufficiently tight, they reinstate common knowledge among the subset of more informed creditors, and lead to self-fulfilling beliefs and two stable equilibria. I refer to the first one, when creditors rely only on their private signals, as the informationally efficient equilibrium, because in the limit when private signals become very precise, agents ignore the market price. This equilibrium is equivalent to the unique equilibrium of the game without short-sale constraints. However, in presence of the constraint there exists a second stable equilibrium, where creditors with high private signals keep relying on the public signal, if they know that other creditors with similarly high signals do as well.

Interestingly, in this second stable equilibrium the bank receives more capital than in the informationally efficient equilibrium, thus they can be called high and low investment equilibrium, respectively. This is because short-sale constraints only improve the precision of agents with signals higher than the market price, so they are the creditors who might react to the 'news' in short-sale constraints. Intuitively, short-sale constraints can only affect the equilibrium outcome if there are some creditors who behave differently in the informationally efficient equilibrium, but due to short-sale constraints learn more about each other's action. Thus it is straightforward that the second equilibrium, whenever it exists, must feature more capital provision than the informationally efficient equilibrium. Short-sale constraints improve the information of some agents who would stay out in absence of the constraint, and create self-fulfilling beliefs and multiplicity in equilibrium actions among these creditors. This leads to more investment, banks with lower asset quality remain solvent, and the equilibrium is closer to the first best. I conclude that short-sale constraints improve allocational efficiency by mitigating the adverse effect of the coordination externality. Therefore, if the gain of short-sale constraints in terms of the increased allocational efficiency of the real economy is higher (or more important) than the loss in terms of informational inefficiency in the financial market, short-sale constraints can be beneficial.

The model presented in this chapter is not the first to highlight the impact of trading constraints on the allocational efficiency of the real economy. Panageas (2003) and Gilchrist et al. (2005) study firms' investment decisions when they raise capital during asset price bubbles, when the cost of capital is low due to short-sale constraints. Both studies rely on the literature initiated by Miller (1977) and Harrison and Kreps (1978), who suggest a link between the level of belief heterogeneity and inflated asset prices (see also Scheinkman and

Xiong (2003), and Rubinstein (2004) for many more 'anomalies' associated with short-sale constraints). In contrast to these papers, in my model agents are rational, and short-sale constraints do not inflate the price, following the insights of Diamond and Verrecchia (1987). In particular, in this study security prices in presence of short-sale constraints are lower than without the constraints, and hence according to Panageas (2003) and Gilchrist et al. (2005) investment should be lower. My focus is on the information provided by market prices instead of price levels.

The financial market model of this chapter is similar to that in Yuan (2005, 2006), who studies a REE with asymmetric information and constraints on borrowing and shorting. She numerically shows that, in presence of borrowing restrictions, a higher market price can reduce uncertainty about the constraint status of informed investors, and that this information effect can be strong enough to cause a backward bending demand curve. In contrast, the first part of this chapter provides a simple closed-form solution of a model that is simplified in one dimension but allows for more generality in other dimensions.[5] Finally, Bai et al. (2006) and Marin and Olivier (2008) study the effects of short-sale constraints when investors trade for informational and allocational purposes. In both papers, trading constraints limit the positions of all informed traders. When the constraints bind, asset prices stop reflecting fundamentals, uninformed investors demand a large discount, and prices exhibit large drops. Therefore, in these models high prices are more informative than low prices. In contrast, in models presented here and in Yuan (2005, 2006), only a subset of informed investors are subject to the short-selling constraint, and uninformed investors need to form beliefs about the size of the demand shock, i.e. the constraint status of informed investors. The most important distinction is that short-sale constraints bind for high prices, making them less informative than low prices.

The model also belongs to the literature on coordination games with strategic complementarities, developed by Carlsson and van Damme (1993) and Morris and Shin (1998), and contributes to discussion about the fragile interaction between private and public information (see, for example, Morris and Shin (1999, 2001, 2002, 2003, 2004) and Hellwig (2002)). In

[5] Also, in a financial market with wealth- and shortsale-constrained risk-neutral agents and an asset supply exponentially distributed, Barlevy and Veronesi (2003) present a theory of stock market crashes, where high asset prices are more informative than low prices.

particular, Morris and Shin (2001) show that when private information becomes arbitrarily precise, a coordination game has a unique equilibrium. In the discussion of Morris and Shin (2001), Atkeson (2001) highlights the potential role of financial markets as the source of endogenous public information, formalized by Angeletos and Werning (2006). They show that a unique equilibrium might not prevail, if the precision of the public signal that aggregates private information increases faster than the precision of the private signal. Hellwig et al. (2006) and Tarashev (2007) also study a coordination game with a financial price as the public signal, while Ozdenoren and Yuan (2008) and Goldstein et al. (2009) study coordination among traders in the market. A common element in these papers is that the informational content of the public signal does not vary across equilibria. In contrast to many previous models, in the model presented in this chapter the informativeness of the public signal varies across its realizations. This is similar in spirit to Angeletos et al. (2006). However, in their analysis the signal is the equilibrium action of a policy maker, whereas in my study the varying informativeness is the result of the asymmetric nature of short-sale constraints. Finally, there are several papers that highlight the adverse effect of short-selling on allocative efficiency in the economy and hence argue in favour of short-sale constraints, see for example Goldstein and Guembel (2008) or Liu (2010). However, to my knowledge, this is the first model that explicitly studies the informational effect of short-sale constraints on real economic activity.

The remainder of the chapter is organized as follows. Section 1.2 presents the financial market. Section 1.3 studies the equilibrium of the financial market and examines the effect of short-sale constraints on the equilibrium price. Section 1.4 analyzes the information content of stock prices with and without short-sale constraints for outside observers. Section 1.5 embeds the credit run model into the economy, and Section 1.6 presents the equilibrium of the coordination game with the skewed public signal. Section 1.7 discusses the results, contrasts the findings with the related literature, and provides some comparative statics and policy implications. Finally, Section 1.8 concludes.

1.2 Model

This section introduces the financial market model. I consider a two-period economy with dates $t = 0$ and 1. At date 0 investors trade, and at date 1 assets pay off. The market is populated by three types of agents: informed and uninformed rational investors, and noise traders.

1.2.1 Assets

There are two securities traded in a competitive market, a risk-free bond and a risky stock. The bond is in perfectly elastic supply and is used as numeraire, with the risk-free rate normalized to 0. The risky asset is assumed to be in net supply of $S \geq 0$, and has final dividend payoff d at date 1, that is the sum of two random components: $d = f + n$. The first risky component of the dividend payoff, f, can be regarded as the fundamental value of the asset. The second component, n, is thought of as additional noise, preventing agents from knowing the exact dividend payoff. The price of the stock at date 0 is denoted by p.

1.2.2 Traders

I assume that the asset market is populated by a continuum of rational traders in unit mass. Traders do not hold endowments in the risky security. They are risk averse and, for tractability, I assume that they have a mean-variance objective function over terminal wealth.[6] Agent k, for $k \in [0, 1]$, maximizes

$$E\left[W_k|\mathcal{I}_k\right] - \frac{\rho}{2} Var\left[W_k|\mathcal{I}_k\right],$$

where ρ is the risk aversion parameter, common across agents. The final wealth $W_k = W_k^0 + x_k\left(d - p\right)$ is given by the initial wealth W_k^0 plus the number of shares purchased, x_k, multiplied by the profit per share, $d - p$. $\mathcal{I}_k$ is the information set of trader k, and $E\left[.|\mathcal{I}_k\right]$ and $Var\left[.|\mathcal{I}_k\right]$ denote the expectation and variance conditional on the information set $\mathcal{I}_k$, respectively.

Rational investors can be either informed or uninformed. Informed traders, who have a measure of λ and are indexed with $k \in [0, \lambda)$, observe the realization of the fundamental f

[6] The mean-variance objective function is equivalent to maximizing exponential (i.e. CARA) utility as long as the uncertainty faced by traders is Gaussian. With short-sale constraints this is not the case for uninformed investors, but the model is nevertheless solvable and yields qualitatively the same result as the one dicussed here.

but not n. The other set of rational traders, with measure $1-\lambda$, and indexed with $k \in [\lambda, 1]$, are uninformed, and do not observe any (private) signals about f. Instead, all agents of the model observe the market price p. These assumptions imply that the risk associated with n is unlearnable for everyone, thus uninformed traders try to best guess component f. Formally, the information set of informed traders is $\mathcal{I}^i = \{f, p\} = \{f\}$, as the price cannot provide more information about the final payoff than their private observation. The information set of uninformed traders is $\mathcal{I}^{ui} = \{p\}$.

Further, I assume that informed traders might be subject to short-sale constraints.[7] In particular, short-sale constraints mean that trader k's stock position is bounded below by zero, $x_k \geq 0$. Short-sale constraints can be thought of as an extreme case of infinite costs when selling short. I assume that $0 \leq w < 1$ proportion of informed traders are subject to short-sale constraints, and index them by $k \in [0, w\lambda)$, while the remaining, with mass $(1-w)\lambda$, for $k \in [w\lambda, \lambda)$, are unconstrained. When $w = 0$, none of the informed traders are restricted from shorting.[8] Throughout the chapter, a higher w can be (broadly) interpreted as higher cost and/or more difficult shorting. This includes regulatory restrictions (such as the short-sale ban of 2008 or the uptick rule), legal restrictions, search costs for lenders, rebate rates, costs of derivative trading, and even the amount of institutional trading in the market.[9] As agents inside the different investor classes are identical, I drop the subscript k from now on.

Finally, there are noise traders in the market, whose trading behavior is not derived from utility maximization. Noise traders simply buy u shares. I will refer to their trade order as demand shock.[10]

Regarding the distribution of random variables, I assume that fundamental f is drawn

[7] For simplicity, I assume that uninformed traders are not subject to short-sale constraints. Such an extension would only affect the equilibrium price level by influencing the demand of uninformed traders, but would not change the information content of the market price.

[8] The qualitative results of the model do not depend on the exact proportions of the three different trader classes. The cardinal question is whether $w = 0$ or $w > 0$. As discussed later, the assumption $w < 1$ impies that there are always unconstrained informed traders, the stock price always reflects the fundamental f up to some noise, and the equilibrium stock price does not exhibit a jump.

[9] See the Securities and Exchange Commission Rule 10a-1, Almazan et al. (2004), Duffie et al. (2002), Jones and Lamont (2002), Ofek and Richardson (2003), and Nagel (2005), respectively, for these proxies on the difficulty of short-selling.

[10] As it is standard in models with informational heterogeneity, the presence of noise trading u makes sure that the price does not reveal f perfectly, and hence the Grossman-Stiglitz paradox does not apply.

from an improper uniform distribution on the real line. The unlearnable noise component is given by $n \sim N\left(0, \sigma_n^2 = \tau_n^{-1}\right)$, and the demand shock is given by $u \sim N\left(0, \sigma_u^2 = \tau_u^{-1}\right)$, where σ_x^2 denotes the variance of random variable x, and τ_x denotes its precision. Throughout the chapter $\phi(.)$ denotes the probability density function (pdf) of a standard normal distribution, $\Phi(.)$ is its corresponding cumulative distribution function (cdf), and $\Phi^{-1}(.)$ is the inverse of the cdf.

1.2.3 Equilibrium concept

I define an equilibrium of the financial market as follows.

Definition 1 *A rational expectations equilibrium (REE) of the asset market is a collection of a price function $P(f,u)$, and individual strategies for constrained informed, unconstrained informed and uninformed traders, $x^c(f,p)$, $x^{uc}(f,p)$, and $x^{ui}(p)$, respectively, such that*

1. *demand is optimal for informed traders:*

$$x^c(f,p) \in \arg\max_{x \in \mathbb{R}^+} E\left[W^c|f\right] - \frac{\rho}{2} Var\left[W^c|f\right], \tag{1.1}$$

and

$$x^{uc}(f,p) \in \arg\max_{x \in \mathbb{R}} E\left[W^{uc}|f\right] - \frac{\rho}{2} Var\left[W^{uc}|f\right]; \tag{1.2}$$

2. *demand is optimal for uninformed traders:*

$$x^{ui}(p) \in \arg\max_{x \in \mathbb{R}} E\left[W^{ui}|P(f,u)=p\right] - \frac{\rho}{2} Var\left[W^{ui}|P(f,u)=p\right]; \tag{1.3}$$

3. *market clearing:*

$$w\lambda x^c(f,p) + (1-w)\lambda x^{uc}(f,p) + (1-\lambda)x^{ui}(p) + u = S, \tag{1.4}$$

Conditions (1.1)-(1.4) define a competitive noisy rational expectations equilibrium for the trading round. In particular, condition (1.1) states that individual asset demands are optimal for informed traders subject to short-sale constraints, conditioned on their private information. Similarly, condition (1.2) states that individual asset demands are optimal for informed traders with no restriction on shorting, given their private information. Also, condition (1.3) states that individual asset demands are optimal for uninformed traders, conditioned on any information they infer from the price. Finally, (1.4) imposes that the asset market clears: aggregate demand equals supply.

1.3 Equilibrium in the financial market

This section solves for the equilibrium of the trading round and studies the informational effects of short-sale constraints on market prices. The model is solved in the general case with $w \geq 0$, then I contrast the results for $w = 0$ and $w > 0$, that is in absence and presence of short-selling constraints, respectively.

Given the optimization problems (1.1), (1.2) and (1.3), optimal demands are the following: an unconstrained informed trader submits demand function

$$x^{uc}\left(f,p\right) = \frac{f-p}{\rho\sigma_n^2}, \tag{1.5}$$

a short-sale constrained informed trader demands

$$x^{c}\left(f,p\right) = \max\left\{\frac{f-p}{\rho\sigma_n^2}, 0\right\} = 1_{f\geq p}\frac{f-p}{\rho\sigma_n^2}, \tag{1.6}$$

and an uninformed trader demands

$$x^{ui}\left(p\right) = \frac{E\left[f|P=p\right]-p}{\rho\left(Var\left[f|P=p\right]+\sigma_n^2\right)}. \tag{1.7}$$

Solving for equilibrium requires three fairly standard steps. First, I postulate a REE price function. Given the price, I derive the optimal demand of uninformed traders. Finally, I show that the market indeed clears at the conjectured price.

I conjecture the equilibrium price of the form

$$P = f + \begin{cases} A\left(u-C\right) & \text{if } u \leq C \\ B\left(u-C\right) & \text{if } u > C \end{cases}, \tag{1.8}$$

with constants A, B and C to be determined in equilibrium, where $A, B > 0$.

Uninformed agents' information set is given by $\mathcal{I} = \{P(f,u) = p\}$. They observe neither f, nor u, only the price realization p; and they know that in equilibrium this is a piecewise linear function of the two unknown variables, described in (1.8). From the price realization p they form a probabilistic estimate about the fundamental f, while also guessing whether short-sale restrictions bind for constrained traders. Given the conjectured price function (1.8) and the Gaussian distribution of u, uninformed investors' posterior is characterized by the conditional probability density function

$$\begin{aligned} g(f|P=p) &= 1_{f<p}\, g(f|P=p) + 1_{f\geq p}\, g(f|P=p) \\ &= \frac{1}{B\sigma_u} 1_{f<p}\, \phi\left(\frac{f-(p+BC)}{B\sigma_u}\right) + \frac{1}{A\sigma_u} 1_{f\geq p}\, \phi\left(\frac{f-(p+AC)}{A\sigma_u}\right), \end{aligned} \tag{1.9}$$

where I use that for any X random variable with density function $g_X(x)$ and a $\varphi(.)$ continuous, differentiable, and injective transformation, the density function of $Y = \varphi(X)$ is given by

$$g_Y(y) = g_X\left(\varphi^{-1}(y)\right) \left|\varphi^{-1\prime}\right|.$$

The above density function in turn allows uninformed traders to compute the conditional expectation and variance of payoff f given p:

$$E[f|P=p] = p + D \text{ and } Var[f|P=p] = E - D^2, \tag{1.10}$$

where

$$\begin{aligned} D &\equiv -A \int_{-\infty}^{C} (u-C)\,\phi(u)\,du - B \int_{C}^{\infty} (u-C)\,\phi(u)\,du, \text{ and} \\ E &\equiv A^2 \int_{-\infty}^{C} (u-C)^2\,\phi(u)\,du + B^2 \int_{C}^{\infty} (u-C)^2\,\phi(u)\,du. \end{aligned}$$

The conjectured REE price function must equate demand and supply for each possible resolution of f and u. Substituting the optimal demands $(1.6)-(1.7)$ into the market clearing condition (1.4) gives

$$w\lambda 1_{f\geq p}\frac{f-p}{\rho\sigma_n^2}+(1-w)\,\lambda\frac{f-p}{\rho\sigma_n^2}+(1-\lambda)\,\frac{p+D-p}{\rho\,(E-D^2+\sigma_n^2)}+u=S,$$

where the resulting coefficients must equal the conjectured A, B and C, which leads to the following result:

Theorem 2 *A piecewise linear REE of the model exists in the form*

$$P=f+\begin{cases} A\,(u-C) & \text{if } u<C \\ B\,(u-C) & \text{if } u\geq C \end{cases}, \tag{1.11}$$

where

$$A=\frac{\rho\sigma_n^2}{\lambda}\ \text{and}\ B=\frac{\rho\sigma_n^2}{(1-w)\,\lambda},$$

and C is the solution of

$$0=L\,(C)\equiv C+\frac{1-\lambda}{\rho}\frac{D\,(C)}{E\,(C)-D^2\,(C)+\sigma_n^2}-S. \tag{1.12}$$

The omitted technicalities are provided in the appendix.

To see why the model has such an elegant solution, regardless of the distributional assumptions on n and u, notice that the demand of uninformed traders is constant, independent of the price p:

$$x^{ui}=\frac{E\,[f|P=p]-p}{\rho\,(Var\,[f|P=p]+\sigma_n^2)}=\frac{D}{\rho\,(E-D^2+\sigma_n^2)}.$$

It is due to the diffuse prior assumption, which implies that uninformed traders have only one source of information, namely the market price. Thus, a change in the price p is fully offset by a change in their expectation $E\,[f|p]$, while the precision of their information, given by $Var\,[f|p]$, remains constant. Hence, the diffuse prior assumption makes the the inference problem of uninformed traders trivial, and simplifies the analysis relative to Yuan (2005, 2006).

To see the intuition behind the piecewise linear structure and the presence of a kink at $u=C$, consider the aggregate demand of informed investors, given by $wx^c\,(f,p)+(1-w)\,x^{uc}\,(f,p)$, for the price p being close to fundamental payoff f. As long as $p\leq f$, the short-selling constraint does not bind, and a unit mass of informed investors are present in the market, submitting a total demand of $(f-p)/A$. However, for price and fundamental realizations such that $p > f$, some informed traders are barred from the market, and informed investors' aggregate demand is $(1-w)(f-p)/A$, less in absolute terms. It implies that when $p > f$, a less aggressive informed demand meets the residual demand, defined as the demand of uninformed traders, plus the demand of noise traders, minus the asset supply, i.e. $u + x^{ui} - S$, which is simply a linear function of the demand shock u. Therefore, the equilibrium price is more sensitive to large demand shocks, implying $B \geq A$, and is a linear function of state variables f and u, conditional on both the constraint binding or not.

Figure 1-1 illustrates the main result of this section. The graph shows the asymmetric change in the equilibrium price due to the presence of short-sale constraints. When shorting is allowed (left panel), the slope of the price p as a function of demand shock u is the same for every realization of the shock. When shorting is prohibited (right panel), the price function is steeper for large demand shocks than for small (negative) demand shocks. It means that when the constraint binds for some speculators, a small increase in the demand shock has a larger upward price impact. Thus, the price reveals information about the payoff f at different rates in the two regions: it provides more information when the constraint does not bind, i.e. the demand shock is low, and less information, when the constraint does bind, i.e. the demand shock is high.

1.3.1 Properties of the equilibrium price

The rest of the section illustrates how certain properties of the equilibrium price change due to short-sale constraints. In order to determine the direct effect of introducing short-sale constraints in a market, one can compare conditional moments of the fundamental f. My main focus is on the information content of the price, illustrated by the conditional variance and skewness. All the results are proven in the Appendix.

Notice first that $A = \rho\sigma_n^2/\lambda$ does not depend on w. Let B_w, C_w, D_w and E_w denote the equilibrium constants B, C, D and E as a function of the w proportion of short-sale-constrained informed traders. Similarly, one can define P_w to be the equilibrium price function as a function of w, for given f and u realizations and with the corresponding equilibrium coefficients B_w and C_w. The absence of short-sale constraints, i.e. $w = 0$, implies $A = B_0 = \rho\sigma_n^2/\lambda$, and solving for the equilibrium price, (1.12) yields

$$C_0 = \frac{\lambda + \rho A \sigma_u^2}{1 + \rho A \sigma_u^2} S.$$

If the asset is in positive net supply, $S > 0$, C_0 is positive, and $D_0 = AC_0 > 0$, which means that uninformed investors' expectation about the asset payoff is above the market price, $E[f|P_0 = p] = p + D_0 > p$, and they demand a discount of $D_0 > 0$ to hold the asset. The equilibrium price thus becomes

$$P_0(f, u) = f + A(u - C_0), \tag{1.13}$$

which implies a conditional second moment of

$$Var[f|P_0 = p] = A^2\sigma_u^2. \tag{1.14}$$

As (1.9) shows, in the presence of short-sale constraints, the conditional distribution changes because of the different impact of the demand shock on the price for high and low prices, that is when the constraint binds or not. The following proposition compares the informativeness of market prices with and without short-sale constraints:

Proposition 3 *Short-sale constraints lead to a decrease in price informativeness, which is defined as the inverse of the conditional variance of the payoff. Formally, for any price realization p,*

$$Var[f|P_w = p] > Var[f|P_0 = p]. \tag{1.15}$$

Condition (1.15) shows that short-sale constraints increase uninformed traders' perceived uncertainty about the asset payoff, because they decrease the information content of the market price for high demand shock realizations.[11] Uninformed investors demand a larger discount for this increase in uncertainty, implying $D_w > D_0$.

[11] The increase in the conditional variance is present in Bai et al. (2006) and Marin and Olivier (2008) as well, but, as discussed shortly, in those models short-sale constraints decrease the information content of the market price for low price realizations.

It is also interesting to see the implications of short-sale constraints on the equilibrium price volatility. From equations (1.11) and (1.13) one can obtain

$$Var\left[P_w|f\right] = E_w - D_w^2 \text{ and } Var\left[P_0|f\right] = A^2\sigma_u^2.$$

Comparing the volatility with and without short-sale constraints gives the following result:

Proposition 4 *Short-sale constraints lead to an increase in price volatility:*

$$Var\left[P_w|f\right] > Var\left[P_0|f\right].$$

This finding is in line with previous empirical results. Indeed, Ho (1996) finds an increase in stock return volatility when short sales were restricted during the Pan Electric crisis in the Singapore market in 1985-1986. Boehmer et al. (2009) document a sharp increase in intraday volatility during the September 2008 emergency order.

The asymmetric effect of short-sale constraints on prices and price informativeness can be easily tested by analyzing return skewness. In the static model presented here, one can define two returns. Following Bai et al. (2006), I define the *announcement-day return* of the stock as the dollar return made between the trading round, date 0, and the final date 1, and the *market return* as the return made between a hypothetical date -1, before trading commences, and date 0. For simplicity, I assume that the price at this date -1, denoted by p^{-1} is constant. Formally, the announcement-day return is given by $r(f,u) = f - P(f,u)$, and the market return is given by $R(f,u) = P(f,u) - p^{-1}$.

Hong and Stein (2003) argue that short-sale constraints can lead to negative skewness in stock returns, which they relate to market crashes. On the empirical side, Reed (2007) documents that under short-sale constraints, the distribution of announcement day stock returns is more left-skewed. He also reports that returns have larger absolute values, when short-selling is constrained. Calculating properties of the announcement-day return with and without short-sale constraints gives the following results:

Proposition 5 *Short-sale constraints lead to more negatively skewed announcement-day returns:*

$$Skew\left[r_w(f,u)\right] < Skew\left[r_0(f,u)\right],$$

and an increase in the absolute value of returns:

$$E\left[\left|r_w\left(f,u\right)\right|\right] > E\left[\left|r_0\left(f,u\right)\right|\right].$$

The intuition for the negative skewness and is that short-sale constraints impede the negative information to be incorporated into the price, which leads to larger realized losses when the final payoff becomes public knowledge. Market prices reflect positive information more, and hence announcement day returns are smaller in this case. Moreover, absolute returns increase simply because losses become larger.

Regarding empirical evidence, Bris et al. (2007) find that in markets where short-selling is either prohibited or not practiced, market returns display significantly less negative skewness. Analyzing market returns with and without short-sale constraints gives the following result:

Proposition 6 *Short-sale constraints lead to less negatively skewed market returns:*

$$Skew\left[R_w\left(f,u\right)\right] > Skew\left[R_0\left(f,u\right)\right].$$

Because of short sale constraints, negative information is less incorporated to the market price and hence downward price movements and negative market returns are smaller in markets where shorting is prohibited.

To conclude this section with a technical sidenote, it is interesting to mention that there are differences in the asset pricing implications of two branches of asymmetric information models with portfolio constraints. The first type of these models includes Bai et al. (2006) and Marin and Olivier (2008). In both of these papers, noise in the market (from the point of view of uninformed traders) comes from the unknown endowment of insiders, and trading constraints limit the positions of all informed traders. These assumptions have two implications: the constraint status of informed investors can be directly inferred from the equilibrium price, and the constraint for insiders is binding for low prices. Therefore, in these models, high prices are more informative than low prices. The model presented here belongs to the other branch, together with Barlevy and Veronesi (2003), and Yuan (2005, 2006). In these studies noise arrives to the market from noise traders' demand or random supply, and only a subset of informed investors are subject to the short-selling constraint. Importantly, uninformed traders have to guess the probability that the constraint binds, and the constraint binds for high prices. Therefore, low prices are more informative than high prices.[12]

1.4 Short-sale constraints and conditional variance

According to the prevailing view, the introduction of short-sale constraints reduces the informativeness of the market price, which is confirmed by the analysis of the previous section. Indeed, (1.15) states that the perceived uncertainty of uninformed traders increases with a partial ban on shorting. This section investigates the effect of short-sale constraints, when an outside observer (e.g. a creditor from Section 1.5) with additional private information tries to learn from the market price. I show that in presence of short-sale constraints the information content of the market price (which constitutes a public signal) varies with the private signal of this agent. In particular, if this information content is measured by the variance conditional on the private and the public signal, then it is a non-monotonic function of the private signal. Moreover, for some private signal realizations the conditional variance is lower in presence of short-sale constraints than for the same private signal in absence of the constraint.

First, I restate the equilibrium price provided in (1.11), with the emphasis on the information content of the price, characterized by the pdf of the payoff, conditional on observing only the market price p:

Proposition 7 *A piecewise linear REE of the financial market exists with*

$$P = f + \begin{cases} A\,(u - C) & \text{if } u < C \\ B\,(u - C) & \text{if } u \geq C, \end{cases}$$

where the equilibrium constants A, $B \geq A$, and C are uniquely determined. Moreover, conditional on the price observation, the payoff f is only 'locally' Gaussian, with a jump

[12] The model presented in this paper here does not cover the $w = 1$ case, which is the subject of Bai et al. (2006) and Marin and Olivier (2008). When the constraint binds, the aggregate demand of all rational traders would be price-inelastic, and hence no price could clear the market with the random noise trading.

around the price realization p:

$$g\left(f|P=p\right)=\frac{1}{B\sigma_u}1_{f<p}\phi\left(\frac{f-(p+BC)}{B\sigma_u}\right)+\frac{1}{A\sigma_u}1_{f\geq p}\phi\left(\frac{f-(p+AC)}{A\sigma_u}\right). \tag{1.16}$$

When $w=0$, that is $A=B$, the conditional distribution simplifies to a normal distribution:

$$g\left(f|P_0=p\right)=\frac{1}{A\sigma_u}\phi\left(\frac{f-(p+AC_0)}{A\sigma_u}\right).$$

Figure 1-2 illustrates the distribution of f conditional on the market-clearing price p in absence and presence of short-sale constraints. The left panel shows that the distribution without short-sale constraints is normally distributed with mean $p+AC_0$ and precision $\tau_{Au}\equiv 1/\left(A^2\sigma_u^2\right)$. The right panel shows that under short-sale constraints the distribution is only locally normal, but not globally. For states of the world when the constraint does not bind, i.e. $f\geq p$, it is normally distributed with mean $p+AC$ and precision τ_{Au}. For states of the world when the constraint binds in the financial market, i.e. $f<p$, it is normally distributed with mean $p+BC$ and precision $\tau_{Bu}\equiv 1/\left(B^2\sigma_u^2\right)$. The variance increases, that is the precision decreases, because in this case there is less informed trading in the market.

Consider now a creditor who, besides observing the market price realization p, is also endowed with private signal t. For the tractability of the analysis, I assume that this private signal is given by $t=f+\xi$, where $\xi\sim N\left(0,\sigma_t^2=\tau_t^{-1}\right)$.

Suppose first that there are no short-sale constraints in the market. Due to the jointly Gaussian distribution of f, t and p, the inference problem of the agent is simple: her posterior about the f is normally distributed with mean $\frac{\tau_t}{\tau_t+\tau_{Au}}t+\frac{\tau_{Au}}{\tau_t+\tau_{Au}}\left(p+AC_0\right)$ and precision $\tau_t+\tau_{Au}$. That is, her conditional pdf is given by

$$g\left(f|t,P_0=p\right)=\frac{1}{\left(\tau_t+\tau_{Au}\right)^{-1}}\phi\left(\frac{f-\left[\frac{\tau_t}{\tau_t+\tau_{Au}}t+\frac{\tau_{Au}}{\tau_t+\tau_{Au}}\left(p+AC_0\right)\right]}{\left(\tau_t+\tau_{Au}\right)^{-1/2}}\right), \tag{1.17}$$

and due to the characteristics of normal distributions, her conditional variance is independent of the private signal realization t:

$$Var\left[f|t,P_0=p\right]=\left(\tau_t+\tau_{Au}\right)^{-1}.$$

Consider now the case with short-sale constraints. A creditor must combine her private signal t, which is normally distributed, with the public signal p, whose distribution is only locally normal, given in (1.16). A simple application of Bayes' rule implies that her posterior becomes

$$g\left(f|t, P=p\right) = \pi \cdot g\left(f|t, P=p, f<p\right) + (1-\pi) \cdot g\left(f|t, P=p, f \geq p\right), \tag{1.18}$$

where $\pi \equiv \Pr\left(f<p|t,p\right)$ and $1-\pi \equiv \Pr\left(f \geq p|t,p\right)$ are the probabilities the agent assigns to the constraint binding or not, respectively, and the conditional pdfs $g\left(f|t, P=p, f<p\right)$ and $g\left(f|t, P=p, f \geq p\right)$ belong to truncated normal distributions on the respectable ranges $f<p$ and $f \geq p$. In particular,

$$g\left(f|t, P=p, f<p\right) = 1_{f<p} \frac{1}{\left(\tau_t+\tau_{Bu}\right)^{-1/2}} \frac{\phi\left(\frac{f-\frac{\tau_t t+\tau_{Bu}(p+BC)}{\tau_t+\tau_{Bu}}}{\left(\tau_t+\tau_{Bu}\right)^{-1/2}}\right)}{\Phi\left(\frac{p-\frac{\tau_t t+\tau_{Bu}(p+BC)}{\tau_t+\tau_{Bu}}}{\left(\tau_t+\tau_{Bu}\right)^{-1/2}}\right)}$$

is the pdf of a truncated normal distribution with mean $\frac{\tau_t}{\tau_t+\tau_{Bu}}t + \frac{\tau_{Bu}}{\tau_t+\tau_{Bu}}\left(p+CB\right)$ and precision $\tau_t+\tau_{Bu}$, because if the short-sale constraint binds in the financial market, the price equals $p = f + B\left(u-C\right)$. Similarly,

$$g\left(f|t, P=p, f \geq p\right) = 1_{f \geq p} \frac{1}{\left(\tau_{Au}+\tau_t\right)^{-1/2}} \frac{\phi\left(\frac{f-\frac{\tau_t t+\tau_{Au}(p+AC)}{\tau_t+\tau_{Au}}}{\left(\tau_t+\tau_{Au}\right)^{-1/2}}\right)}{1-\Phi\left(\frac{p-\frac{\tau_t t+\tau_{Au}(p+AC)}{\tau_t+\tau_{Au}}}{\left(\tau_t+\tau_{Au}\right)^{-1/2}}\right)}$$

is the pdf of a truncated normal distribution with mean $\frac{\tau_t}{\tau_t+\tau_{Au}}t + \frac{\tau_{Au}}{\tau_t+\tau_{Au}}\left(p+AC\right)$ and precision $\tau_t+\tau_{Au}$, because if the short-sale constraint does not bind in the financial market, the price equals $p = f + A\left(u-C\right)$.

The conditional variance of a creditor in presence of short-sale constraints, as a function of the private signal t is illustrated on Figure 1-3. In general, the computation of this conditional variance becomes analytically intractable, but Figures 1-4 and 1-5 help to understand the intuition behind it.

Let us fix p, and consider two special cases. First, suppose that the creditor receives a much higher private signal, i.e. $t \to \infty$. As it implies

$$\lim_{t\to\infty} \pi = \Pr\left(f < p | t, p\right) = 0 \text{ and } \lim_{t\to\infty} \Pr\left(f \geq p | t, p\right) = 1,$$

the creditor is sure that the constraint does not bind in the financial market, which implies that all informed traders trade, and hence the precision of the price signal is τ_{Au}. Therefore, the posterior precision of the information available to her is given by $\tau_t + \tau_{Au}$, as if there were no short-sale constraints in the market at all. Figure 1-3 illustrates that in this case the conditional variances do not differ with and without the constraint.

Suppose now that the creditor receives a private signal much lower than the market price, i.e. $t \to -\infty$. It implies that

$$\lim_{t\to-\infty} \pi = \Pr\left(f < p | t, p\right) = 1 \text{ and } \lim_{t\to-\infty} \Pr\left(f \geq p | t, p\right) = 0,$$

hence she is sure that the constraint binds in the financial market, which implies that only a subset of informed traders trade, and hence the precision of the price signal is τ_{Bu}, lower than without the short-sale constraints. Therefore, the posterior precision of the information available to her is given by $\tau_t + \tau_{Bu}$, again lower than without the constraint. Put it differently, her posterior variance, as illustrated on Figure 1-3, increases. Thus, short-sale constraints decrease price informativeness for agents with private signals much smaller than the price realization.

Finally, consider the cases when the two signals are close to each other. Figures 1-4 and 1-5 illustrate the change in the conditional distribution due to short-sale constraints for a creditor with a private signal greater than the price, $t > p$, and for a creditor with a private signal smaller that the price, $t < p$, respectively. Contrasting the signal distributions without and with short-sale constraints, it is easy to see that when $t > p$, the introduction of short-sale constraints means that the creditor puts smaller weights on low payoff states of the world that she would consider unlikely based only on her private signal. The reason for this is that when constraint binds in the security market, i.e. when $f < p$, the demand shock is amplified due to the short-sale constraint. Therefore, the same price realization means a lower fundamental. However, if $t > p$, the private signal of the creditor suggests that the fundamental is high, thus low fundamental states are even more improbable. The agent knows that the market price is more likely to contain more positive information in general, therefore she becomes more certain about the payoff being high. Short-sale constraints hence confirm and strengthen her private information, as Figure 1-4 suggests. When $t \to \infty$, this effect gets weaker, and in the limit disappears. Hence, for $t \to \infty$, short-sale constraints do not alter the conditional distribution, and thus the conditional variance is not affected either.

When $t < p$, the opposite effect arises. Comparing the signal distributions without and with short-sale constraints, when $t < p$, the introduction of short-sale constraints means that the creditor puts larger weights on states of the world that she thought to be unlikely based on her private signal. That is, short-sale constraints force the agent to consider some previously irrelevant states of the world. She knows that the market price is more likely to contain more positive information in general, therefore when her private signal is below the price realization, her uncertainty about whether the constraint binds in the financial market increases, and hence her uncertainty about the payoff increases, too. Short-sale constraints dispute her private information, and hence weaken her posterior precision, as Figure 1-5 suggests. When $t \to -\infty$, this effect gets weaker, and in the limit disappears. However, the agent becomes certain that the constraint binds, and in this case the precision of the price signal is lower. Hence, for $t \to -\infty$, short-sale constraints alter the conditional distribution by affecting its precision, and thus the conditional variance increases.

As the following proposition states, short-sale constraints can increase the information content of the price for high enough private signal realizations, measured by the variance conditional on the private and the public signal:

Proposition 8 *There exists constant $\bar{t}$ such that the posterior variance of the asset payoff conditional on observing price p and private signal t, $Var\,[f|t,p]$, is lower under short-sale constraints (i.e. when $w > 0$ or $B > A$), if and only if $t - p \geq \bar{t}$. The threshold $\bar{t}$ is a decreasing function of w, and* $\lim_{\sigma_t} \bar{t} = 0$.

The following section studies how this non-monotonic change in the conditional variance due to short-sale constraints affects coordination in a game with strategic complementarities.

1.5 Economy with a financial market and creditors

This section extends the previous setup by embedding a coordination game between dates 0 and 1. Suppose that a financial institution (e.g. investment bank, or bank, for short) is financed through a combination of short-term and long-term debt. Long-term debt holders are passive - in the past they have decided to provide capital that cannot be withdrawn. Short-term debt matures at date $t = 1$, on which occasion it can be renewed.

The state of fundamentals is characterized by θ that is interpreted as the cash-flow the bank's assets generate at date 1. Higher values of θ correspond to higher quality/liquidity projects. I assume that the bank has outstanding debt with size normalized to 1, from which the short-term debt amounts to ω and the long-term debt is $1-\omega$. Short-term debt holders (creditors, for short from now on) can decide to roll over their debt. For simplicity, I assume that the bank's assets generate sufficiently large cash-flows in the long run, but they only have θ to pay out creditors who demand capital payoff at date 1. Therefore, the bank remains solvent if and only if $\theta \geq \omega(1-I)$, where I denotes the proportion of creditors who roll over, and hence $\omega(1-I)$ is the amount to be paid out to creditors who recall their loans.

Creditors are a continuum of risk-neutral agents with measure one, and indexed by $j \in [0,1]$.[13] Each creditor can choose between two actions. They either provide capital (i.e. roll over the short-term debt), $i_j = 1$, the risky action, or refrain from doing so (i.e. recall the loan or withdraw money), $i_j = 0$, the safe action. The net payoff from withdrawing is normalized to zero. The net payoff from lending to the bank is $1-c$ if the bank remains solvent and $-c$ otherwise, where $c \in (0,1)$ parametrizes the private costs of lending, which can be interpreted, for example, as transaction costs, administrative fees, or taxes.[14] It follows that the payoff of creditor j is

$$U(i_j, I, \theta) = i_j \left(1_{\theta \geq \omega(1-I)} - c\right), \tag{1.19}$$

[13] The security market and debt market (i.e. the capital provision environment of creditors) are assumed to be segmented markets, that is the asset price is fully exogenous from the point of view of creditors, and hence it does not incorporate their private information, as in Angeletos and Werning (2006). See the discussion later.

[14] As creditors are assumed to be risk-neutral, this setting is equivalent to any set of payoffs $\{\pi_H, \pi_L, \pi_0\}$, where providing capital pays either π_H in case of the bank remaining solvent and $\pi_L < \pi_H$ in case of failure, while recalling the loan gives a sure payoff π_0 that satisfies $\pi_L < \pi_0 < \pi_H$. The utility of a creditor in this setting would simply be a linear function of the utility given in (1.19), and hence would lead to the same optimal action.

where $1_{\theta \geq \omega(1-I)}$ is the indicator of the bank remaining solvent, and takes the value of 1 if $\theta \geq \omega(1-I)$ and 0 otherwise.[15]

If creditors know the value of θ perfectly before making their decision, there exist a tripartite classification of the state, in the spirit of Obstfeld (2004). Based on this, the optimal strategy of creditors is as follow: If $\theta \leq 0$, then the dominant strategy is to withdraw deposits from the bank, irrespective of what other capital providers do, because the bank always fails. In turn, if $\theta \geq \omega$, then the dominant strategy is to give money to the bank, irrespective of what other creditors do, because it always remains solvent. When the bank asset value θ lies in the interval $(0, \omega)$, there is a coordination problem among capital providers. On one hand, if every other creditor rolls over the debt, the bank survives, and lending yields more than withdrawing: $1 - c > 0$. On the other hand, if every other creditor withdraws, the bank fails, and withdrawing yields more than financing the bank: $0 > -c$. Therefore, both $I = 1$ and $I = 0$ is an equilibrium whenever $\theta \in (0, \omega)$: the former outcome represents the first best, while the latter is considered a coordination failure. In this interval the bank's future depends on the size of the credit run.[16]

Following standard global game setups in the spirit of Carlsson and van Damme (1993) and Morris and Shin (1998), I assume that information is imperfect, so that the state θ is not common knowledge. In the beginning of the game, nature draws θ from a diffuse uniform distribution over the real line, which constitutes the agents' initial common prior about the state of the world. Investor j then receives a private signal $t_j = \theta + \xi_j$, where ξ_j has a Gaussian distribution with mean 0 and standard deviation σ_t, ξ_j is independent of θ, and independently and identically distributed across short-term debt holders. The precision of the private signal is given by $\tau_t = 1/\sigma_t^2$.

To connect the security trading and the credit run, I assume that the payoff of the asset,

[15] The coordination setup presented here is a simplified version of models on bank runs, e.g. Diamond and Dybvig (1983), Rochet and Vives (2004) and Goldstein and Pauzner (2005); or Morris and Shin (2004), who study coordination among creditors of a distressed borrower. In contrast to those papers, I choose to work with a parsimonious model, as my aim is to analyze the effect of short-sale constraints on coordination, instead of providing a more realistic setting. In particular, I will abstract away from the first mover's advantage and demand-deposit insurance, emphasized by Diamond and Dybvig (1983) and Goldstein and Pauzner (2005), or the price at which the debt is issued, as in Morris and Shin (2004).

[16] There exist other interpretations of coordination, here presented using the terminology of creditor runs. In models of currency crises, as in Obstfeld (1986, 2004) or Morris and Shin (1998), speculators decide whether to attack a currency by shorting it. Chamley (1999), Morris and Shin (1999) and Dasgupta (2007) consider investment complementarities.

f, correlates with the quality of the banks' assets. Thus, the price of the financial asset, p, can provide additional (public) information regarding the state of the world beyond the private signals, and hence can facilitate or hurt coordination among capital providers. For simplicity, I assume $f = \theta$, and think about the traded asset as the (only) security that the bank has on the asset side of the balance sheet, a zero-net-supply financial derivative on the bank's equity, or as an industry index that includes the bank. Thus, the price is an exogenous signal from the viewpoint of creditors, in the sense that the , but is nevertheless correlated with the fundamental θ.

Because the two parts of the economy are segmented, with an information spillover from the financial market to the credit run in the form of the price p, without the outcome of the coordination game affecting the market price, the equilibrium of the whole economy is also separable into two parts. In fact, it is just a simple conjugate of the equilibrium of the trading round, defined in Definition 1 and discussed in Section 1.3, and the equilibrium of the run, conditional on the realization of the market price. Thereby, I only define the equilibrium of the coordination game:

Definition 9 *Let p denote the price of the asset with payoff $f = \theta$ emerging from the financial market. A perfect Bayesian equilibrium of the credit run consists of individual strategies for investing, $i(t_j, p)$, and the corresponding aggregate, $I(\theta, p)$, such that*

1. *decision is optimal for creditors:*

$$i(t_j, p) \in \arg\max_{i \in \{0,1\}} E[U(i, I(\theta, p), \theta) | t_j, p] \text{ for } j \in [0,1]; \tag{1.20}$$

2. *proportion of short-term debt rolled over is*

$$I(\theta, p) = \textstyle\int_0^1 i(t_j, p)\, dj; \tag{1.21}$$

3. *agents update their beliefs according to Bayes' rule.*

This section hence only solves for the equilibrium of a standard global game setup with private and public information, fully characterized by conditions (1.20) and (1.21). Combining the equilibrium of the credit run with the equilibrium of the financial market would provide an equilibrium of the whole economy.

I restrict my attention to monotone equilibria, defined as perfect Bayesian equilibria such that, for a given realization p of the public signal, a creditor provides capital to the bank if and only if the realization of her private signal is at least some threshold $t^*(p)$; that is $i(t_j, p) = 1$ iff $t_j \geq t^*(p)$. It implies that the bank can be characterized in a similar way: the bank with asset quality θ survives if and only if this quality is higher that some threshold $\theta^*(p)$; formally if $\theta \geq \theta^*(p)$.[17]

1.6 Credit runs and portfolio constraints

After trading in the financial market has taken place, but before the payoff at date 1 happens, creditors decide whether to roll over short-term debt, thereby providing capital to the bank in need of liquidity, or to withdraw it. Since the payoff of the financial asset f and the bank asset value θ are correlated, the equilibrium price of the financial market, p, provides an observable public signal regarding the unknown parameter θ, and creditors can coordinate their actions based on it. In the following subsections, I solve the coordination model, first without constraints on short-selling, then with the short-sale constraints.

1.6.1 Equilibrium analysis with no short-sale constraints

In this section I provide a solution to the coordination game among capital providers when short-selling is allowed for everyone. To pin down the equilibrium of the model, characterized by the pair $\{t^*, \theta^*\}$, I solve for the optimal θ^* while taking t^* as given, and for the optimal t^* if θ^* is assumed to be given. The joint solutions of these two conditions describe the equilibria of the credit run.

In a monotone equilibrium described above, creditors with private signals $t_j \geq t^*$ provide capital. Based on the joint distribution of θ and the private signals t_j, the aggregate proportion of creditors who roll over is given by

$$I(\theta, p) = \Pr(t \geq t^*(p) | \theta) = 1 - \Phi\left(\sqrt{\tau_t}\left(t^*(p) - \theta\right)\right).$$

[17] My results concerning multiple equilibria are obtained even within this restricted class. Moreover, in absence of short-sale constraints, uniqueness within this class implies overall uniqueness, see Morris and Shin (1998, 1999).

The right hand side of this equation increases in θ, therefore a better bank receives more capital rolled over. The bank avoids bankruptcy if and only if $\theta \geq \theta^* (p)$, where $\theta^* (p)$ is the quality of the marginal bank that solves $\theta = \omega (1 - I(\theta, p))$. Therefore,

$$t^* (p) = \theta^* (p) + \frac{1}{\sqrt{\tau_t}} \Phi^{-1} \left(\frac{\theta^* (p)}{\omega} \right). \tag{1.22}$$

Condition (1.22) characterizes the banks that survives withdrawals for a given switching strategy $t^* (p)$. There are several remarks to be made about this equation. First, notice that the right-hand side of (1.22) is strictly increasing in $\theta^* (p)$, therefore there is a unique $\theta^* (p)$ that satisfies the equation for a given $t^* (p)$. Secondly, the bank survival threshold $\theta^* (p)$ is an increasing function of the creditor cutoff $t^* (p)$, as a lower switching strategy from creditors implies more capital rolled over, and hence a bank with lower asset payoff surviving. Thirdly, as

$$\frac{dt^* (p)}{d\theta^* (p)} = 1 + \frac{1}{\sqrt{\tau_t} \omega} \frac{1}{\phi \left(\Phi^{-1} \left(\frac{\theta^* (p)}{\omega} \right) \right)} > 1,$$

it must be that $d\theta^* / dt^* < 1$. The presence of strategic complementarities implies that any increase in the cutoff $t^* (p)$ results in a smaller increase in the marginal bank's value, because no creditor can be certain about the signals received by others and hence the strategy of others. Finally, in the limit when private signals become arbitrarily precise, $\tau_t \to \infty$, creditors become certain about others' signals as well, and the bank survival threshold $\theta^* (p)$ becomes exactly the individual capital provision threshold $t^* (p)$.

Next, consider the derivation of the equilibrium cutoff strategy $t^* (p)$ as a function of the threshold $\theta^* (p)$. Creditors receive payoff 1 if the bank avoids distress, and 0 if not, while paying a cost c. Because they do not observe the state θ directly, the payoff from rolling over the loan must be calculated from the posterior distribution over the states, conditional on the private and public signal. If creditor j knows that the bank solvency threshold is $\theta^* (p)$, she assigns probability $\Pr (\theta \geq \theta^* (p) | t_j, p)$ to the bank surviving, based on all her information, which implies that the expected payoff from rolling over is $\Pr (\theta \geq \theta^* (p) | t_j, p) - c$. As withdrawing yields a payoff normalized to 0, the signal of the marginal agent, who is indifferent between withdrawing or not, must solve the indifference condition $\Pr (\theta \geq \theta^* (p) | t_j, p) = c$.

In absence of short-sale constraints, the market price is $p = \theta + A(u - C_0)$. As u is normally distributed with mean zero and precision $\tau_u = 1/\sigma_u^2$, the precision of the price signal is $\tau_{Au} \equiv \tau_u/A^2$. Therefore, the posterior of agent j about θ is normally distributed with mean $\frac{\tau_t}{\tau_t+\tau_{Au}} t_j + \frac{\tau_{Au}}{\tau_t+\tau_{Au}} (p + AC_0)$ and precision $\tau_t + \tau_{Au}$. Thus, the indifference condition becomes

$$\Phi\left(\sqrt{\tau_t + \tau_{Au}}\left(\theta^*(p) - \frac{\tau_t}{\tau_t + \tau_{Au}} t^*(p) - \frac{\tau_{Au}}{\tau_t + \tau_{Au}}(p + AC_0)\right)\right) = 1 - c,$$

which is equivalent to

$$\theta^*(p) = \frac{\tau_t}{\tau_t + \tau_{Au}} t^*(p) + \frac{\tau_{Au}}{\tau_t + \tau_{Au}}(p + AC_0) + \frac{1}{\sqrt{\tau_t + \tau_{Au}}}\Phi^{-1}(1 - c), \tag{1.23}$$

and implies a linear relationship between θ^* and t^*. Figure 1-6 illustrates the the critical mass condition, (1.22), and the individual optimality condition without short-sale constraints, (1.23), respectively.

An equilibrium is the joint solution to conditions (1.22) and (1.23), which lead to

$$\frac{\tau_{Au}}{\sqrt{\tau_t}}\theta^*(p) - \Phi^{-1}\left(\frac{\theta^*(p)}{\omega}\right) = \sqrt{1 + \frac{\tau_{Au}}{\tau_t}}\Phi^{-1}(1 - c) + \frac{\tau_{Au}}{\sqrt{\tau_t}}(p + AC_0).$$

As the left-hand side of the equation is a continuous function of θ^*, which takes the value $-\infty$ for $\theta^* = \omega$ and ∞ for $\theta^* = 0$, the equation always has a solution. Moreover, the solution is unique for every p_0 if and only if the left-hand side of the equation is a strictly decreasing function of θ^*, that is if and only if $\tau_{Au} \leq \sqrt{2\pi\tau_t}$.[18]

The following proposition states the above result:

Proposition 10 (Morris and Shin) *In absence of short-sale constraints, the equilibrium is unique if and only if the private noise is small relative to the price noise, that is for* $\sigma_t \leq \sqrt{2\pi}A^2\sigma_u^2$. *Moreover, in the limit as private noise vanishes so that* $\sigma_t \to 0$, *a creditor with private signal below* $t^*(p) = c\omega$ *recalls her loan, and the bank with asset quality below* $\theta^*(p) = c\omega$ *fails.*

Proposition 10 confirms the uniqueness result of Morris and Shin (1999, 2001). For any positive level of noise in the public signal, $\sigma_u > 0$, uniqueness is ensured by sufficiently

[18]As Figure 1-6 suggests, this is equivalent to the slope of the critical mass condition (1.22) always being below the slope of the individual optimality condition, (1.23).

small noise in the private signal. The intuition is that as the private signal becomes much more precise than the public signal, creditors stop relying on the public signal and use only their private information. This implies that the equilibrium dependence on the common noise component u vanishes, and makes it harder to predict the actions of others, heightening strategic uncertainty. When strategic uncertainty is strong enough, multiplicity breaks down. It is interesting to note that the equilibrium run size and outcome outcome does not depend on public signal p (or common noise component u), which is the second finding of Morris and Shin (1999). In what follows, I will refer to this equilibrium as the informationally efficient equilibrium,. It is important to mention that this informationally efficient equilibrium is different from the first best or allocationally efficient equilibrium, i.e. $I^* = \omega$ and $\theta^* = 0$. This difference is due to the presence of the coordination externality.

1.6.2 Equilibrium analysis with short-sale constraints

As shown in Section 1.3, the introduction of short-sale constraints has an adverse effect on the market price. The fact that the price reveals information about the payoff at different rates for high and low realizations of the demand shock implies that short-sale constraints notably change the inference problem of creditors, as presented in Section 1.4.

To solve for the equilibrium in presence of short-sale constraints, one needs to repeat the steps of the previous subsection. First, given that the joint distribution of the state θ and the private signals does not change, the critical mass condition (1.22) that determines the quality of the marginal bank as a function of individual strategies, does not change either. However, short-sale constraints do affect the posterior of creditors after observing both the price and the private signal. The public signal p is now only locally Gaussian, but not globally, as given in (1.16) and illustrated on Figure 1-2.

As shown in Section 1.4, in presence of short-sale constraints the posterior pdf can be given in the following way (see (1.18)):

$$g\left(\theta|t, P=p\right) = \pi \cdot g\left(\theta|t, P=p, \theta<p\right) + (1-\pi) \cdot g\left(\theta|t, P=p, \theta \geq p\right).$$

In this equation, $\pi \equiv \Pr\left(\theta < p|t, P=p\right)$ and $1-\pi \equiv \Pr\left(\theta \geq p|t, P=p\right)$ denote probabilities that the creditor with private signal t associates with the short-sale constraint binding in the financial market or not, respectively. Moreover, the conditional pdfs $g\left(\theta|t, P=p, \theta<p\right)$ and $g\left(\theta|t, P=p, \theta\geq p\right)$ belong to the class of truncated normal distributions, with means $\frac{\tau_t}{\tau_t+\tau_{Bu}}t+\frac{\tau_{Bu}}{\tau_t+\tau_{Bu}}\left(p+CB\right)$ and $\frac{\tau_t}{\tau_t+\tau_{Au}}t+\frac{\tau_{Au}}{\tau_t+\tau_{Au}}\left(p+CA\right)$, and precisions $\tau_t+\tau_{Bu}$ and $\tau_t+\tau_{Au}$, respectively, because in the first case the creditor knows the short-sale constraint binds in the financial market, and hence the price equals $p=\theta+B\left(u-C\right)$, while in the second case this creditor knows the constraint does not bind, and hence the price equals $p=\theta+A\left(u-C\right)$.

As before, the expected net payoff of agent j from providing capital to the bank, for a fixed success threshold θ^*, is $\Pr\left(\theta>\theta^*|t_j, P=p\right)-c$ and hence t^* must solve the indifference condition $\Pr\left(\theta\geq\theta^*|t_j, P=p\right)=c$, which is equivalent to

$$\theta^*\left(p\right)=\begin{cases}\frac{\tau_t}{\tau_t+\tau_{Bu}}t^*\left(p\right)+\frac{\tau_{Bu}}{\tau_t+\tau_{Bu}}\left(p+BC\right)+\frac{1}{\sqrt{\tau_t+\tau_{Bu}}}\Phi^{-1}\left(\left(1-c\right)\frac{\pi_B^*}{\pi^*}\right) & \text{if } \theta^*\left(p\right)\leq p\\ \frac{\tau_t}{\tau_t+\tau_{Au}}t^*\left(p\right)+\frac{\tau_{Au}}{\tau_t+\tau_{Au}}\left(p+AC\right)+\frac{1}{\sqrt{\tau_t+\tau_{Au}}}\Phi^{-1}\left(1-c\frac{1-\pi_A^*}{1-\pi^*}\right) & \text{if } \theta^*\left(p\right)>p,\end{cases}\tag{1.24}$$

where $\pi^*\equiv\Pr\left(\theta<p|t^*, p\right)$ is the probability the marginal agent assigns to the short-sale constraint binding in the market, $\pi_B^*\equiv\Pr\left(\theta<p|t^*, p=\theta+B\left(u-C\right)\right)$ is the probability that the marginal agent assigns to informed traders shorting/selling in a market with no constraints but volatility τ_{Bu}^{-1}, and $\pi_A^*\equiv\Pr\left(\theta<p|t^*, p=\theta+A\left(u-C\right)\right)$ is the probability that the marginal agent assigns to informed traders shorting/selling in a market with no constraints but volatility τ_{Au}^{-1}. It is easy to see that when there are no short-sale constraints, i.e. $B=A$, $\pi^*=\pi_A^*=\pi_B^*$, and (1.24) is equivalent to (1.23).

Figure 1-7 illustrates the critical mass condition, (1.22), and the individual optimality condition in presence of short-sale constraints, (1.24), respectively. The former displays the quality of the marginal bank, θ^*, given that creditors follow the threshold strategy with t^*, that is a capital provider leaves her money in the bank if and only if she receives private signal $t_j\geq t^*$. As the θ^* threshold is determined only by the joint distribution of the fundamental and the private signals, short-selling constraints do not alter it.

What changes is the optimal switching strategy of creditors for a fixed θ^* bank solvency threshold. However, as the distributions are not jointly Gaussian, the posterior (1.18) is not Gaussian any more, and hence it is not possible to simplify condition (1.24) further more, and to provide a simple necessary and sufficient condition for the number of equilibria. The reason for this is that, as seen on Figure 1-7, the slope of (1.24) is not monotonic. Instead, short-sale constraints create a hump shape on the individual optimality condition, with the slope $d\theta^*/dt^*$ of the individual optimality condition taking values between the upper slope at the kink, $d\theta^*/dt^*|_{\theta^* \to p+}$, where it is clearly the smallest, and when $\theta^* \to -\infty$, where it is the largest, $\frac{\tau_t + \tau_{Bu}}{\tau_t}$. Thus, a sufficient condition for uniqueness would be that

$$\frac{d\theta^*}{dt^*}|_{\theta^* \to p+} > \frac{1}{1 + \frac{1}{\sqrt{\tau_t}\omega}\sqrt{2\pi}}. \tag{1.25}$$

However, as shown in Appendix 1.9.3, there exists a constant $\overline{\sigma}_t > 0$ such that for every $0 < \sigma_t < \overline{\sigma}_t$, there is a price realization p such that (1.25) does not hold, and hence the individual optimality condition and the critical mass condition have three intersections. Thereby, there are multiple equilibria of the system of equations (1.22) and (1.24). The following proposition formally states this result:

Proposition 11 *In presence of short-sale constraints, there are multiple equilibria in investment strategies when σ_t is sufficiently small. Moreover, multiplicity remains as private noise vanishes so that $\sigma_t \to 0$: the switching strategies become*

$$t^*(p) = \theta^*(p) = \begin{cases} c\omega & \text{for all } p \\ p & \text{if } \beta c\omega < p < c\omega, \end{cases} \tag{1.26}$$

and hence for every $\theta \in [p, c\omega)$ both the informationally efficient equilibrium and a 'high capital provision' equilibrium exist whenever $\beta c\omega \leq p \leq c\omega$, where $\beta = \frac{A^2}{(1-c)B^2 + cA^2} < 1$.

The technical bits of the proof are in Appendix 1.9.3.

The informationally efficient equilibrium is the same as the unique equilibrium of the unconstrained economy: a bank with asset quality above $c\omega$ remains solvent. However, there exists an equilibrium with more capital provision: creditors also finance banks with lower asset quality, between the public signal realization p and $c\omega$. This is an informationally inefficient equilibrium, as agents put excessive weight on the public signal. It is characterized by overinvestment compared to the informationally efficient equilibrium, because agents with lower signals provide capital too, hence I refer to it as the 'high investment' equilibrium.

1.7 Discussion

In this section I discuss my results on multiplicity, allocational efficiency, provide comparative statics, and refer to some policy implications.

1.7.1 Multiplicity

Canonical papers in the literature on transparency show that releasing more information is not necessarily good. Indeed, in Morris and Shin, while without a public signal the market may be in a uniqueness region, by adding a precise enough public signal, the economy has multiple equilibria.

Since Morris and Shin (2001), several authors have considered ways that reinstate multiplicity in coordination games. The existing literature mainly focuses on the endogenous nature of the public signal. For example, Angeletos and Werning (2006) study financial market prices, which aggregate the dispersed information of agents, or direct noisy signals about others' activity. Information aggregation can overturn the Morris and Shin (1998) uniqueness result and lead to multiplicity if the precision of public information increases faster than the precision of the private information. Hellwig et al. (2006) and Tarashev (2007) also study coordination games with financial prices being endogenous public signals. Because all these papers stay in the class of jointly Gaussian distributions, the informational content of the public signal does not vary for its different realizations and hence across the multiple equilibria.

In contrast, the model presented here provides a fundamentally different setting. What is cardinal for the analysis is that agents with different private signals interpret the same public signal in different ways. In particular, the information they infer from the public signal changes with the distance of their private and public signal. Holding the price constant and increasing the private signal can provide more information about the composition of the market price: a high price is more likely to be the result of a high demand shock than a low price to be the result of a low demand shock, because in the first case the fewer informed traders have a smaller corrective effect on the market price.

According to the prevailing view, the introduction of short-sale constraints reduces the informativeness of the market price, i.e. decrease its precision, and hence, following the

Morris and Shin logic, should not lead to coordination failures. Indeed, (1.15) states that the perceived uncertainty of uninformed traders increases with a ban on shorting. The surprising finding of this model is that, in contrast to the existing literature, I show that short-sale constraints can make asset prices contain more information for some creditors with additional information, as demonstrated in Proposition 8.[19]

Although both the setup and the motivation are different, the results of this chapter are close in spirit to Angeletos et al. (2006). They examine the informational role of policy decisions in a coordination setting. They show that policy interventions create endogenous public information and can lead to multiple equilibria.[20] There are two differences though. First, in their paper the public signal reveals that the state of the world is neither too high, nor too low. In contrast, short-sale constraints 'help' to rule out only lower states of the world by making creditors' posterior distributions more left-skewed. This has strong implications on allocational efficiency, discussed below. Second, in their analysis the signal is the equilibrium action of a policy maker, whereas the present article takes the constraint as given. I show that, even abstracting from signaling and analyzing the constraint on short-selling as an endogenous decision of regulators, short-sale constraints are nevertheless capable of suggesting that prices, influenced by demand shocks, are lower than economic fundamentals would imply. It would be interesting to see how introducing signaling (i.e. endogenizing the authority's decision to introduce short-sale constraints in a security market) would influence the results of the model.

[19] It is interesting to refer back to similarities and differences with Bai et al. (2006), and Marin and Olivier (2008). What is crucial in the analysis is that short-sale or other trading constraints result in a varying information content across different price levels. Therefore, even if the asset pricing implications of the two types of models are different, qualitative results, such as the increasing information content of the price under short-sale constraints for some agents with additional private information, and the possibility of multiple equilibria would not be affected. However, with a financial market model, where high asset prices are more informative than low prices, the informationally efficient equilibrium would be allocationally more efficient as well.

[20] The two types of equilibria that they identify are also in line with the findings of this paper. Their inactive-policy equilibrium, where agents coordinate on a strategy that is insensitive to the policy, is analogous to my informationally efficient level of creditor run, and their continuum of active-policy equilibria correspond to equilibria when capital provision depends on the price p.

1.7.2 Efficiency

An interesting result of the chapter is that in the second equilibrium creditors always provide more capital than in the informationally efficient equilibrium. It is shown in 1.26: the second equilibrium only exists for $\beta c\omega < p < c\omega$.

The intuition is the following. As shown in Proposition 8, asset prices under short-sale constraints provide more information to creditors with high private signals. First, consider the case when $p > c\omega$. Without short-sale constraints, in the informationally efficient (unique) equilibrium, creditors rely only on their private signals, and based on their assessments about the bank's asset value, agents with private signals $t_j \geq c\omega$ provide capital to the bank. In presence of short-sale constraints, agents with private signals above p will become more informed about both the fundamental θ and hence about other creditors' beliefs. Therefore short-sale constraints weaken strategic uncertainty among these creditors, leading to the possible multiplicity of equilibria. In one equilibrium they all provide capital, but as they all provide capital in the informationally efficient equilibrium, it would not change bankruptcy outcomes. In the other equilibrium they all refrain from doing so, which implies that only agents with medium signals (between $c\omega$ and p) would invest, which is not a monotone equilibrium. Therefore when $p > c\omega$, the uniqueness of the equilibrium survives.

Consider now the case when $p < c\omega$. In this case, agents with signals above p become more informed due to short-sale constraints. Those with signals above c will become more certain about what others do, which only reinforces their willingness to invest. The main difference is that now creditors with signals $p < t_j < c\omega$, who would have stayed out in absence of the constraint, obtain more precise information. Therefore strategic uncertainty weakens among these creditors, and they all become more informed about both the fundamental θ and about the beliefs of other creditors who have signal realizations between p and $c\omega$. Self-fulfilling beliefs and the resulting multiplicity hence arise in this group of creditors with medium realizations of the private signal. If they all stay out, we obtain an equilibrium equivalent to the informationally efficient equilibrium. However, there exist another equilibrium in which they all provide capital. In this second equilibrium creditors rely more on the public signal.

As the second equilibrium only emerges when $p < c\omega$, the second equilibrium, with bankruptcy threshold $\theta^* = p$ is closer to the first best ($\theta^* = 0$) than the informationally efficient equilibrium ($\theta^* = c\omega$). I conclude that short-sale constraints improve economic efficiency by mitigating the adverse effect of the coordination externality. In contrast to Morris and Shin (2002), who show that an increase in transparency might decrease welfare, short-sale constraints provide a 'good type' of transparency, recreating multiplicity only when it is desirable.[21]

1.7.3 Comparative statics and policy implications

The two main parameters of the coordination game are the proportion of informed investors barred from shorting, w, and creditors' private cost of providing capital, c.

As motivated in Section 1.2, the interpretation of w is quite broad. Here I focus mainly on regulatory restrictions such as a short-sale ban, the uptick rule, or legal restrictions on institutional trading. As shown in Section 1.6, w only affects the lower threshold for existence of the high investment equilibrium, through influencing

$$\beta = \frac{A^2}{(1-c)\,B^2 + cA^2},$$

which simplifies to

$$\beta = \frac{(1-w)^2}{(1-c) + c\,(1-w)^2}.$$

It is easy to verify that

$$\frac{\partial \beta}{\partial w} = -\frac{2\,(1-w)\,(1-c)}{\left[(1-c) + c\,(1-w)^2\right]^2} < 0,$$

i.e. tighter short-selling constraints lead to a higher probability of multiple equilibria. One interpretation of this multiplicity in the bankruptcy outcome is an increase in ex ante uncertainty about the outcome of the coordination, which can be interpreted as undesirable excess volatility. Clearly, to make predictions about the impact of certain policy measures, one needs to be able to find robust patterns across certain equilibria, as in Angeletos et al. (2008).

[21] This is clearly not a welfare analysis of the whole economy, which would have to take into account that short-sale constraints compromise market liquidity and price discovery, and certainly make constrained informed investors worse off.

The other crucial parameter of the coordination model is the private cost of capital provision, c. Parameter c affects the net benefit or loss for creditors if they choose the risky action. Clearly, a higher c makes capital provision less desirable from the point of view of creditors, which implies that in the informationally efficient equilibrium, with $\theta^* = c\omega$ in the limit, banks receive less capital and hence they need a higher asset quality to remain solvent. The effect on the lower threshold for a high investment equilibrium, $\beta c\omega$, is more subtle. After some simple algebra one finds that

$$\frac{\partial \beta}{\partial c} = (1-w)^2 \frac{1-(1-w)^2}{\left[1-c+c(1-w)^2\right]^2} \geq 0,$$

which also implies that as long as short-selling is restricted, i.e. $w > 0$, $\beta c\omega$ increases in c.

Moreover, one can characterize the benefit of short-sale constraints by the proportion of additional banks that get financed, $c\omega - \beta c\omega = (1-\beta)\, c\omega$, which can be interpreted as the ex ante probability of multiple equilibria. Here the lower bound, $\beta c\omega$ decreases in w, hence tighter short-sale constraints increase the benefits in the real economy. Furthermore, it satisfies

$$\frac{\partial}{\partial c}\left((1-\beta)\, c\omega\right) = \left[\left(\frac{1-c}{c}\right)^2 - (1-w)^2\right] \omega \frac{1-(1-w)^2}{\left[\frac{1-c}{c} + (1-w)^2\right]^2},$$

which implies an inverse U-shaped relationship. For small c values the derivative is positive, hence the ex ante probability of multiplicity, or the potential benefit of short-sale constraints, increases, while for c close to 1, this benefit decreases.

Finally, as the analysis of the previous section shows that tighter short-sale constrains can promote allocational efficiency, one can reflect on the short-sale bans around the globe in late 2008 and the following year. In fact, a sudden jump in c, implied for example by news that the investment opportunity worsens, increases the bankruptcy threshold for the bank. Introducing strict enough shorting restrictions, by increasing w, can create a second equilibrium and hence partly offset the increase in c. Empirical studies about the effect of short-sale bans in and after 2008, such as Boehmer et al. (2009) and Beber and Pagano (2011), conclude that if the SEC's and other regulators' goal with the short-sale ban was to artificially raise prices on financial stocks, they failed, and in the meantime compromised market quality. However, the SEC might have just been trying to avert a credit run on the largest investment banks. My model shows that while short-sale constraints increase market volatility, they can also affect the information that agents learn from prices, and can lead to outcomes where creditors do not withdraw money from low quality banks. Washington Mutual and Wachovia did go bankrupt during the 3-week shorting ban, collapsing under the weight of their bad loans, suggesting that their fundamentals were below the threshold $\beta c\omega$. But, while it is now clear that other financial firms such as Citigroup had extremely troubled fundamentals, the introduction of short-sale constraints could have contributed to their survival.

1.8 Concluding remarks

The model presented in this chapter examines the informational effects of short-sale constraints when asset prices provide guidance for decisions made in a coordination environment. I present a model that shows although short-selling constraints make asset prices more volatile and decrease price informativeness, they can provide more information for certain agents of the economy, who are endowed with additional private information too. Due to learning more in presence of short-sale constraints, creditors with moderate private signals are willing to lend more, if they think others with similar signals lend as well, which leads to a second equilibrium with higher allocational efficiency. My result thus implies that the decrease in average informativeness is more than compensated by an increase in informativeness to some agents.

The existing literature studying the effects of short-sale constraints identifies both benefits and detriments of these restrictions. The first group include prevention from speculative shorting that otherwise could lead to bear raids. On the other hand, introducing a ban on short-selling has been shown to decrease market liquidity and reduce price informativeness. In this chapter, I show that, allowing for a richer structure than in previous models, short-sale constraints can increase the information content of market prices. Although it leads to informational inefficiency in capital provision, it can increase allocative efficiency and prevent financial institutions from collapsing in uncertain times, when a fear of distress prevents creditors to roll over short-term debt. In particular, short-sale constraints improve the information of creditors with private signals above the market price realization. If this increase in precision is strong enough, short-sale constraints can create a second equilibrium in which creditors provide more capital, leading to less severe credit runs. My model hence suggests that emergency orders such as the one in September 2008 can increase efficiency even in absence of manipulative shorting, if the foregone costs of a potential collapse of part of the banking industry and systemic risk (i.e. the increase in allocational efficiency in the real economy) are large enough to dominate the costs of compromised market quality (i.e. the fall in informational efficiency in the financial market) in troubled times.

The model considered here studies information aggregation in a coordination game, with an external public signal emerging from a market subject to trading constraints. A more straightforward way to study information aggregation and portfolio restrictions would be to assume that investors with dispersed information are actually participants in the market, and hence the market price aggregates their information in presence of the short-selling constraint. Such a model must be more complicated because of the dual role of the price (for the inference and market clearing), but the present analysis suggests that it could shed more light on the interaction between asymmetric information and portfolio constraints.

1.9 Appendix

1.9.1 REE in the financial market

Optimal demands. Investor k's optimization problem is given by

$$\begin{aligned}\max_{x_k} U(W_k) &= E[W_k|\mathcal{I}_k] - \frac{\rho}{2} Var[W_k|\mathcal{I}_k] \\ &= x_k (E[f|\mathcal{I}_k] - p) - \frac{\rho}{2} x_k^2 \left(Var[f|\mathcal{I}_k] + \sigma_n^2\right).\end{aligned}$$

Solving the FOC without short-sale constraints, one obtains

$$x_k = \frac{E[f|\mathcal{I}_k] - p}{\rho \left(Var[f|\mathcal{I}_k] + \sigma_n^2\right)}.$$

From here the optimal demands for all three types of traders are straightforward. ■

Proof of Theorem 2. The derivation in the main text provides the step-by-step solution to the problem. There are three issues left for this appendix: (i) to derive the conditional expectations $E[f|p]$ and $E\left[f^2|p\right]$, and the conditional variance $Var[f|p]$, (ii) to prove the existence of the equilibrium, and (iii) to analyze uniqueness.

(i) The conditional distribution (1.9) implies that the expectation simply becomes

$$E[f|P=p] = \int_{-\infty}^{p} \frac{1}{B\sigma_u} f \phi\left(\frac{f-(p+BC)}{B\sigma_u}\right) df + \int_{p}^{\infty} \frac{1}{A\sigma_u} f\phi\left(\frac{f-(p+AC)}{A\sigma_u}\right) df = p + D,$$

where

$$\begin{aligned}D \equiv B\sigma_u \int_{-\infty}^{-\frac{C}{\sigma_u}} \left(v + \frac{C}{\sigma_u}\right) \phi(v)\, dv + A\sigma_u \int_{-\frac{C}{\sigma_u}}^{\infty} \left(v + \frac{C}{\sigma_u}\right)\phi(v)\, dv \\ = AC - (B-A)\sigma_u \int_{\frac{C}{\sigma_u}}^{\infty} \left(w - \frac{C}{\sigma_u}\right)\phi(w)\, dw,\end{aligned}$$

and similarly

$$E\left[f^2|p\right] = \int\limits_{-\infty}^{C} \left[p - A\left(u - C\right)\right]^2 \phi\left(u\right) du + \int\limits_{C}^{\infty} \left[p - B\left(u - C\right)\right]^2 \phi\left(u\right) du = p^2 + 2pD + E,$$

with

$$E \equiv A^2 \int\limits_{-\infty}^{C} \left(u - C\right)^2 \phi\left(u\right) du + B^2 \int\limits_{C}^{\infty} \left(u - C\right)^2 \phi\left(u\right) du$$

$$= A^2 \left[\sigma_u^2 + C^2\right] + \left(B^2 - A^2\right) \int\limits_{C}^{\infty} \left(u - C\right)^2 \phi\left(u\right) du.$$

Therefore

$$Var\left[f|p\right] = E\left[f^2|p\right] - E^2\left[f|p\right] = E - D^2,$$

that is independent of p. From here, the calculations end in the main text, and C solves

$$0 = L\left(C\right) \equiv C + \frac{1-\lambda}{\rho} \frac{D\left(C\right)}{E\left(C\right) - D^2\left(C\right) + \sigma_n^2} - S. \tag{1.27}$$

Consider the case when $w = 0$, that is $B_0 = A$; it implies that

$$D_0 = AC_0 \text{ and } E_0 = A^2\left[\sigma_u^2 + C_0^2\right],$$

therefore

$$Var\left[f|p_0\right] = E_0 - D_0^2 = A^2\sigma_u^2,$$

and hence

$$C_0 = \frac{\lambda + \rho A\sigma_u^2}{1 + \rho A\sigma_u^2} S = kS,$$

where $0 \leq k \leq 1$. As $S \geq 0$, we also get that $C_0 \geq 0$. In particular, if the asset is in positive net supply, $S > 0$, C_0 is positive, and $D_0 = AC_0 > 0$, which means that uninformed investors demand a discount of $D_0 > 0$ to hold the asset.

(ii) To show the existence of a real C that satisfies $L\left(C\right) = 0$, notice that when $B > A$, for $C = C_0$,

$$L(C_0) = C_0 + \frac{1-\lambda}{\rho} \frac{D(C_0)}{E(C_0) - D^2(C_0) + \sigma_n^2} - S$$
$$< -\frac{1-\lambda}{\rho} \frac{B-A}{E(C_0) - D^2(C_0) + \sigma_n^2} \int_{C_0}^{\infty} (u - C_0)\, \phi(u)\, du < 0.$$

Moreover, from (1.27) one can rewrite $L(C)$ as

$$L(C) = C + \frac{1-\lambda}{\rho} \frac{D(C)}{E(C) - D^2(C) + \sigma_n^2} - S$$
$$= \frac{\lambda Var[f|p] + \sigma_n^2}{\lambda Var[f|p] + \lambda \sigma_n^2} C - \frac{1-\lambda}{\rho} \frac{B-A}{Var[f|p] + \sigma_n^2} \int_{C}^{\infty} (u - C)\, \phi(u)\, du - S,$$

where S is constant, $Var[f|p]$ is finite (see below in the proof of Proposition 3), and for $C \to \infty$, $\int_{C}^{\infty} (u - C)\, \phi(u)\, du \to 0$. Therefore,

$$\lim_{C \to \infty} L(C) = \infty. \tag{1.28}$$

As $L(C)$ is continuos, combining it with $L(C_0) < 0$ and (1.28), it must have a real root above C_0.

(iii) For the proof of uniqueness, notice first that for every $C < C_0$, $D(C) < D(C_0)$ and $Var[f|p] > Var[f|p_0]$, therefore

$$L(C) < L(C_0) < 0.$$

Thus, there is no such $C < C_0$ that satisfies $L(C) = 0$. Regarding the case $C > C_0$, simple algebra shows that

$$\frac{dD}{dC} = A - (B - A) \frac{d}{dC} \int_{C}^{\infty} (u - C)\, \phi(u)\, du > A > 0,$$

and

$$E - D^2 = A^2 \left[\sigma_u^2 + C^2\right] + \left(B^2 - A^2\right) \int_C^{\infty} (u - C)^2 \phi(u)\, du$$
$$= A^2 \sigma_u^2 + 2(B - A) A \int_C^{\infty} u (u - C) \phi(u)\, du + (B - A)^2 Var\left[\max\{0, u - C\}\right],$$

hence

$$\frac{d}{dC}\left[E - D^2\right] < 0.$$

Therefore, $\frac{D(C)}{E(C) - (D(C))^2 + \sigma_n^2}$ is increasing in C, and

$$\frac{d}{dC} L(C) = 1 + \frac{d}{dC} \frac{D(C)}{\rho\left[E(C) - (D(C))^2\right] + A} > 1.$$

As $L(C)$ is strictly increasing and continuous, the $L(C) = 0$ equation must have a unique solution. ■

Proof of Proposition 3. From (1.10),

$$Var[f|p] = E - D^2$$
$$= A^2 \sigma_u^2 + \left(B^2 - A^2\right) \left[\int_C^{\infty} (u - C)^2 \phi(u)\, du - \left(\int_C^{\infty} (u - C) \phi(u)\, du\right)^2\right]$$
$$+ 2(B - A) AC \int_C^{\infty} (u - C) \phi(u)\, du + 2A(B - A) \left(\int_C^{\infty} (u - C) \phi(u)\, du\right)^2,$$

where the second term of the RHS is nonnegative due to Jensen's inequality applied on the random variable $w \equiv \max\{0, u - C\}$ and the convex function $x \mapsto x^2$: $E\left[w^2\right] \geq E^2[w]$, and the third and fourth components are trivially non-negative too. Therefore

$$Var[f|p] \geq A^2 \sigma_u^2 = Var[f|p_0]. \tag{1.29}$$

■

Proof of Proposition 4. Due to the improper prior assumption, $Var[p|f] = Var[f|p] =$

$E - D^2$ and $Var[p_0|f] = Var[f|p_0] = A^2\sigma_u^2$, hence (1.29) also implies that $Var[p|f] \geq Var[p_0|f]$. ■

Proof of Proposition 5. As $r_0 = f - p_0 = A(u - C_0)$, where u has a symmetric distribution around 0, which implies $Skew[r_0] = 0$. Therefore, to have $Skew[r] < Skew[r_0]$, it is sufficient to show $E\left[(r - E[r])^3\right] < 0$. From its definition,

$$r = f - p = -\begin{cases} A(u - C) & \text{if } u \leq C \\ B(u - C) & \text{if } u > C \end{cases},$$

and

$$E\left[(r - E[r])^3\right] = E\left[r^3\right] - 3E\left[r^2\right]E[r] + 2E^3[r]$$

where

$$\begin{aligned} E[r] &= D = AC - (B - A)\int_C^\infty (u - C)\,\phi(u)\,du, \\ E\left[r^2\right] &= E = A^2\left(\sigma_u^2 + C^2\right) + \left(B^2 - A^2\right)\int_C^\infty (u - C)^2\,\phi(u)\,du, \text{ and} \\ E\left[r^3\right] &= A^3C\left(3\sigma_u^2 + C^2\right) - \left(B^3 - A^3\right)\int_C^\infty (u - C)^3\,\phi(u)\,du. \end{aligned}$$

After some tedious algebra, the negativity of the skewness follows from the fact that the skewness of the random variable $w = \max\{0, u - C\}$ is positive.

To prove the second part of the proposition, notice that

$$E[|r_0|] = E[A|u|] = 2A\int_0^\infty \max\{u, C_0\}\,\phi(u)\,du$$

and

$$
\begin{aligned}
E\left[|r|\right] &= B\int\limits_{C}^{\infty}(u-C)\,\phi(u)\,du - A\int\limits_{-\infty}^{C}(u-C)\,\phi(u)\,du \\
&= (B+A)\int\limits_{C}^{\infty}u\phi(u)\,du - (B-A)\,C\int\limits_{C}^{\infty}\phi(u)\,du + 2AC\int\limits_{0}^{C}\phi(u)\,du,
\end{aligned}
$$

thus simple algebra yields

$$
E\left[|r|\right] - E\left[|r_0|\right] = (B-A)\int\limits_{C}^{\infty}(u-C)\,\phi(u)\,du + 2A\int\limits_{0}^{\infty}\left(\max\{u,C\} - \max\{u,C_0\}\right)\phi(u)\,du.
$$

On the right-hand side both components are non-negative, therefore $E\left[|r|\right] \geq E\left[|r_0|\right]$. ■

Proof of Proposition 6. As $R = p - p^{-1} = f - r - p^{-1}$, $Skew\left[R|f\right] = Skew\left[-r\right]$ and $Skew\left[R_0|f\right] = Skew\left[-r_0\right]$, hence $Skew\left[R|f\right] > Skew\left[R_0|f\right]$ is straightforward from $Skew\left[r\right] < Skew\left[r_0\right]$. ■

1.9.2 Information content of the price under short-sale constraints

The posterior of a creditor with private signal t and price signal p that comes from a financial market with short-sale constraints is given by

$$
g\left(f|t, P=p\right) = \pi \cdot g\left(f|t, P=p, f<p\right) + (1-\pi)\cdot g\left(f|t, P=p, f\geq p\right),
$$

where, using the simplifying notation

$$
\Phi_1 = \Phi\left(\frac{p - \frac{\tau_t t + \tau_{Bu}(p+BC)}{\tau_t+\tau_{Bu}}}{\left(\tau_{Bu}+\tau_t\right)^{-1/2}}\right) \text{ and } \Phi_2 = 1 - \Phi\left(\frac{p - \frac{\tau_t t + \tau_{Au}(p+AC)}{\tau_t+\tau_{Au}}}{\left(\tau_{Au}+\tau_t\right)^{-1/2}}\right),
$$

the variable

$$
\begin{aligned}
\pi &= \Pr\left(f<p|t, P=p\right) \\
&= \frac{\frac{\left(\tau_{Bu}+\tau_t\right)^{-1/2}}{\tau_t^{-1}\tau_{Bu}^{-1}}\phi\left(\frac{t-(p+BC)}{\left(\tau_{Bu}^{-1}+\tau_t^{-1}\right)^{1/2}}\right)\Phi_1}{\frac{\left(\tau_{Bu}+\tau_t\right)^{-1/2}}{\tau_t^{-1}\tau_{Bu}^{-1}}\phi\left(\frac{t-(p+BC)}{\left(\tau_{Bu}^{-1}+\tau_t^{-1}\right)^{1/2}}\right)\Phi_1 + \frac{\left(\tau_{Au}+\tau_t\right)^{-1/2}}{\tau_t^{-1}\tau_{Au}^{-1}}\phi\left(\frac{t-(p+AC)}{\left(\tau_{Au}^{-1}+\tau_t^{-1}\right)^{1/2}}\right)\Phi_2}
\end{aligned}
$$

gives the probability that the creditor assigns to the constraint binding in the financial market, $\Pr(f \geq p|t, P = p) = 1 - \pi$, and the conditional distributions are given by

$$g(f|t, P = p, f < p) = 1_{f<p} \frac{1}{(\tau_{Bu} + \tau_t)^{-1/2} \Phi_1} \phi \left(\frac{f - \frac{\tau_t t + \tau_{Bu}(p+BC)}{\tau_t + \tau_{Bu}}}{(\tau_{Bu} + \tau_t)^{-1/2}} \right)$$

and

$$g(f|t, P = p, f \geq p) = 1_{f\geq p} \frac{1}{(\tau_{Au} + \tau_t)^{-1/2} \Phi_2} \phi \left(\frac{f - \frac{\tau_t t + \tau_{Au}(p+AC)}{\tau_t + \tau_{Au}}}{(\tau_{Au} + \tau_t)^{-1/2}} \right).$$

When no informed trader is subject to the short-sale constraint, i.e. $w = 0$ or $B = A$, both truncated normal pdfs belong to the same normal distribution, with mean $\frac{\tau_t}{\tau_t + \tau_{Au}} t + \frac{\tau_{Au}}{\tau_t + \tau_{Au}} (p + AC_0)$ and precision $\tau_t + \tau_{Au}$, and the probabilities simplify to

$$\begin{aligned} \Pr(f < p|t, P_0 = p) &= \Phi \left(\frac{p - \frac{\tau_t t + \tau_{Bu}(p+AC_0)}{\tau_t + \tau_{Au}}}{(\tau_{Bu} + \tau_t)^{-1/2}} \right) \text{ and} \\ \Pr(f \geq p|t, P_0 = p) &= 1 - \Phi \left(\frac{p - \frac{\tau_t t + \tau_{Bu}(p+AC_0)}{\tau_t + \tau_{Au}}}{(\tau_{Bu} + \tau_t)^{-1/2}} \right). \end{aligned}$$

1.9.3 Global game solution under short-sale constraints

General notations

As shown in Section 1.4, in presence of short-sale constraints the posterior pdf becomes

$$g(\theta|t, P = p) = \pi \cdot g(\theta|t, P = p, \theta < p) + (1 - \pi) \cdot g(\theta|t, P = p, \theta \geq p),$$

hence a simple integration yields that

$$G(x|t, P = p) = \pi \cdot G(x|t, P = p, \theta < p) \text{ if } x \leq p,$$

and

$$G(x|t, P = p) = \pi \cdot G(p|t, P = p, \theta < p) + (1 - \pi) \cdot G(x|t, P = p, \theta \geq p) \text{ if } x > p.$$

Combining it with the indifference condition $\Pr(\theta \geq \theta^* | t^*, P = p) = c$ gives

$$1 - c = \pi \cdot G(\theta^* | t^*, P = p, \theta < p) \text{ if } \theta^* \leq p,$$

and

$$1 - c = \pi \cdot G(p | t^*, P = p, \theta < p) + (1 - \pi) \cdot G(\theta^* | t^*, P = p, \theta \geq p) \text{ if } \theta^* > p,$$

which leads to (1.24) with the notation

$$\pi_A^* = \Phi\left(\frac{p - \frac{\tau_t t^* + \tau_{Au}(p+AC)}{\tau_t + \tau_{Au}}}{(\tau_t + \tau_{Au})^{-1/2}}\right) \text{ and } \pi_B^* = \Phi\left(\frac{p - \frac{\tau_t t^* + \tau_{Bu}(p+BC)}{\tau_{Bu} + \tau_t}}{(\tau_t + \tau_{Bu})^{-1/2}}\right)$$

and

$$\begin{aligned} \pi^* &\equiv \Pr(\theta < p | t^*, p) \\ &= \frac{\frac{(\tau_{Bu}+\tau_t)^{-1/2}}{\tau_t^{-1}\tau_{Bu}^{-1}} \phi\left(\frac{t^* - (p+BC)}{\left(\tau_{Bu}^{-1} + \tau_t^{-1}\right)^{1/2}}\right) \pi_B^*}{\frac{(\tau_{Bu}+\tau_t)^{-1/2}}{\tau_t^{-1}\tau_{Bu}^{-1}} \phi\left(\frac{t^* - (p+BC)}{\left(\tau_{Bu}^{-1} + \tau_t^{-1}\right)^{1/2}}\right) \pi_B^* + \frac{(\tau_{Au}+\tau_t)^{-1/2}}{\tau_t^{-1}\tau_{Au}^{-1}} \phi\left(\frac{t^* - (p+AC)}{\left(\tau_{Au}^{-1} + \tau_t^{-1}\right)^{1/2}}\right) \left[1 - \pi_A^*\right]}. \end{aligned}$$

Multiplicity

First, I characterize the critical mass and individual optimality curves, (1.22) and (1.24), respectively. It is easy to see that both of them imply θ^* is a continuous and strictly increasing function of t^*.

Condition (1.22) yields that the slope of the critical mass curve (CM, for simplicity) is given by

$$\frac{dt^*}{d\theta^*} = 1 + \frac{1}{\sqrt{\tau_t}\omega} \frac{1}{\phi\left(\Phi^{-1}\left(\frac{\theta^*}{\omega}\right)\right)},$$

or by its inverse

$$\delta_{CM} \equiv \frac{d\theta^*}{dt^*} = \frac{1}{1 + \frac{1}{\sqrt{\tau_t}\omega} \frac{1}{\phi\left(\Phi^{-1}\left(\frac{\theta^*}{\omega}\right)\right)}}$$

for the θ^* that solves

$$t^* = \theta^* + \frac{1}{\sqrt{\tau_t}} \Phi^{-1}\left(\frac{\theta^*}{\omega}\right).$$

Thus, δ_{CM} can be interpreted as a function of t^*. In particular, as $0 < \phi(x) < 1/\sqrt{2\pi}$ for every $x \in R$, we have a lower and an upper threshold for the slope:

$$0 < \delta_{CM} < \overline{\delta}_{CM} \equiv \frac{1}{1 + \frac{1}{\sqrt{\tau_t}\omega}\sqrt{2\pi}},$$

and it reaches its maximum for

$$\theta^* = \frac{\omega}{2},$$

or for the t^* value of

$$t^* = \frac{\omega}{2}.$$

Moreover, as $\phi(x)$ is strictly increasing for $x < 0$ and strictly decreasing for $x > 0$, and θ^* is an increasing function of t^*, the slope δ_{CM} is strictly increasing in t^* for $t^* < \frac{\omega}{2}$ and strictly decreasing for $t^* > \frac{\omega}{2}$. Therefore, δ_{CM} takes every value in $\left(0, \overline{\delta}_{CM}\right]$ for $t^* \leq \frac{\omega}{2}$, and every value between $\overline{\delta}_{CM}$ and 0, when t^* increases from $\frac{\omega}{2}$ to ∞, where $\overline{\delta}_{CM} < 1$.

Now I turn my attention to the individual optimality curve (IO, for simplicity), given in (1.24), which, for tractability, is restated here:

$$\theta^* = \begin{cases} \frac{\tau_t}{\tau_t+\tau_{Bu}} t^* + \frac{\tau_{Bu}}{\tau_t+\tau_{Bu}} (p + BC) + \frac{1}{\sqrt{\tau_t+\tau_{Bu}}} \Phi^{-1}\left((1-c)\frac{\pi_B^*(p,t^*)}{\pi^*(p,t^*)}\right) & \text{if } \theta^* \leq p \\ \frac{\tau_t}{\tau_t+\tau_{Au}} t^* + \frac{\tau_{Au}}{\tau_t+\tau_{Au}} (p + AC) + \frac{1}{\sqrt{\tau_t+\tau_{Au}}} \Phi^{-1}\left(1 - c\frac{1-\pi_A^*(p,t^*)}{1-\pi^*(p,t^*)}\right) & \text{if } \theta^* > p. \end{cases}$$

The first observation I make is that this condition can be rewritten with the introduction of $\Delta\theta \equiv \theta^* - p$ and $\Delta t \equiv t^* - p$:

$$\Delta\theta = \begin{cases} \frac{\tau_t}{\tau_t+\tau_{Bu}} \Delta t + \frac{\tau_{Bu}}{\tau_t+\tau_{Bu}} BC + \frac{1}{\sqrt{\tau_t+\tau_{Bu}}} \Phi^{-1}\left((1-c)\frac{\pi_B^*(\Delta t)}{\pi^*(\Delta t)}\right) & \text{if } \Delta\theta \leq 0 \\ \frac{\tau_t}{\tau_t+\tau_{Au}} \Delta t + \frac{\tau_{Au}}{\tau_t+\tau_{Au}} AC + \frac{1}{\sqrt{\tau_t+\tau_{Au}}} \Phi^{-1}\left(1 - c\frac{1-\pi_A^*(\Delta t)}{1-\pi^*(\Delta t)}\right) & \text{if } \Delta\theta > 0, \end{cases} \tag{1.30}$$

where

$$\pi_A^*(\Delta t) = \Phi\left(-\frac{\tau_t \Delta t + \tau_{Au} AC}{(\tau_t + \tau_{Au})^{1/2}}\right) \text{ and } \pi_B^*(\Delta t) = \Phi\left(-\frac{\tau_t \Delta t + \tau_{Bu} BC}{(\tau_t + \tau_{Bu})^{1/2}}\right), \tag{1.31}$$

and

$$\pi^*(\Delta t) =$$

$$= \frac{\frac{(\tau_{Bu}+\tau_t)^{-1/2}}{\tau_t^{-1}\tau_{Bu}^{-1}}\phi\left(\frac{\Delta t - BC}{\left(\tau_{Bu}^{-1}+\tau_t^{-1}\right)^{1/2}}\right)\pi_B^*(\Delta t)}{\frac{(\tau_{Bu}+\tau_t)^{-1/2}}{\tau_t^{-1}\tau_{Bu}^{-1}}\phi\left(\frac{\Delta t - BC}{\left(\tau_{Bu}^{-1}+\tau_t^{-1}\right)^{1/2}}\right)\pi_B^*(\Delta t) + \frac{(\tau_{Au}+\tau_t)^{-1/2}}{\tau_t^{-1}\tau_{Au}^{-1}}\phi\left(\frac{\Delta t - AC}{\left(\tau_{Au}^{-1}+\tau_t^{-1}\right)^{1/2}}\right)\left[1-\pi_A^*(\Delta t)\right]}. \tag{1.32}$$

It means that on the (t^*, θ^*) plane every solution-pair $(t^*(p), \theta^*(p))$ of (1.24) is given by an appropriate shift of the point $(\Delta t, \Delta\theta)$ along the 45-degree line, where Δt and $\Delta\theta$ solve (1.30) – (1.32), and are only functions of the parameters of the model and the equilibrium constants of the financial market, and do not depend on p. It also implies that the slope and convexity attributes of the curve do not depend on p either.

The characterization of the solution to (1.30) – (1.32) is as follows. The slope of the curve is given by

$$\delta_{IO}(\Delta t) = \begin{cases} \frac{\tau_t}{\tau_t+\tau_{Bu}} + \frac{1}{\sqrt{\tau_t+\tau_{Bu}}}\frac{d}{d(\Delta t)}\Phi^{-1}\left((1-c)\frac{\pi_B^*(\Delta t)}{\pi^*(\Delta t)}\right) & \text{if } \Delta\theta \le 0 \\ \frac{\tau_t}{\tau_t+\tau_{Au}} + \frac{1}{\sqrt{\tau_t+\tau_{Au}}}\frac{d}{d(\Delta t)}\Phi^{-1}\left(1-c\frac{1-\pi_A^*(\Delta t)}{1-\pi^*(\Delta t)}\right) & \text{if } \Delta\theta > 0, \end{cases}$$

$$= \begin{cases} \frac{\tau_t}{\tau_t+\tau_{Bu}} + \frac{(1-c)}{\sqrt{\tau_t+\tau_{Bu}}\phi\left(\Phi^{-1}\left((1-c)\frac{\pi_B^*(\Delta t)}{\pi^*(\Delta t)}\right)\right)}\frac{d}{d(\Delta t)}\left(\frac{\pi_B^*(\Delta t)}{\pi^*(\Delta t)}\right) & \text{if } \Delta t \le \Delta t_0 \\ \frac{\tau_t}{\tau_t+\tau_{Au}} - \frac{c}{\sqrt{\tau_t+\tau_{Au}}\phi\left(\Phi^{-1}\left(1-c\frac{1-\pi_A^*(\Delta t)}{1-\pi^*(\Delta t)}\right)\right)}\frac{d}{d(\Delta t)}\left(\frac{1-\pi_A^*(\Delta t)}{1-\pi^*(\Delta t)}\right) & \text{if } \Delta t > \Delta t_0. \end{cases}$$

where Δt_0 is the unique solution to

$$\pi^*(\Delta t_0) = 1-c,$$

i.e. Δt_0 is the Δt value that gives $\Delta\theta = 0$ in (1.31). Moreover, after some demanding calculations, omitted here, it is possible to show that:

1. $\frac{d}{d(\Delta t)}\left(\frac{\pi_B^*(\Delta t)}{\pi^*(\Delta t)}\right) > 0$ for $\Delta t \le \Delta t_0$ and $\frac{d}{d(\Delta t)}\left(\frac{1-\pi_A^*(\Delta t)}{1-\pi^*(\Delta t)}\right) < 0$ for $\Delta t > \Delta t_0$;

2. δ_{IO} is increasing in Δt for both $(-\infty, \Delta t_0)$ and $(\Delta t_0, \infty)$;

3. In the limits $\Delta t \to \pm\infty$ we have $\lim_{\Delta t\to-\infty}\delta_{IO}(\Delta t) = \frac{\tau_t}{\tau_t+\tau_{Bu}}$ and $\lim_{\Delta t\to\infty}\delta_{IO}(\Delta t) =$

$\frac{\tau_t}{\tau_t+\tau_{Au}}$

4. There is a 'kink' at Δt_0, and thus the IO curve is not differentiable: on the two sides of $\Delta t = \Delta t_0$, the slopes are finite but different. In particular, I define $\alpha(\Delta t_0)$ and $\beta(\Delta t_0)$ such that $\frac{\tau_t}{\tau_t+\tau_{Bu}} < \frac{1}{\beta(\Delta t_0)} \equiv \lim_{\Delta t \to \Delta t_0 -} \delta_{IO}(\Delta t) < \infty$ and $0 < \alpha(\Delta t_0) \equiv \lim_{\Delta t \to \Delta t_0 +} \delta_{IO}(\Delta t) < \frac{\tau_t}{\tau_t+\tau_{Au}}$, are well-defined, and satisfy $\alpha(\Delta t_0) < \frac{1}{\beta(\Delta t_0)}$;

5. The slope $\delta_{IO}(\Delta t)$ is increasing in Δt and takes every value in $\left(\frac{\tau_t}{\tau_t+\tau_{Bu}}, \frac{1}{\beta(\Delta t_0)}\right)$ for $\Delta t < \Delta t_0$, and is increasing and takes every value in $\left(\alpha(\Delta t_0), \frac{\tau_t}{\tau_t+\tau_{Au}}\right)$ for $\Delta t > \Delta t_0$.

In what follows, for simplicity, I refer to this $(\Delta t_0, 0)$ point as the 'kink' of the IO curve.

The next observation is that the uniqueness or multiplicity of the solutions for a given parameter set depends only on which part of the CM curve the 'kink' of the IO condition would get shifted to. In particular, if the 'kink' is shifted to a part of the CM curve where its slope is sufficiently small such that $\delta_{CM} \leq \alpha(\Delta t_0)$, there is a unique solution. This is because both before and after the kink the IO curve is steeper than the CM curve, and hence there cannot be any more intersections. However, if at this point the slope satisfies $\delta_{CM} > \alpha(\Delta t_0)$, multiplicity can be ensured by choosing the appropriate p: if the kink is before $t^* = \frac{\omega}{2}$, as δ_{IO} starts to increase from $\alpha(\Delta t_0)$, and as δ_{CM} decreases, a slight increase in p would ensure that they have multiple equilibria, and if the kink is after $t^* = \frac{\omega}{2}$, as δ_{IO} increases from $\alpha(\Delta t_0)$, and as δ_{CM} increases too, a slight decrease in p would ensure that they have multiple equilibria. Therefore, in what follows, I solve for the point where the 'kink' gets shifted to, and determine the relationship of the two slopes δ_{CM} and δ_{IO}.

First, suppose that $\alpha(\Delta t_0) \geq \overline{\delta}_{CM}$; in this case there is a unique solution. This condition hence requires

$$\alpha(\Delta t_0) \geq 1 - \frac{\sqrt{2\pi}}{\sqrt{2\pi} + \sqrt{\tau_t}\omega}.$$

Second, suppose that $0 < \alpha(\Delta t_0) < \overline{\delta}_{CM}$. It means that there are two points of the CM curve such that the slope is exactly $\alpha(\Delta t_0)$: they are pinned down by

$$\begin{aligned}
\underline{\theta}^* &= \omega\left(1 - \Phi\left[\phi^{-1}\left(\frac{1}{\omega\sqrt{\tau_t}}\frac{\alpha(\Delta t_0)}{1-\alpha(\Delta t_0)}\right)\right]\right) < \frac{\omega}{2} \text{ and} \\
\overline{\theta}^* &= \omega\Phi\left[\phi^{-1}\left(\frac{1}{\omega\sqrt{\tau_t}}\frac{\alpha(\Delta t_0)}{1-\alpha(\Delta t_0)}\right)\right] > \frac{\omega}{2},
\end{aligned}$$

where $\phi^{-1}(y)$ denotes the unique non-negative x such that $\phi(x) = y$ for $y \leq 1/\sqrt{2\pi}$. Therefore, if the shifted kink has $\theta^* \leq \underline{\theta}^*$ or $\theta^* \geq \overline{\theta}^*$, the solution is unique, and if $\underline{\theta}^* < \theta^* < \overline{\theta}^*$, there are multiple equilibria.

As the 'kink' has coordinates $(\Delta t_0, 0)$ on the (t^*, θ^*) plane, shifting it is equivalent to moving it to the point $(p + \Delta t_0, p)$. It is thus on the CM curve if and only if

$$p + \Delta t_0 = p + \frac{1}{\sqrt{\tau_t}} \Phi^{-1}\left(\frac{p}{\omega}\right),$$

that is if it is shifted by $\widehat{p} = \omega\Phi\left(\sqrt{\tau_t}\Delta t_0\right)$, to the point $\left(\omega\Phi\left(\sqrt{\tau_t}\Delta t_0\right) + \Delta t_0, \omega\Phi\left(\sqrt{\tau_t}\Delta t_0\right)\right)$. Therefore, there are multiple equilibria if and only if

$$\underline{\theta}^* < \omega\Phi\left(\sqrt{\tau_t}\Delta t_0\right) < \overline{\theta}^*,$$

that is

$$\frac{1}{\omega\sqrt{\tau_t}} \frac{\alpha(\Delta t_0)}{1 - \alpha(\Delta t_0)} < \phi\left(\sqrt{\tau_t}\Delta t_0\right) < \frac{1}{\sqrt{2\pi}},$$

or, equivalently,

$$\alpha(\Delta t_0) < 1 - \frac{1}{1 + \omega\sqrt{\tau_t}\phi\left(\sqrt{\tau_t}\Delta t_0\right)}.$$

But what is exactly $\alpha(\Delta t_0) \equiv \lim_{\Delta t \to \Delta t_0 +} \delta_{IO}(\Delta t)$? Using the relevant part of the δ_{IO} function, and the fact that $\pi^*(\Delta t_0) = 1 - c$, the (upper) slope δ_{IO} close to the 'kink' becomes

$$\begin{aligned}
\alpha(\Delta t_0) &= \lim_{\Delta t \to \Delta t_0 +} \delta_{IO}(\Delta t) \\
&= (1-c)\frac{\tau_t}{\tau_t + \tau_{Au}} + c\frac{\tau_t}{\tau_t + \tau_{Bu}}\frac{\tau_{Bu}}{\tau_{Au}} \\
&\quad -c\frac{\tau_{Bu}}{\tau_{Au}}\frac{1}{\sqrt{\tau_t + \tau_{Bu}}} \frac{\phi\left(\frac{\Delta t_0 - BC}{\left(\tau_t^{-1} + \tau_{Bu}^{-1}\right)^{1/2}}\right)}{\phi\left(\frac{\Delta t_0 - AC}{\left(\tau_t^{-1} + \tau_{Au}^{-1}\right)^{1/2}}\right)} \frac{\left(\frac{\Delta t_0 - AC}{\tau_t^{-1} + \tau_{Au}^{-1}} - \frac{\Delta t_0 - BC}{\tau_t^{-1} + \tau_{Bu}^{-1}}\right)\Phi\left(-\frac{\tau_t \Delta t_0 + \tau_{Bu} BC}{\sqrt{\tau_t + \tau_{Bu}}}\right)}{\phi\left(-\frac{\tau_t \Delta t_0 + \tau_{Au} AC}{\sqrt{\tau_t + \tau_{Au}}}\right)}.
\end{aligned}$$

Because of the elaborate expression above, it is impossible to characterize the number of equilibria as a function of the precisions τ_t and τ_u in the general case. Instead, I only consider the special case, when τ_u is held constant and $\tau_t \to \infty$, because this is the case where Morris and Shin provide uniqueness. In fact, it is easy to show that $\pi^*(\Delta t_0) = 1 - c$

implies $\Delta t_0 = 0$,and hence

$$\alpha \equiv \lim_{\tau_t \to \infty} \alpha (\Delta t_0) = (1-c) + c\frac{\tau_{Bu}}{\tau_{Au}} = \frac{(1-c) B^2 + cA^2}{B^2} < 1.$$

For multiplicity it must be that

$$\alpha (\Delta t_0) < 1 - \frac{1}{1 + \omega \sqrt{\tau_t} \phi \left(\sqrt{\tau_t} \Delta t_0 \right)},$$

but when $\tau_t \to \infty$, the RHS converges to 1, and in the limit indeed

$$\alpha = \frac{(1-c) B^2 + cA^2}{B^2} < 1.$$

Therefore, when the private signal becomes arbitrarily precise, there are still multiple equilibria of the coordination problem. Similarly one can show that in the limit $\tau_t \to \infty$ the 'lower' slope of the IM curve becomes

$$\frac{1}{\beta} \equiv \lim_{\tau_t \to \infty} \frac{1}{\beta (\Delta t_0)} = (1-c) + c\frac{\tau_{Au}}{\tau_{Bu}} = \frac{(1-c) B^2 + cA^2}{A^2} > 1.$$

The next question is what the exact thresholds are in these multiple equilibria. First, one derived intersection is at the kink, which provides an equilibrium of the model. By its definition, the kink satisfies $\Delta \theta_0 = 0$, which in the limit $\tau_t \to \infty$ implies $\Delta t_0 = 0$, and thus the shifting by p gives the solution $t^* = \theta^* = p$.

For a second intersection to be derived, one needs to find the joint solution of equations (1.22) and (1.24) in the limit when $\tau_t \to \infty$. Instead of solving for the explicit joint solutions, I instead guess and verify that the solution is $t^* = \theta^* = c\omega$, and only when $p < c\omega$. Indeed, assuming that $\theta^* = c\omega$, the CM equation gives

$$t^* = \theta^* + \frac{1}{\sqrt{\tau_t}} \Phi^{-1} \left(\frac{\theta^*}{\omega} \right) = c\omega + \frac{1}{\sqrt{\tau_t}} \Phi^{-1} (c),$$

and hence when $\tau_t \to \infty$, the RHS converges to $c\omega$, therefore we have $\lim_{\tau_t \to \infty} t^* = c\omega$.

Suppose now that $t^* = c\omega$, and plug it in the IM equation. First, if $\theta^* > p$,

$$\lim_{\tau_t \to \infty} \frac{1 - \pi_A^* (p, t^* = c\omega)}{1 - \pi^* (p, t^* = c\omega)} = 1 + \left[\frac{A^2}{B^2} \frac{\phi \left(- \frac{(p - c\omega) + BC}{\tau_{Bu}^{-1/2}} \right)}{\phi \left(- \frac{(p - c\omega) + AC}{\tau_{Au}^{-1/2}} \right)} - 1 \right] \lim_{\tau_t \to \infty} \Phi \left(\sqrt{\tau_t} (p - c\omega) \right),$$

where $\lim_{\tau_t \to \infty} \Phi \left(\sqrt{\tau_t} (p - c\omega) \right)$ is bounded and hence always finite. Therefore, in the limit it satisfies

$$\lim_{\tau_t \to \infty} \frac{1}{\sqrt{\tau_t + \tau_{Au}}} \Phi^{-1} (.) = 0,$$

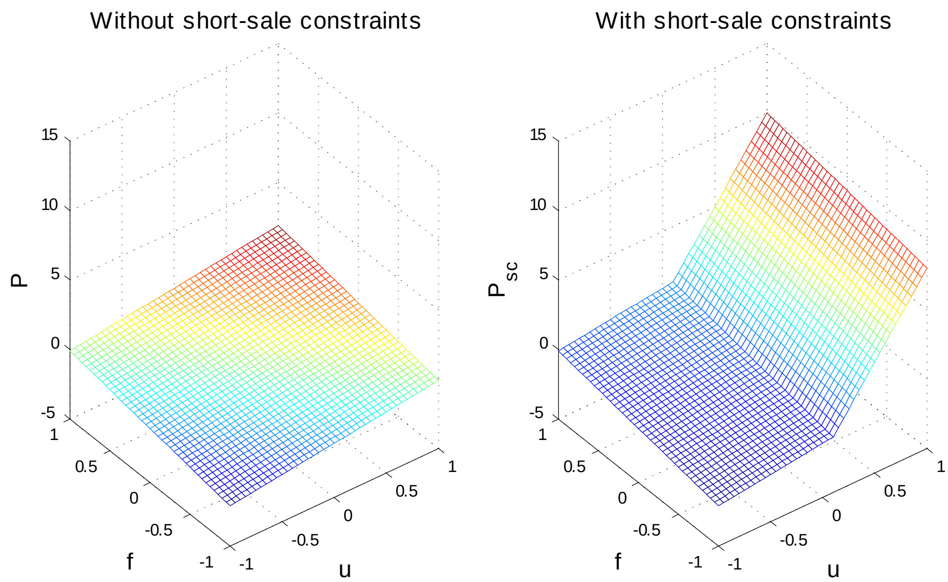
and hence $\lim_{\tau_t \to \infty} \theta^* = c\omega$. Therefore, in the limit the other intersection satisfies $t^* = \theta^* = c\omega$.

Second, suppose that $\theta^* \leq p$. When $t^* = c\omega$, in the limit $\tau_t \to \infty$ again

$$\begin{aligned} \lim_{\tau_t \to \infty} \frac{\pi_B^* (p, t^* = c\omega)}{\pi^* (p, t^* = c\omega)} = & \lim_{\tau_t \to \infty} \Phi \left(\sqrt{\tau_t} (p - c\omega) \right) \\ & + \frac{B^2}{A^2} \frac{\phi \left(- \frac{(p - c\omega) + AC}{\tau_{Au}^{-1/2}} \right)}{\phi \left(- \frac{(p - c\omega) + BC}{\tau_{Bu}^{-1/2}} \right)} \left(1 - \lim_{\tau_t \to \infty} \Phi \left(\sqrt{\tau_t} (p - c\omega) \right) \right), \end{aligned}$$

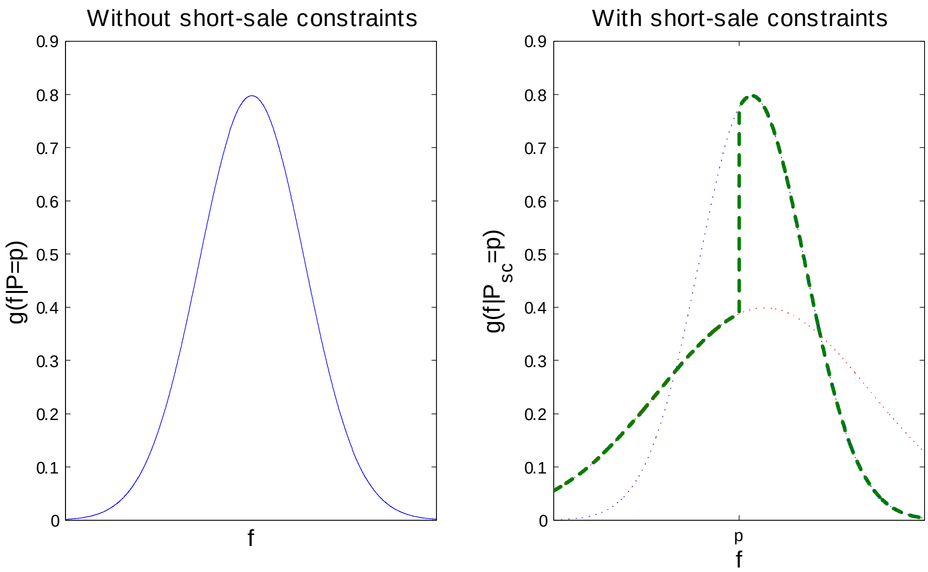
which is always finite, and hence in the limit (1.24) simplifies to $\lim_{\tau_t \to \infty} \theta^* = c\omega$. Therefore, in the limit the other intersection satisfies $t^* = \theta^* = c\omega$.

Figure 1-1: The equilibrium price as a function of the fundamental and demand shock, in absence and presence of short-sale constraints.



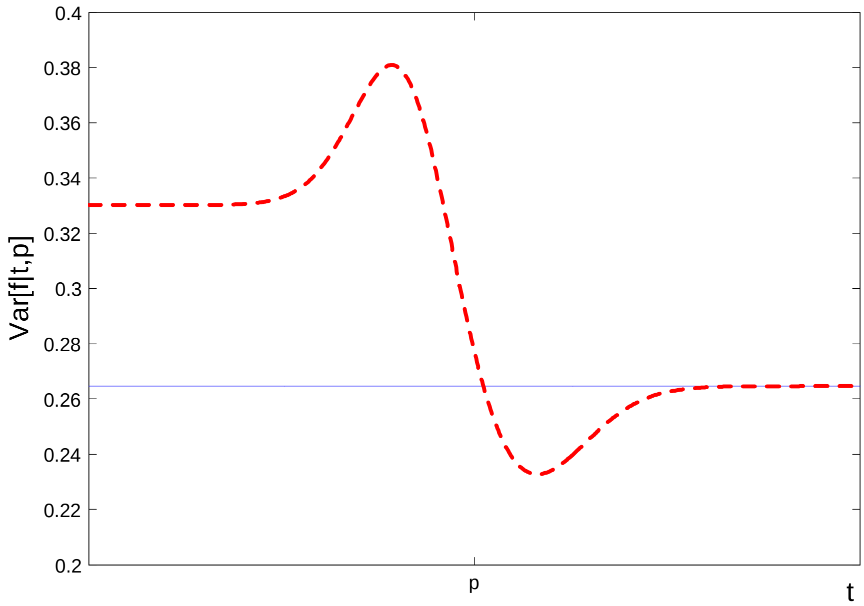
The demand shock (u) is on the x axis and the payoff (f) is on the y axis. The left panel shows the price when informed investors are not short-sale constrained, and the right panel shows the price when $w > 0$ proportion of informed investors are subject to short-sale constraints. The parameters are set to $S = 0$, $\sigma_n^2 = 1$, $\lambda = 0.5$ and $w = 0.9$, which imply $A = 1$ and $B = 10$. The equilibrium value of C depends on the assumption about the demand shock distribution g_u; without making any distributional assumptions I set $C = 0.1$.

Figure 1-2: Distribution of f conditional on p, in absence and presence of short-sale constraints.



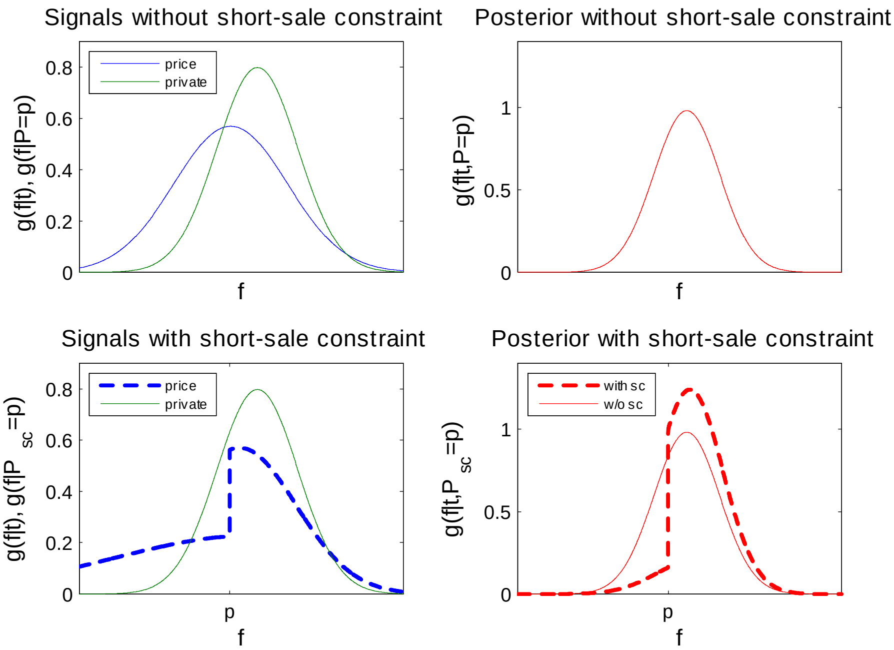
The left panel shows that the distribution without short-sale constraints is Gaussian. The right panel shows that in presence of short-sale constraints (dashed line), the distribution is only locally normal, with different means and variances on the two segment, and with a jump at the price p. For comparison, the continuous line represents the conditional distribution without the constraint.

Figure 1-3: Variance of f, conditional on the private signal t and the price p, without and with short-sale constraints.



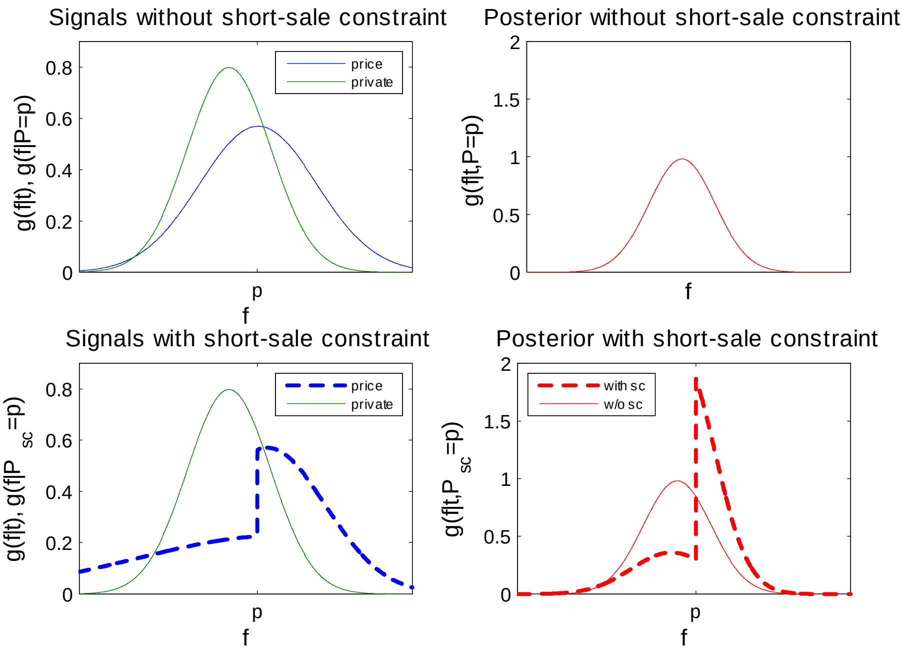
The solid line shows the variance of f conditional on the private signal t and the price p with no short-sale constraints. The dashed line shows the same in presence of short-sale constraints. When $t \to \infty$, short-sale constraints do not change price informativeness. When $t \to -\infty$, short-sale constraints, intuitively, decrease price informativeness. For intermediate t values, creditors can actually learn more from prices under short-sale constraints. The parameters are set to $\sigma_t = 0.2$, $\sigma_u = 0.5$, $\sigma_n = 1$, $\rho = 1$, $\lambda = 0.5$ and $w = 0.9$, implying $A = 2$ and $B = 20$.

Figure 1-4: Conditional distribution of f based on t, p, and both t and p, in absence and presence of short-sale constraints, when $t > p$.



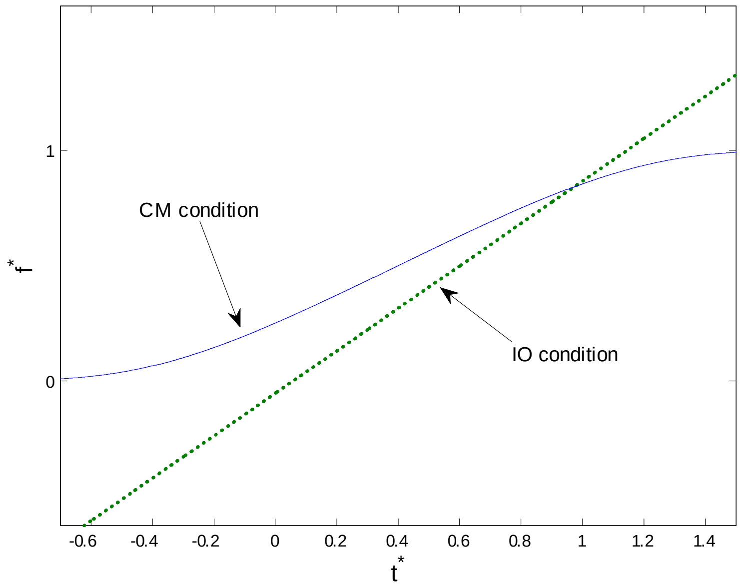
The top left panel shows $g(f|t)$ and $g(f|p)$, and the top right panel shows $g(f|t,p)$, without short-sale constraints. The bottom left panel shows $g(f|t)$ and $g(f|p)$, and the bottom right panel shows $g(f|t,p)$, with short-sale constraints. The parameters are set to $\sigma_t = 0.5$, $\sigma_u = 0.7$, $\sigma_n = 1$, $\rho = 1$, $\lambda = 0.5$, and $w = 0.6$. Signal realizations are $t = -0.2$ and $p = 0.14$.

Figure 1-5: Conditional distribution of f based on t, p, and both t and p, in absence and presence of short-sale constraints, when $t < p$.



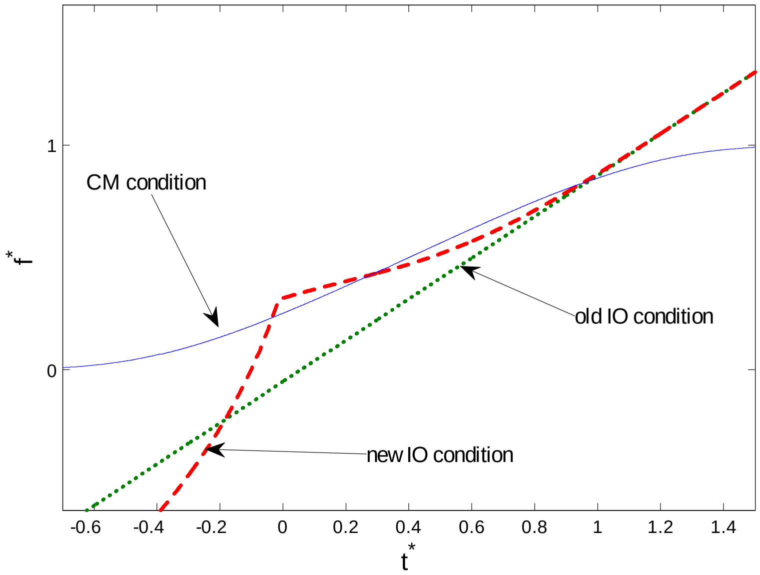
The top left panel shows $g(f|t)$ and $g(f|p)$, and the top right panel shows $g(f|t,p)$, without short-sale constraints. The bottom left panel shows $g(f|t)$ and $g(f|p)$, and the bottom right panel shows $g(f|t,p)$, with short-sale constraints. Signal realizations are $t = 0.2$ and $p = -0.14$.

Figure 1-6: IO and CM conditions without short-sale constraints.



This figure plots the critical mass condition, (1.22), thin line, and the individual optimality condition without short-sale constraints, (1.23), dotted line. The cutoff strategy for investment, t^*, is shown on the x axis, and the success threshold, θ^*, is on the y axis. The parameters and variable realizations used here are $\sigma_n = 1$, $\sigma_u = 0.5$, $\sigma_t = 0.3$, $\omega = 0.8$, $c = 0.7$ and $p = 0.25$.

Figure 1-7: IO and CM conditions with short-sale constraints.



This figure plots the critical mass condition, (1.22), thin line, and the individual optimality condition with short-sale constraints, (1.24), dashed line. The cutoff strategy for investment, t^*, is shown on the x axis, and the success threshold, θ^*, is on the y axis. The parameters and variable realizations used here are $\sigma_n = 1$, $\sigma_u = 0.5$, $\sigma_t = 0.3$, $\omega = 0.8$, $c = 0.7$ and $p = 0.25$. For comparison, the dotted line shows the individual optimality condition without short-sale constraints, (1.23).

Chapter 2

Financially constrained strategic arbitrage

2.1 Introduction

Large traders, such as dealers, hedge funds and other financial institutions play an important role in financial markets when exploiting the relative mispricing of assets: through their trading, these arbitrageurs bring prices closer to fundamentals and provide liquidity to other market participants. However, their willingness to provide liquidity can be subject to many factors. For example, institutional investors' trades can have significant price impact as their strategies often involve dealing with large positions in assets held by a relatively few number of investors. Also, wealth constraints and risk management policies crucially affect arbitrageurs' allocation of capital to trading opportunities. Therefore, when a small number of arbitrageurs are present in a market, in addition to internalizing their own price impact when making investment decisions, they also internalize the impact of their trades on the constraints and portfolio decisions of other large traders.

This chapter studies how wealth constraints of strategic arbitrageurs affect their willingness to invest, and the dynamics of prices. Arbitrageurs can invest in a fundamentally riskless arbitrage opportunity. They are required to have positive mark-to-market capital at all times, and if they violate this constraint, they have to liquidate their risky positions. As their portfolio is evaluated at up-to-the-minute market information, some arbitrageurs can adversely affect market prices and hence trigger the liquidation of others. I show that whether arbitrageurs behave cooperatively or engage in predatory behaviour depends on their size of investment in the arbitrage opportunity. When arbitrageurs have similar proportions invested in the arbitrage, they behave cooperatively, and spread their orders over several trading periods to minimize price impact. However, if there is significant difference in this ratio, the trader with low proportion of wealth invested in the arbitrage predates on the trader with high proportion of wealth in the arbitrage, and forces her to exit the market. Moreover, I show that the threat of predation can make arbitrageurs reluctant to invest in the first place, and they only exploit the mispricing shortly before it disappears.

To analyze the effect of wealth constraints on arbitrage trading I consider the following setup, which partially builds on the models of Gromb and Vayanos (2002) and Kondor (2009). Two assets with identical payoffs are traded in segmented markets at different prices, and arbitrageurs take long-short positions to exploit this mispricing. In the absence of arbitrageurs, the gap between prices would be constant for a finite time horizon, then it would exogenously disappear. Therefore, the arbitrage is fundamentally riskless. Arbitrageurs, by trading, endogenously determine the size of the gap. If arbitrageurs on aggregate buy more of the cheap asset and short more of the expensive asset, i.e. they short the gap, prices of the assets converge, and the gap shrinks. On the other hand, if arbitrageurs sell the cheap asset and buy the more expensive, i.e. they go long in the gap, prices diverge, and the gap widens. I consider a finite set of large arbitrageurs who invest in this arbitrage opportunity. Arbitrageurs have two important features. They are strategic, that is, they realize hey have a price impact on the gap, and they face wealth constraints, that is, they must fully collateralize for losses. Moreover, when their capital is insufficient, arbitrageurs must close their positions and leave the market. The wealth constraint thus implies that arbitrageurs' capital limits the positions they can take if they do not want to violate the constraint. However, the liquidation constraint can also provide incentives for some arbitrageurs to make prices diverge and trigger the insolvency of other traders.

The main results are obtained in a framework with two arbitrageurs. Suppose first that arbitrageurs already have some bets in place about the gap. I show that their behaviour depends on their exposure to the arbitrage opportunity. In particular, if traders have similar proportion of capital invested in the assets, they behave cooperatively, and the equilibrium gap decreases quickly. Arbitrageurs compete with each other and rush to the market, hence prices converge, and the wealth constraint never binds. However, if there is a significant difference in the proportion of their capital invested in the arbitrage opportunity, the trader with lower proportion of wealth invested in the gap predates on the trader with high proportion of wealth invested in the gap: the former (short-)sells the cheap asset and buys the expensive one, thus prices diverge. Arbitrageurs suffer losses, but these losses are higher for the arbitrageur who has invested more in the gap. If she violates her wealth constraint, she is forced to close her positions in the following period. This in turn widens the price gap even more, and makes future investment opportunities even better for the sole solvent arbitrageur.[1]

Given the cooperative or predatory behaviour discussed above, I also examine whether arbitrageurs are willing to invest in the arbitrage opportunity at the first place if they know they can become exposed to predation by other arbitrageurs. It is important to emphasize that the possible future losses are all due to predatory behaviour as opposed to unforeseen shocks, and are all subject to more than one arbitrageur being present in the market. As liquidation is costly, the threat of predation by other arbitrageurs implies that strategic traders reduce their initial investments so that liquidation does not happen in equilibrium. However, as long as one arbitrageur has a much higher level of capital than the other, it does not affect the gap path significantly, because the increased investment of the former compensates for the small position taken by the latter. I show that the wealth constraint has its strongest effect on the gap process when arbitrageurs start with similarly low level of capital. In this case arbitrageurs are reluctant to invest much, as shorting one more unit of

[1] The following quote provides an insight on the recent forced liquidation of Focus Capital, by suggesting that arbitrageurs occasionally decide to withdraw liquidity from markets, making prices diverge from fundamentals and forcing distressed institutions to unwind some of their positions at great losses:

> "In a letter to investors, the founders of Focus, Tim O'Brien and Philippe Bubb, said it had been hit by "violent short-selling by other market participants", which accelerated when rumors that it was in trouble circulated. Sharp drops in the value of its investments led its two main banks to force it to sell last Tuesday, according to the letter." (Financial Times, March 4, 2008)

Other famous examples of predatory trading include the near-collapse of Long-Term Capital Management (LTCM) in 1998, when Goldman Sachs and other counterparties strategically traded against LTCM to aggravate its situation. The proposal of UBS Warburg, to take over Enron's traders without taking over its trading positions, was opposed on the same ground - it presented potential predatory risk (AFX News Limited, AFX-Asia, January 18, 2002). See Edwards (1999) and Loewenstein (2000) for detailed analyses on the LTCM crisis, and Table I of Brunnermeier and Pedersen (2005) for an extensive list on examples of predatory trading.

the gap has a large effect on the proportion of wealth put into the arbitrage opportunity, and exposes the trader to become a prey of the other arbitrageur. Therefore, the gap changes very little initially, and agents only race to the arbitrage opportunity later.

These results are very much in contrast to the case with a single (monopolistic) arbitrageur. She knows that she faces a one-sided bet: if the trader shorts the gap, prices converge. This implies that her mark-to-market wealth never decreases, and the wealth constraint never binds. In the absence of other arbitrageurs, she gradually provides liquidity to the local markets to minimize her price impact, and her profits are not competed away. My analysis suggests that as the presence of other arbitrageurs creates predatory risk, increased competition in liquidity provision does not always imply that market segmentation and abnormal profits disappear quickly.

The model presented here is related to several strands of the literature, in addition to that on financial constraints. It is connected to models of limited arbitrage, including Shleifer and Vishny (1997), Xiong (2001), Gromb and Vayanos (2002), and Liu and Longstaff (2004). A large part of this literature focuses on potential losses in convergence trading due to institutional frictions or capital constraints. The common element in these models is that their mechanisms amplify exogenous shocks: arbitrageurs have to liquidate part of their positions after an initial shock to prices which creates further adverse price movements and liquidations. In my model, the amplification mechanism is endogenized and entirely strategic. Arbitrageurs are not fully competitive, and hence some of them can exploit their price impact to force others into distress. This type of strategic interaction, which is missing from the above papers, makes a fundamentally riskless arbitrage opportunity risky.

The two papers closest to my analysis on financially constrained arbitrage are Kondor (2009) and Attari and Mello (2006). Kondor (2009) develops an equilibrium model of convergence trading and its impact on asset prices, where arbitrageurs optimally decide how to allocate their limited capital over time. He shows that prices of identical assets can diverge even if the constraints faced by arbitrageurs are not binding, and that in equilibrium arbitrageurs' activity endogenously generates losses with positive probability, even if the trading opportunity is fundamentally riskless. Whereas he works with one representative arbitrageur and his focus is on the endogenous determination of the price gap, I study the trading behaviour of imperfectly competitive arbitrageurs, who try to exploit the vulnerability of each other by engaging in predatory trading. Attari and Mello (2006) analyze the trading strategy of a monopolistic arbitrageur who can, to some extent, influence the dynamics of prices on which capital requirements are based. They show that financial constraints are responsible for volatile prices and for time variation in the correlations of prices across markets. In contrast, my model allows for heterogeneity among arbitrageurs and focuses on the strategic interaction among them. Moreover, the lack of uncertainty allows me to provide analytical solution in my setting, while they can only numerically solve their model.

The model also belongs to those on predatory trading (i.e. trading that induces and/or exploits the need of other investors to reduce their positions) and forced liquidation. Brunnermeier and Pedersen (2005) show that if a distressed trader needs to sell for exogenous reasons, others also sell and subsequently buy back the asset. This leads to price overshooting and a reduced liquidation value for the distressed trader. Hence, the market is illiquid when liquidity is most needed. Carlin et al. (2007) analyze how episodic illiquidity can arise from a breakdown in cooperation between market participants. They consider a repeated setting of a predatory stage game and show that while most of the time traders provide apparent liquidity to each other, when the steaks are high, cooperation breaks down, leading to sudden and short-lived illiquidity. In these papers liquidation is exogenously imposed on some agents, as arbitrageurs become distressed due to an adverse shock and have to liquidate, while solvent traders take advantage of them. In contrast, the model presented here endogenizes the solvency of arbitrageurs: as capital requirements depend on observed prices, arbitrageurs might be able to induce the distress of others by manipulating the price, thus giving rise to predatory risk, which discourages investors from investing in the arbitrage opportunity.[2]

Abreu and Brunnermeier (2002, 2003) also provide a model with limited willingness of arbitrageurs to exploit a mispricing. They consider a setup where arbitrageurs want to invest while other arbitrageurs are investing, but asymmetric information causes a coordination problem. In contrast, in the model of this chapter information is symmetric, and arbitrageurs want to invest when others do not. It creates an incentive to drive other investors out from the market, which in turn prevents arbitrageurs with limited capital from investing much in the first place.

[2]See also papers that concentrate on endogenous risk as a result of amplification due to financial constraints, e. g. Bernardo and Welch (2004), Danielsson et al. (2004, 2011), and Morris and Shin (2004).

The chapter proceeds as follows. Section 2.2 presents the general model. Section 2.3 solves the case with a single arbitrageur. Section 2.4 derives the equilibrium of the model with two strategic arbitrageurs. Section 2.5 analyzes the effect of predatory threat on the initial investment decisions. Finally, Section 2.6 concludes.

2.2 Model

The model is similar to the setups of Gromb and Vayanos (2002) and Kondor (2009). Time is discrete and there are four periods, $t = 0, 1, 2,$ and 3. There is a set of arbitrageurs who can invest in two traded assets: a riskless bond and a fundamentally riskless arbitrage opportunity. The riskless bond has a constant return, normalized to one. The arbitrage opportunity is called a (price) gap, denoted by g_t in period $t = 0, ..., 3$. I assume that this gap starts at an initial level of $\overline{g} > 0$ and disappears due to an exogenous shock at date 3, i.e. $g_3 = 0$. I also assume, and then confirm in equilibrium, that it is always non-negative.

The natural interpretation of the gap is the difference between the prices of two risky assets with identical payoffs that are traded in segmented markets by local traders, and only a set of arbitrageurs can trade in both of them.[3] The prices can be different due to an initial supply shock to the local traders in one market, which disappears at date 3. In this setting, arbitrageurs can take long-short positions by buying the cheaper asset and shorting the expensive asset. This strategy gives a fundamentally riskless arbitrage opportunity if held until the price difference disappears at date 3, which can also be thought of as the maturity of the gap. Investing more into the arbitrage opportunity, which is essentially betting on the converge of the prices of the two assets, happens by increasing the long position in the cheap asset and increasing the short position (in absolute terms) in the expensive asset. I also refer to this as shorting the gap.[4]

There are a finite number arbitrageurs, denoted by I, which for simplicity is either one

[3] See Gromb and Vayanos (2002) for a microfoundation in this spirit. Not modelling the local markets and using the shortcut of a gap asset means that arbitrageurs are not allowed to take asymmetric positions in the two assets.

[4] As the focus of this analysis is on the strategic interaction among large traders facing an arbitrage opportunity, I take market segmentation for local traders as given. Gromb and Vayanos (2002), Zigrand (2004) and Kondor (2009) use similar assumptions. See Van Nieuwerburgh and Veldkamp (2009, 2010) for an information-based mechanism that results in endogenous market segmentation.

or two. Arbitrageurs behave strategically. In particular, when there is a single arbitrageur, $I = 1$, she has monopoly power over providing liquidity in the local markets. I refer to the case with two arbitrageurs, $I = 2$, as a duopoly of strategic traders. Arbitrageurs, indexed by $i = 1, ..., I$, are assumed to be risk neutral. They start with positive capital M^i in the bond and no initial endowment in the arbitrage opportunity, $x^i = 0$, and maximize expected utility of their date 3 wealth.

Arbitrageurs' activity affects the difference between the prices of local assets. In particular, when arbitrageurs in aggregate short x_t units of the gap, its level is given by

$$g_t = \overline{g} - \lambda x_t,\ t = 0, 1, 2, \tag{2.1}$$

where $\lambda > 0$ is an exogenously given illiquidity parameter that describes the price impact of arbitrage trades.[5] Equation (2.1) can also be given in the dynamic form:

$$g_t = g_{t-1} - \lambda \left(x_t - x_{t-1} \right) \tag{2.2}$$

for $t = 0, 1, 2$, and with $g_{-1} \equiv \overline{g}$ and $x_{-1} \equiv 0$. Equation (2.2) shows that when arbitrageurs increase their long position in the cheaper asset and their short position in the expensive asset by one unit, the price difference decreases, and the gap shrinks by λ.

Moreover, arbitrageurs are subject to wealth constraints. In particular, they are required to have non-negative marked-to-market wealth at all times.[6] If a trader violates this constraint, i.e. she defaults, she has to close all her positions in the following period. I refer to this as fire-sale or liquidation. Formally, if arbitrageur i has M^i_{t-1} in the riskless bond and a short position of x^i_{t-1} units of the gap after trading at date $t-1$, then her mark-to-market wealth is $M^i_{t-1} - g_{t-1} x^i_{t-1}$, and the constraint can be written as:

$$\text{if } M^i_{t-1} - g_{t-1} x^i_{t-1} < 0, \text{ it must be that } x^i_t = 0.$$

[5] I assume that increasing the short position in the gap by one unit always has the same price impact (as long as $x_t < \overline{g}/\lambda$). It holds, for example, if local traders having exponential utility and asset payoffs are normally distributed.

[6] The specific wealth constraint considered in this model is just one of many financial constraints that are based on market prices, e.g. margin constraints (Brunnermeier and Pedersen (2009) and Garleanu and Pedersen (2011)), or value at risk (VaR) constraints (Garleanu and Pedersen (2007)). They would lead to qualitatively similar results.

The wealth constraint requires that arbitrageurs can always cover their accumulated losses from their bond positions. As long as they do not default, they do not face any restrictions on their orders in the following trading period. However, when they do default, they must close their positions immediately, i.e. sell all the risky assets that they hold and buy back what they short in the following period.[7]

Arbitrageur i's optimization problem is as follows:

$$\max_{\left\{x_t^i\right\}_{t=0}^2} W_3^i\left(M^i\right) = M^i + \sum_{t=0}^{2} g_t\left(x_t^i\right)\left(x_t^i - x_{t-1}^i\right), \tag{2.3}$$

subject to the evolution of the gap:

$$g_t = g_{t-1} - \lambda \sum_{i=1}^{I}\left(x_t^i - x_{t-1}^i\right) \text{ for } t = 0, 1, 2, \tag{2.4}$$

and the wealth constraint:

$$x_t^i = 0 \text{ if } M_{t-1}^i < g_{t-1} x_{t-1}^i \text{ for } t = 0, 1, 2, \tag{2.5}$$

where $g_{-1} \equiv \overline{g}$, $M_{-1}^i \equiv M^i$ and $x_{-1}^i \equiv 0$ for $i = 1, ..., I$. In each trading period $t = 0, 1, 2$, first it is determined whether an arbitrageur is solvent. Second, the risky asset is traded.

The equilibrium of the economy is defined as follows:

Definition 12 *A dynamic Nash-equilibrium of the trading game consists of the gap* $\{g_t\}_{t=0}^2$ *and the holdings of arbitrageurs* $\left\{x_t^i\right\}_{t=0}^2$ *for* $i = 1, ..., I$*, such that* $\left\{x_t^i\right\}_{t=0}^2$ *solve* (2.3) *subject to* (2.4) *and* (2.5).

Before proceeding to the solution of the model, I make two observations about the optimization problem and the wealth constraint.

First, it is important to notice that as long as there is a single arbitrageur, i.e. she has monopoly power in providing liquidity, the market price used to evaluate her portfolio only

[7] The combination of the wealth constraint and the liquidation can be thought of as a shortcut for the joint effect of two well-known phenomena. On one hand, the relationship between past performance and fund flows has been documented for various asset classes. See, for example, Chevalier and Ellison (1997) and Sirri and Tufano (1998), or Berk and Green (2004), who provide a model of active portfolio management when fund flows rationally respond to past performance. On the other hand, Coval and Stafford (2007) show that funds experiencing large outflows decrease existing positions by engaging in fire-sales, which creates price pressure.

depends on her risky holdings. However, when there are at least two strategic agents, the trade order of one of them influences the market clearing price and hence affects the constraint status of the other arbitrageur. In particular, widening the gap between the prices of the two assets creates losses to someone who is betting on the convergence of prices, and might even trigger her fire-sale. When this distressed trader is forced to close her positions, this further widens the gap, and creates a more profitable opportunity to agents still solvent. Therefore, although it is costly to trade against price convergence, there is also a benefit of having a better investment opportunity later on. Moreover, an arbitrageur close to bankruptcy might not mind violating her constraint at all. When others are betting on divergence and thus are effectively widening the gap, it can be very costly to support the price to ensure that she remains solvent.

Second, there is a natural way to simplify the wealth constraint (2.5). Since the dynamics of the riskless position can be expressed as

$$M_t^i = M_{t-1}^i + g_t \left(x_t^i - x_{t-1}^i \right) \tag{2.6}$$

for $t = 0, 1, 2$, it is easy to show that requiring non-negative capital at time t, $M_t^i - g_t x_t^i \geq 0$, is equivalent to

$$M_{t-1}^i \geq g_t x_{t-1}^i. \tag{2.7}$$

If it does not hold, arbitrageur i is forced to liquidate in the following period: $x_{t+1}^i = 0$. However, in this 4-period economy, marking to market is only relevant after period 1. This is because for $t = 0$, condition (2.7) is equivalent to $M^i \geq g_0 x^i$, which always holds as arbitrageurs start with positive bond positions ($M^i > 0$) and no endowment in risky assets ($x^i = 0$). In addition, violating the constraint at $t = 2$ would mean that an arbitrageur has to liquidate her risky position in period 3, but there is no trading at date 3 as assets already pay off. Therefore the wealth constraint is only relevant after period $t = 1$: if arbitrageur i fails to satisfy

$$M_0^i \geq g_1 x_0^i, \tag{2.8}$$

she must liquidate at period 2, i.e. have $x_2^i = 0$.

Further simplification of (2.8) can provide additional intuition regarding the nature of the constraint. In particular, from (2.6), (2.8) is equivalent to

$$M^i \geq (g_1 - g_0)\, x_0^i. \tag{2.9}$$

The left hand side of this inequality is the mark-to-market wealth of arbitrageur i at date 0, which is positive by assumption, and hence the agent is not distressed at date 0. The right hand side of the inequality represents the loss arbitrageur i makes on her positions between date 0 and 1. Hence, (2.9) requires the arbitrageur's wealth before trading at date 1 to be enough to cover all the losses suffered on her initial position. However, it might not always hold. In particular, when initially arbitrageur i is shorting the gap, $x_0^i > 0$, but it actually widens, $g_1 > g_0$, the wealth constraint gets tighter and she can become distressed if her starting capital is not sufficiently high. Similarly, arbitrageur i's wealth constraint gets tighter if she bets on price divergence, $x_0^i < 0$, while the gap shrinks, $g_1 < g_0$. On the other hand, as long as arbitrageur i bets on the convergence (divergence), and prices do converge (diverge), the constraint gets relaxed.

2.3 Monopoly

In this section I solve for the optimal trades of the unconstrained and the constrained monopolist arbitrageur. With a sole arbitrageur, $I = 1$, the trading game simplifies to a portfolio choice problem, subject to a wealth constraint that affects the trading speed of the agent.

Dropping the superscript referring to the only arbitrageur $i = 1$, her optimization problem can be written as:

$$\max_{\{x_t\}_{t=0}^2} W_3^i(M) = M + \sum_{t=0}^{2} g_t(x_t)(x_t - x_{t-1}) \tag{2.10}$$

subject to market clearing:

$$g_t = g_{t-1} - \lambda (x_t - x_{t-1}) \text{ for } t = 0, 1, 2, \text{ and } g_{-1} \equiv \overline{g},$$

and the insolvency constraint:

$$x_2 = 0 \text{ if } M < (g_1 - g_0)\, x_0. \tag{2.11}$$

First, I solve the optimization problem without (2.11). The optimal trades and the gap process in absence of the wealth constraint are summarized in the following result:

Proposition 13 *The unconstrained monopolist arbitrageur gradually provides liquidity in the local markets, i.e. she trades the same amount in every period. Formally,*

$$x_{0,u} = \frac{1}{4\lambda}\overline{g},\ x_{1,u} = \frac{1}{2\lambda}\overline{g},\ \text{and}\ x_{2,u} = \frac{3}{4\lambda}\overline{g},$$

and the gap decreases linearly over time:

$$g_{0,u} = \frac{3}{4}\overline{g},\ g_{1,u} = \frac{1}{2}\overline{g},\ \text{and}\ g_{2,u} = \frac{1}{4}\overline{g}.$$

Proposition 13 states that in case there is a single strategic trader taking advantage of the mispricing across markets, her early trades only compete with her later trades. As she can commit to a strategy that minimizes her price impact, she smoothes her orders across several dates, and hence trades the same amount in each period. This is illustrated on Figure 2-1.

Suppose now that the monopolist arbitrageur is subject to wealth constraint (2.11), which might prevent her to supply liquidity as in Proposition 13. The main question is whether a trader endowed with positive capital and facing a riskless arbitrage opportunity would ever get to a state where she faces liquidation. The answer is negative:

Proposition 14 *The wealth constraint never binds on the equilibrium gap path. Therefore it does not affect the trading of a monopolist arbitrageur, and does not influence the convergence speed of the two prices.*

The result of Proposition 14 is rather straightforward. It is obvious that the constrained arbitrageur can never be better off than the unconstrained arbitrageur of Proposition 13. However, she can achieve the same terminal wealth. This is because when a single strategic trader shorts the gap, the convergence is purely the effect of her trade. Consequently, she is making profits throughout the whole process, and the gap decreases, $g_1 - g_0 < 0$. The wealth constraint thus never binds, and in fact never affects the equilibrium trading of the arbitrageur.

2.4 Duopoly

2.4.1 Benchmark case

Similarly to the monopoly case, I start with characterizing the equilibrium orders and the gap process when there are two strategic arbitrageurs, $I = 2$, and they face no constraints on the positions taken in the gap asset. However they are aware that investing one more unit of capital at a certain date decreases the return on future investments of both arbitrageurs. It has two contrasting implications regarding their trading behaviour. First, they would like to trade slowly to minimize their price impact. Second, both of them would still like to trade faster than the other arbitrageur. Formally, I obtain the following results:

Proposition 15 *The equilibrium holdings of unconstrained duopolist arbitrageurs are give by*

$$x_{0,u}^i = \frac{385}{1299\lambda}\overline{g},\ x_{1,u}^i = \frac{182}{433\lambda}\overline{g},\ \text{and}\ x_{2,u}^i = \frac{205}{433\lambda}\overline{g},\ \text{for}\ i = 1, 2,$$

and the gap decreases as

$$g_{0,u} = \frac{529}{1299}\overline{g},\ g_{1,u} = \frac{69}{433}\overline{g},\ \text{and}\ g_{2,u} = \frac{23}{433}\overline{g}.$$

Figure 2-2 illustrates the evolution of the gap and the holdings of the duopolist arbitrageurs, and contrasts the gap processes in the monopoly and duopoly cases. The main message of Proposition 15 is that when there are two strategic traders taking advantage of the mispricing across markets, these competing arbitrageurs race to the market, and the price gap decreases much faster than with a single arbitrageur.

This result is clearly intuitive. As before, illiquidity gives arbitrageurs an incentive to spread trades over time, in order to minimize their price impact. However, now the trade order of an arbitrageur at a certain date not only competes with her later investments, but also with all the present and future investments of the other arbitrageur. As arbitrageurs face a downward sloping demand curve, they both try to trade before the other arbitrageur trades, and the presence of another arbitrageur leads to competition between them. The equilibrium strategy shows that the second effect is stronger than the first. This is why duopolist strategic traders cannot commit to a strategy that minimizes their joint price impact and takes advantage of the mispricing the most efficient way (from the viewpoint of arbitrageurs in aggregate). Instead they both race to the market at date 0. As Figure 2-2 shows, trading volume is large in the early periods; and the gap converges faster than with a single arbitrageur, and slows down later.

2.4.2 Constrained case

In the remainder of this section I consider a subgame of the optimization program (2.3) to study how wealth constraints affect arbitrage activity with two strategic traders. I assume that some trading at date 0 has already taken place: the price gap is given by g_0, and arbitrageurs already have short positions x_0^i in the gap asset and bond holdings M_0^i, $i = 1, 2$.[8] I proceed to the overall solution in Section 2.5 after discussing the equilibria of the subgame and the notion of predatory threat.

The optimization problem of agent i is the following:

$$\max_{x_1^i, x_2^i} W_3^i \left(M_0^i, x_0^i, g_0\right) = M_0^i + g_1\left(x_1^i\right)\left(x_1^i - x_0^i\right) + g_2\left(x_2^i\right)\left(x_2^i - x_1^i\right).$$

subject to market clearing:

$$g_t = g_{t-1} - \lambda\left(x_t^i - x_{t-1}^i + x_t^{-i} - x_{t-1}^{-i}\right) \text{ for } t = 1, 2 \text{ and } i = 1, 2,$$

where $-i$ denotes the other agent; and the insolvency constraints:

$$x_2^i = 0 \text{ if } M^i < \left(g_1\left(x_1^i\right) - g_0\right)x_0^i, \text{ and } x_2^{-i} = 0 \text{ if } M^{-i} < \left(g_1\left(x_1^i\right) - g_0\right)x_0^{-i}.$$

The second wealth constraint indicates that arbitrageur i is aware of the constraint for arbitrageur $-i$, and hence can influence the price to trigger her fire-sale.

To define an equilibrium, I define the states of the world and two notions of value functions as follows:

Definition 16 *At date* 1 *each arbitrageur can be in one of three states: (i) state* n *for the*

[8] One reason for being endowed with the risky assets before trade starts would be because traders previously enjoyed some (unmodelled) private benefits from holding them.

constraint being satisfied and not binding *at the equilibrium holding and gap, i.e.* $M^i > \left(g_1\left(x_1^i\right) - g_0\right) x_0^i$*; (ii) state* b *for the constraint* binding, $M^i = \left(g_1\left(x_1^i\right) - g_0\right) x_0^i$*; or (iii) state* v *for the constraint being* violated, $M^i < \left(g_1\left(x_1^i\right) - g_0\right) x_0^i$.

At date 2 *each arbitrageur can be in one of two states: (i) state* s *for* solvent *(i.e. trade freely), or (ii) state* l *for* liquidated/insolvent *(i.e. having to close her risky position).*

The dynamics of states are as follows: (i) If arbitrageur i satisfies her wealth constraint, she can freely trade in period 2*. Formally, if arbitrageur i is in state* n *or* b *at date* 1*, she gets to state* s *at date* 2*; (ii) On the other hand, if the arbitrageur violates the constraint, she must liquidate in period* 2*. Formally, if agent i is in state* v *at date* 1*, she gets to state* l *at date* 2*.*

Given the definition of states, one can define the state-dependent value functions:

Definition 17 *The* state-dependent (or conditional) value function *of agent* $i = 1, 2$ *in period* $t = 1, 2$ *and arbitrageur states* $\{jk\}$ *is denoted by* $V_{t,jk}\left(M_t^i, x_t^i, M_t^{-i}, x_t^{-i}\right)$*, where*

j *and* k *are the states of arbitrageur* i *and* $-i$*, respectively;* $j, k \in \{n, b, v\}$ *if* $t = 1$*, and* $j, k \in \{s, l\}$ *if* $t = 2$*;*

M_t^i *and* x_t^i *are the after-trade holdings of arbitrageur* i*; and*

M_t^{-i} *and* x_t^{-i} *are the after-trade holdings of arbitrageur* $-i$*.*

Based on the state-dependent value functions I define the value function such that the optimization problem is the problem of choosing the optimal demand and the state jointly:

Definition 18 *The* value function *of agent* i at date t *is the merger of different conditional value functions from different states of the world given as*

$$V_t^i\left(M_t^i, x_t^i, M_t^{-i}, x_t^{-i}\right) = \sum_{j,k} 1_{jk} V_{t,jk}\left(M_t^i, x_t^i, M_t^{-i}, x_t^{-i}\right)$$

where 1_{jk} *is an indicator, and takes the value of* 1 *if, based on their date* 1 *mark-to-market portfolio value, arbitrageur* i *is in state* j *and arbitrageur* $-i$ *is in state* k*, and zero otherwise.*

Finally, given the value function, I take the standard definition of a Nash-equilibrium:

Definition 19 *A* Nash-equilibrium *of the economy is a vector of demands* $\{x_t^i\}_{i=1,2;t=1,2}$ *such that* x_t^i *solves the program*

$$\max_x V_t^i\left(M_t^i, x_t^i, M_t^{-i}, x_t^{-i} | x_t^{-i}; M_{t-1}^i, x_{t-1}^i, M_{t-1}^{-i}, x_{t-1}^{-i}\right)$$
$$= V_t^i\left(M_{t-1}^i + g_t(x)\left(x - x_{t-1}^i\right), x, M_{t-1}^{-i} - g_t(x)\left(x_t^{-i} - x_{t-1}^{-i}\right), x_t^{-i}\right)$$

where $g_t(x)$ *is the market-clearing gap in period* t *when agent* i *submits the demand* x*, and agent* $-i$ *submits her equilibrium demand* x_t^{-i}.

Before proceeding to the equilibria of this game, let me make an observation about the wealth constraint. As described in (2.8), the wealth constraint can be expressed as $M^i \geq (g_1 - g_0)\, x_0^i$ for $i = 1, 2$. Thus, if arbitrageur i enters period 1 with a zero position in the risky assets, $x_0^i = 0$, her constraint will never bind. Suppose now that both arbitrageurs have taken non-zero positions at date 0. Then M^1/x_0^1 and M^2/x_0^2 exist, and they describe the inverse of the proportion of wealth invested in the gap. Suppose further that both x_0^1 and x_0^2 are positive (as it is going to be in equilibrium), that is arbitrageurs initially bet on the convergence of prices. It implies that the wealth constraints can be rewritten in the form

$$\frac{M^1}{x_0^1} \geq g_1 - g_0 \text{ and } \frac{M^2}{x_0^2} \geq g_1 - g_0.$$

It is easy to see that as long as the proportion invested in the gap asset is different between agents, for example $M^1/x_0^1 > M^2/x_0^2$, there is a natural order between arbitrageurs. If arbitrageur 2 is solvent, arbitrageur 1 remains solvent too. On the other hand, if arbitrageur 1 is insolvent, arbitrageur 2 has to liquidate too. Moreover, there always exists a gap level g_1 such that arbitrageur 1 remains solvent while arbitrageur 2 goes bankrupt. Therefore the trader with higher M^i/x_0^i ratio, i.e. lower proportion of wealth invested in the arbitrage opportunity, can always be more *aggressive*, while the arbitrageur with higher proportion of wealth invested in the gap must be more *cautious* with her trades. In the characterization of the equilibrium I will refer to them as arbitrageurs a and c.[9]

[9] If $M_0^1/x_0^1 = M_0^2/x_0^2$, the constraint binds for them at the same time. It implies that either both arbitrageurs remain solvent, or they both go bankrupt. Also, when, for example, arbitrageur 1 does not trade in period 0, i.e. $x_0^1 = 0$, the constraint will never bind for her. This case can be thought of as the limit when $M_0^1/x_0^1 \to \infty$.

Before proceeding to the solution of the model, I discuss the methodology of the equilibrium construction. The above problem can be solved backwards. First I solve for the optimal trades at date 2 given the conjectured state arbitrageurs are in (ss, sl, or ll), and obtain value functions representing their continuation utilities. Then I solve for the optimal trades of period 1. The complexity of the solution arises here regarding how to deal with the liquidation constraint. The possibility of forced liquidation implies that the optimization problem of an arbitrageur is globally non-continuous and non-concave, so local conditions for the equilibrium are not sufficient. However, the optimization problem is locally concave almost everywhere. Figures 2-3 and 2-4 illustrate the utility of an arbitrageur as a function of her trade at date 1 while holding the other arbitrageur's date-1 trade constant in two particular cases. It is straightforward that the optimization problem can always be divided into three segments that correspond to the states of the world such that the utility function is concave in each segment.[10] The possible portfolios of an arbitrageur in one segment lead to a different continuation state from portfolios in another segment: if a trader increases her short position sufficiently, the gap shrinks and both arbitrageurs remain solvent. However if an arbitrageur decides to go long in the gap, the gap widens, and can push (at least) one arbitrageur into distress. Consequently, for each portfolio choice of the other trader, an arbitrageur compares the locally optimal investment strategies in the three segments, and picks the one with highest utility.

Because of the local concavity, given the other arbitrageur's investment decision, there is an optimal portfolio within each state of the world. Combining these conditions for the two arbitrageurs gives a set of candidate equilibria, satisfying that none of the traders want to alter their strategies as long as the state of the world remains the same. Therefore, it must be also checked whether these trades are globally optimal too, i.e. whether any arbitrageur would prefer to deviate in such a way that changes the state of the world.

[10] The three states from the viewpoint of the aggressive arbitrageur are ss, sl and ll. From the viewpoint of the cautious arbitrageur the possible states are ss, ls and ll. This is because the roles of arbitrageurs imply that it is impossible to have a case when the cautious arbitrageur remains solvent and the aggressive arbitrageur becomes insolvent.

Candidate equilibria

I describe the equilibria of the economy in two steps. First, I provide the set of *candidate equilibria* with the locally optimal portfolios, which also determine the gap path. The derived date-1 gap g_1, combined with (2.9), thus provides a straightforward necessary condition on the proportion of wealths invested for such an equilibrium to exists. Then I discuss the actual equilibria of the economy for the three cases when (i) both arbitrageurs remain solvent; (ii) the aggressive arbitrageur remains solvent, but the cautious is insolvent; or (iii) both arbitrageurs go bankrupt. These are different from the candidate equilibria because the globally optimal portfolios must satisfy more requirements than local optimality. For tractability, I only discuss the cases when arbitrageurs initially short the gap asset, i.e. $x_0^a, x_0^c > 0$, which will be the case in equilibrium. All the other cases are described in an internet appendix.

Proposition 20 *When both arbitrageurs remain solvent, the locally optimal strategies and the gap path are given by*

$$x_1^i - x_0^i = \frac{7}{23\lambda} g_0 \text{ and } x_2^i - x_1^i = \frac{3}{23\lambda} g_0 \text{ for } i = a, c. \tag{2.12}$$

and

$$g_1 = \frac{9}{23} g_0 \text{ and } g_2 = \frac{3}{23} g_0.$$

Such a candidate equilibrium exists for every $0 < M^c/x_0^c < M^a/x_0^a$. *Moreover, the wealth constraint is not binding for any arbitrageur.*

Suppose that both arbitrageurs remain solvent, and it happens without the constraint binding for the cautious arbitrageur. It implies that the locally optimal strategies are those that would emerge in the equilibrium of the economy with no wealth constraint.[11] As before, since arbitrageurs face a downward sloping demand curve, they both try to trade before the other arbitrageur trades. It leads to competition between them: arbitrageurs race to the market, and the gap shrinks quickly. Since the gap decreases, $g_1 < g_0$, arbitrageurs record profits throughout the convergence, thus they indeed remain solvent even if they start with

[11] For this note that substituting the date-0 unconstrained gap, $g_{0,u}$, into g_0 gives the same portfolios that were derived in Proposition 15.

very low capital. Moreover, the constraint of arbitrageur c cannot bind in equilibrium, because that would imply the gap must widen, $g_1 > g_0$. However, in this case both arbitrageurs would be willing to short the gap a little bit more to trade before the other arbitrageur, hence the gap would shrink, and the constraint would not bind any more.

Proposition 21 *When the aggressive arbitrageur remains solvent and the cautious liquidates:*

(i) There exists a candidate equilibrium where the wealth constraint is not binding for the aggressive arbitrageur. The locally optimal strategies and the gap path are given by

$$\begin{aligned} x_1^i - x_0^i &= \frac{1}{5\lambda}\left(g_0 - \lambda x_0^c\right) \text{ for } i = a, c, \\ x_2^a - x_1^a &= \frac{1}{2\lambda}\left(g_1 + \lambda x_1^c\right) \text{ and } x_2^c = 0, \end{aligned}$$

and

$$g_1 = \frac{3}{5}g_0 + \frac{2}{5}\lambda x_0^c \text{ and } g_2 = \frac{2}{5}g_0 + \frac{3}{5}\lambda x_0^c.$$

The candidate equilibrium requires $x_0^c > \frac{1}{\lambda}g_0$ *and* $0 < M^c/x_0^c < -\frac{2}{5}\left(g_0 - \lambda x_0^c\right) \leq M^a/x_0^a$.

(ii) There exists a candidate equilibrium where the wealth constraint is binding for the aggressive arbitrageur when $0 < M^c/x_0^c < M^a/x_0^a < -\frac{2}{5}\left(g_0 - \lambda x_0^c\right)$. *As the constraint is binding for the aggressive arbitrageur, the locally optimal strategies and the gap path satisfy* $g_1 - g_0 = M^a/x_0^a$. *Moreover, there are many possible optimal trades as this case corresponds to a corner solution.*

The proposition states that as long as the constraint does not bind for arbitrageur a, the locally optimal strategies satisfy that arbitrageurs sell the same amount from the cheap asset and buy the same amount from the expensive asset, driving the gap up at date 1. In fact, the cautious trader knows that if the aggressive trader goes long in the gap asset to widen the gap, she does not have enough capital to cover her losses emerging due to the price divergence, and she will be forced to close her position. As arbitrageurs face a downward sloping demand curve, the cautious trader wants to avoid a round-trip transaction (buying and then being forced to sell, or selling and then buying), because it would lead to additional losses. She also wants to minimize her price impact when liquidating. Therefore, she conducts the fire-sale in two periods, and closes part of her positions already at date 1 and the rest at date 2.

In the meantime, at date 1 the aggressive arbitrageur finds it optimal to do exactly the same the cautious arbitrageur does. Notice that the condition $x_0^c > \frac{1}{\lambda} g_0$ implies that the gap widens through time, i.e. $g_2 > g_1 > g_0$. Hence when trader c finishes the fire-sale, trader a will face a better arbitrage opportunity to invest than in the very beginning, as the gap is wider. In fact, the aggressive trader withdraws liquidity instead of providing liquidity exactly when the cautious arbitrageur would need it the most. This is in the spirit of Brunnermeier and Pedersen (2005). However, in this model predation happens endogenously, unlike in Brunnermeier and Pedersen (2005), where the prey is passive. Here arbitrageur c could avoid bankruptcy by taking a sufficiently large long position and ensuring she suffers no losses between periods 0 and 1, but she realizes it would be too costly for her. The constraints on the proportion of the wealth invested correspond to the fact that the aggressive arbitrageur can indeed cover her losses due to the gap diverging from g_0 to g_1, while the cautious arbitrageur cannot.

Proposition 21 also states that a qualitatively similar candidate equilibrium (but with different trades) can happen even if arbitrageur a has lower level of capital (or higher proportion of capital invested in the arbitrage opportunity). This is because with $M^c/x_0^c < M^a/x_0^a$ the aggressive arbitrageur can always set the gap such that her losses are still covered by her starting wealth while violating the wealth constraint of the cautious arbitrageur.

Proposition 22 *When both arbitrageurs become insolvent, the locally optimal strategies and the gap path are given by*

$$x_1^i - x_0^i = -\frac{1}{3}\left(x_0^i + x_0^{-i}\right) \text{ and } x_2^i = 0 \text{ for } i = a, c.$$

and

$$g_1 = g_0 + \frac{2\lambda}{3}\left(x_0^a + x_0^c\right) \text{ and } g_2 = g_0 + \lambda\left(x_0^a + x_0^c\right).$$

Such a candidate equilibrium exists if $0 < M^c/x_0^c < M^a/x_0^a < 2\lambda\left(x_0^a + x_0^c\right)/3$.

When both arbitrageurs violate the constraint and become insolvent, they have to strategically liquidate their positions through two periods. Arbitrageurs know that they are facing a downward sloping demand curve, and want to minimize their price impact while liquidating.

To close their positions, they have to buy the expensive asset and sell the cheap asset, hence they both want to buy/sell before the other arbitrageur. Thus they race to the market. In equilibrium, they liquidate the same amount at date 1, namely 2/3 of their aggregate asset holdings, and at date 2 they liquidate the remaining 1/3.

Equilibrium characterization

Given the locally optimal strategies, it is possible to analyze under what circumstances they are globally optimal too. Regarding the equilibrium with both arbitrageurs solvent, I obtain the following result:

Proposition 23 *There exists an equilibrium of the trading game with both arbitrageurs remaining solvent (state ss) if and only if*

$$\frac{M^a}{x_0^a} > \frac{M^c}{x_0^c} \geq \Delta^1_{nn,nv}\left(g_0, x_0^c\right), \tag{2.13}$$

where the function $\Delta^1_{nn,nv}\left(\cdot,\cdot\right) > 0$ *is given in Appendix 2.7.2.*

According to Proposition 20, it was possible to have a candidate equilibrium such that both arbitrageurs remain solvent for any proportions of wealth invested in the arbitrage opportunity, because prices converged and arbitrageurs made profits throughout the whole trading process. When looking for an actual equilibrium, turns out this is not the case.

In particular, as the aggressive arbitrageur is aware of the wealth constraint of the cautious agent, arbitrageur a can engage in the manipulation of date-1 prices. Facing a downward sloping demand curve, this manipulation is costly because of the price impact. However, manipulation can be profitable due to two sources of profits. First, if the cautious arbitrageur goes bankrupt, the aggressive arbitrageur has monopoly power in providing liquidity to local traders at date 2. Second, as arbitrageur c has a short position in the gap after period 1, i.e. $x_1^c = x_0^c + \frac{7}{23\lambda} g_0 > 0$, her fire-sale widens the gap and makes forced liquidation even more desirable for the aggressive trader. The cost of manipulation is decreasing in the proportion of arbitrageur c's wealth invested into the arbitrage opportunity, i.e. increasing in M^c/x_0^c, while the profit of the fire-sale is increasing in x_1^c, i.e. in both the cautious arbitrageur's holding before date 1, x_0^c, and the initial gap g_0. Combining these observations, there exists a threshold for M^c/x_0^c such that if the proportion of arbitrageur c's wealth invested into the arbitrage opportunity is low enough, forcing her to liquidate is too costly, and an equilibrium with both agents remaining solvent exists.

Next, I present the conditions under which predation happens.

Proposition 24 *There exists an equilibrium with the aggressive arbitrageur remaining solvent and the cautious arbitrageur becoming insolvent (state sl) if and only if*

$$0 < \frac{M^c}{x_0^c} \leq \Delta^c_{nv,nn}\left(g_0, x_0^c\right) \text{ and } \frac{M^a}{x_0^a} \geq \Delta^c_{nv,vv}\left(g_0, x_0^a, x_0^c\right), \tag{2.14}$$

where $\Delta^c_{nv,nn}\left(\cdot,\cdot\right), \Delta^c_{nv,vv}\left(\cdot,\cdot\right) > 0$ *are given in Appendix 2.7.2.*

Comparing Propositions 21 and 24, the main difference is that the wealth requirements are tighter. For an equilibrium it must be that the locally optimal strategies are globally optimal too. It is apparent that the key is whether the cautious arbitrageur would be better off avoiding liquidation as a result of some costly price manipulation at date 1 that changes the state of the world.

The cautious arbitrageur starts trading with an initial long position in the arbitrage opportunity, but due to her limited capital, she cannot sustain losses caused by the activity of the aggressive arbitrageur in the short run. It is apparent that if arbitrageur c wants to remain solvent, she can always do so. This is because if arbitrageur a widens the gap, trader c can always engage in exactly the opposite trade that leaves the gap unchanged, and thus leaves the state untouched as well. The question is how costly it is.

In particular, suppose the aggressive arbitrageur's strategy is fixed at buying a very large amount of the gap asset, which makes prices diverge. Arbitrageur c, being subject to the wealth constraint, can do two things. First, she can short enough so that she neutralizes the effect of the aggressive arbitrageur's trades and brings g_1 sufficiently close to the the original level g_0. In this case she remains solvent. As she faces a downward sloping demand curve, shorting a large amount of the gap asset is costly, as it diminishes future returns on the assets she is holding. On the other hand, the benefit of this strategy is that the arbitrageur remains solvent and can invest again at date 2. Alternatively, she can accept that she is pushed to insolvency. In that case the optimal liquidation strategy means shorting less at date 1, which leads to smaller price impact. Moreover, since the other trader is still solvent at date 2, the cautious arbitrageur can liquidate at more favourable prices.

Whether the cautious trader thus finds it optimal to liquidate or not, given the selling pressure of the aggressive trader, depends on the relative costs and gains of these two strategies. In particular, the profit from remaining solvent increases in her initial position x_0^c, and in the gap size g_0. This implies that the threshold for equilibrium on the proportion of wealth invested by the cautious arbitrageur, $\Delta_{nv,nn}^c(\cdot,\cdot)$, is an increasing function of both x_0^c and g_0.

Finally, regarding equilibria in which both agents get liquidated I obtain the following result:

Proposition 25 *There exists no equilibrium of the trading game with both arbitrageurs being insolvent.*

This result is rather intuitive. Indeed, liquidation imposes a cost on both agents, because they have to close the positions they previously created to bet on the convergence of prices. Put it differently, arbitrageurs must sell assets at lower prices than they have bought them, or buy back previously shorted assets at prices higher than when they started to short them. Given that the divergence in prices is solely the effect of their own activities, arbitrageurs could avoid these self-imposed costs by not trading at all in period 1. By simply holding on to their existing positions the gap would not change, and the constraint would not get tighter than before. Arbitrageurs would remain solvent and their optimal unconstrained trades in period 2 could not make them worse off than the forced liquidation.

The different regions for the proportions of wealth invested in the arbitrage opportunity described in Propositions 23 and 24 are illustrated on Figure 2-5.

2.5 Predatory threat and arbitrage

So far I have taken the initial positions x_0^i and the gap g_0 as given. In this section I endogenize x_0^i by extending the previous analysis with an investment phase at date 0. Arbitrageurs know that the initial positions they take and hence the gap they face affect which state of the world they get into after date 0. I show that liquidation does not happen in equilibrium, but as long as one arbitrageur has a much higher level of capital than the other, it does not affect the gap path significantly. I show that the wealth constraint has its strongest effect on the gap path when arbitrageurs start with similarly low level of capital. In this case the gap decreases very little in period 0, then both agents rush to the arbitrage opportunity.

When solving the date-0 optimization problem, I restrict the (on- and off-equilibrium) action space of arbitrageurs to trades with which they end up in either an ss or an sl equilibrium. It means that for a given x_0^2, arbitrageur 1 must choose her position x_0^1 in such a way that arbitrage positions maximize her utility while satisfying either (2.13) or (2.14). Of course when deciding on the initial investment x_0^i, arbitrageur i also realizes that as long as her proportion of wealth invested into the arbitrage opportunity is higher than that of the other trader, i.e. $M^i/x_0^i < M^{-i}/x_0^{-i}$, the wealth constraint is tighter for her, and hence she takes the role of the cautious arbitrageur. Formally, I look for a dynamic equilibrium where arbitrageur i solves the problem

$$x_0^i \in \arg\max_x W_3^i = V_0\left(x|M^i, M^{-i}, x_0^{-i}\right),$$

where

$$V_0\left(x|M^i, M^{-i}, x_0^{-i}\right) = \begin{cases} V_{0,ss}\left(M_0^i, x_0^i, M_0^{-i}, x_0^{-i}\right) & \text{if satisfy conditions for } ss \text{ equilibrium} \\ V_{0,sl}\left(M_0^i, x_0^i, M_0^{-i}, x_0^{-i}\right) & \text{if satisfy conditions for } sl \text{ equilibrium} \\ V_{0,ls}\left(M_0^i, x_0^i, M_0^{-i}, x_0^{-i}\right) & \text{if satisfy conditions for } ls \text{ equilibrium,} \end{cases}$$

As the optimization programs of arbitrageurs with these constraints become difficult to solve in closed form (it includes solving 4th order equations), I make some simplifying steps and then solve the problem numerically. In particular, first I solve the optimization problems given that both agents remain solvent while satisfying the constraints for an ss equilibrium, i.e.

$$\max_{x_0^i} V_{0,ss}\left(M_0^i, x_0^i, M_0^{-i}, x_0^{-i}\right) \equiv M_0^i + \frac{72}{23^2\lambda}g_0^2 = M^i + g_0 x_0^i + \frac{72}{23^2\lambda}g_0^2$$

subject to (2.13), and then I confirm that none of the agents have incentives to deviate to the sl state when the other arbitrageur chooses the optimal strategy x_0^{-i} that solves her program.[12]

[12]The deviations allowed here include those when the arbitrageur goes long in the gap, i.e. $x_0^i < 0$, even

Propositions 26 and 27 describe the equilibrium date-0 trading of strategic arbitrageurs:

Proposition 26 *None of the arbitrageurs are forced to liquidate in equilibrium.*

This result is rather intuitive. It shows that arbitrageurs reduce their initial investments such that liquidation does not happen in equilibrium. This is because liquidation is rather costly. As there is no uncertainty in the model, no strategic agent wants to buy an asset that she has to sell later with certainty, since buying an asset pushes its price up while selling decreases its price, both working against the profit of this kind of round-trip transaction. It implies that liquidation does not happen in equilibrium, but the threat of liquidation is still present on the off-equilibrium path.

Proposition 27 *Based on the initial capital of traders, the effect of the wealth constraint on arbitrageur activity can be divided into four cases.*

(I) There exists a constant $\Omega > 0$ *such that for* $M^1, M^2 \geq \frac{1}{\lambda}\Omega\overline{g}^2$*, arbitrageur strategies and the gap path are the same as in the unconstrained case, discussed in Proposition 15.*

(II) There exists a function $\Phi(.)$ *such that when* $0 < M^2 < \frac{1}{\lambda}\Omega\overline{g}^2$ *and* $M^1 \geq \Phi(M^2)$*, arbitrageur* 2 *is the cautious trader, and the constraint* (2.13) *binds for her. As a result, she trades less at date* 0 *than in the unconstrained case.*

(III) Similarly, when $0 < M^1 < \frac{1}{\lambda}\Omega\overline{g}^2$ *and* $M^2 \geq \Phi(M^1)$*, arbitrageur* 1 *becomes the cautious trader, and the constraint* (2.13) *binds for her. She trades less at date* 0 *than in the unconstrained case.*

(IV) When both arbitrageurs have low level of capital, $M^1, M^2 < \frac{1}{\lambda}\Omega\overline{g}^2$*, and they are close to each other such that* $M^1 < \Phi(M^2)$ *and* $M^2 < \Phi(M^1)$*, both arbitrageurs invest less than in the unconstrained case, and hence the gap remains larger.*

The four regions for cases (I)-(IV) are illustrated on Figure 2-6. Arbitrageurs remain solvent in all cases. Moreover, $\Phi(.)$ *is positive, strictly increasing, satisfies* $\Phi(x) > x$ *for* $0 < x < \frac{1}{\lambda}\Omega\overline{g}^2$*, and* $\Phi(x) = x$ *for* $x = 0$ *or* $x = \frac{1}{\lambda}\Omega\overline{g}^2$*.*

Proposition 27 describes the initial trades as a function of arbitrage capital. First, if both arbitrageurs start with sufficiently high level of capital, the wealth constraint does not

though those cases were not discussed in Section 2.4.

affect their trades and hence the dynamic equilibrium of the model is exactly the same as in unconstrained case, described in Proposition 15. As arbitrageurs have a lot of cash on hand that can provide a cushion against very large adverse movements in the gap, they race to the market and take large bets on the convergence of prices. Traders' positions and the evolution of the gap are illustrated on Figure 2-2.

When at least one arbitrageur has a low capital level to start with, while the other has (relatively) more, e.g. $M^1 \geq \Phi\left(M^2\right)$, the wealth constraint affects the date 0 trading through affecting agent 2's willingness to invest. Arbitrageur 2 must short less compared to the case when the constraint is not effective, $x_0^2 < x_{0,u}^i$, because she wants to avoid liquidation later on. In fact, she takes such a small position that it is not worth for arbitrageur 1 to push her to insolvency. On the other hand, arbitrageur 1 can invest more, $x_0^1 > x_{0,u}^i$, as long as she has a lower proportion in the gap asset. This is rather profitable for her, as the threat of potential liquidation restricts the ability of arbitrageur 2 to provide liquidity to local traders, and agent 1 has almost monopoly power in doing so. The large position that arbitrageur 1 takes compensates for the small holdings by arbitrageur 2 so the date 0 gap is not very different from the case when both arbitrageurs have high level of capital. This is illustrated on Figure 2-7.

Finally, when both arbitrageurs have low capital level to start with, and they are close to each other so that $M^1 < \Phi\left(M^2\right)$ and $M^2 < \Phi\left(M^1\right)$, the wealth constraint is important for both arbitrageurs. In particular, suppose that the proportion of wealth invested in the arbitrage opportunity is fixed for both agents, and it is larger for arbitrageur 2, i.e. $M^1/x_0^1 > M^2/x_0^2$. It implies that agent 2 is the cautious arbitrageur and faces a tighter wealth constraint, so she must reduce her holdings if she wants to avoid forced liquidation. However, by investing less she decreases her proportion of wealth in the arbitrage opportunity to below that of arbitrageur 1, that is she makes $M^1/x_0^1 < M^2/x_0^2$. Now arbitrageur 1 becomes the cautious arbitrageur, she is more prone to predatory risk, so she should reduce her initial investment. This drives the M^1/x_0^1 ratio above M^2/x_0^2, and so on. In the end, both arbitrageurs trade very little at date 0, $x_0^1, x_0^2 \ll x_{0,u}^i$, and the gap level remains high, $g_0 \gg g_{0,u}$. Given that both of them remain solvent, in the next period they both race to the arbitrage opportunity, and the gap quickly shrinks. This is illustrated on Figure 2-8.

2.6 Final remarks

This chapter presents an equilibrium model of endogenous predation and forced liquidation among strategic arbitrageurs who are subject to capital constraints. Arbitrageurs bet on the convergence of prices of two assets, but when prices actually diverge and their marked-to-market portfolio value becomes negative, traders have to unwind their risky holdings immediately and leave the market. This implies that arbitrageurs' wealth limits the positions they can take as long as they do not want to violate the constraint. Strategic traders may trigger the bankruptcy of 'weaker' agents, which creates predatory risk and implies that even if the investment opportunity is a fundamentally riskless arbitrage, traders might be reluctant to invest in it.

First I study a model when agents are already endowed with positions in the risky assets. I show that when traders have similar proportion of wealth invested in the arbitrage opportunity, they behave cooperatively, and prices converge through time, as in a benchmark model without the constraint. However, if there is a significant difference in their proportion of wealth invested, the arbitrageur with lower proportion invested in the arbitrage opportunity predates on the other trader by manipulating the price and forcing her to unwind her position at a large discount.

Then I examine whether a strategic trader is willing to build up a portfolio if it makes her prone to predation and hence large losses. I show that in the equilibrium of the full model liquidation never happens, but the threat of predation makes arbitrageurs reluctant to invest much in the arbitrage opportunity because of the presence of other arbitrageurs. In particular, the wealth constraint seriously affects the gap between the asset prices when arbitrageurs have similarly low level of capital, and implies that instead of racing to the opportunity arbitrageurs stay out, and the gap decreases gradually.

In the model presented here there is no informational asymmetry about the opportunity among arbitrageurs, and prices and positions are always deterministic. Naturally, this provides an opportunity to extend the framework in several dimensions. For example, it would be interesting to allow for asymmetric positions in the two risky assets and see what effect it would have if some strategic traders only had information about one leg of the trades of other arbitrageurs, as anecdotal evidence recalls about the trading counterparties of LTCM.

Moreover, it would be important to evaluate the empirical significance of the presented mechanism and to distinguish it from others that result in similarly slow trading of large traders, e.g. Kyle (1985). These are left for future work.

2.7 Appendix

2.7.1 Optimal trading of the monopoly

First I solve the problem without the wealth constraint.

Proof of Proposition 13. The arbitrageur's optimization program is given by

$$\max_{\{x_t\}_{t=0}^{2}} W_3 = M + \sum_{t=0}^{2} g_t(x_t)(x_t - x_{t-1}) \tag{2.15}$$

where $g_t = g_{t-1} - \lambda(x_t - x_{t-1})$ for $t = 0, 1, 2$ and $g_{-1} \equiv \overline{g}$. Writing it as a dynamic program it becomes

$$\max_{x_2} W_3 = M_1 + g_2(x_2)(x_2 - x_1),$$

and the FOC yields $0 = g_2 + \frac{dg_2}{dx_2}(x_2 - x_1) = g_1 - 2\lambda(x_2 - x_1)$, i.e.

$$x_2 - x_1 = \frac{1}{2\lambda} g_1 \text{ and } g_2 = \frac{1}{2} g_1. \tag{2.16}$$

Moreover, $W_3 = M_1 + g_2(x_2)(x_2 - x_1) = M_1 + \frac{1}{4\lambda} g_1^2$. Going back one more period the optimization program becomes

$$\begin{aligned} \max_{x_1} W_3 &= M_1 + \frac{1}{4\lambda} g_1^2 = M_0 + g_1(x_1 - x_0) + \frac{1}{4\lambda} g_1^2 \\ &= M_0 + (g_0 - \lambda(x_1 - x_0))(x_1 - x_0) + \frac{1}{4\lambda}(g_0 - \lambda(x_1 - x_0))^2, \end{aligned}$$

and the FOC yields $0 = g_0 - 2\lambda(x_1 - x_0) - \lambda\frac{1}{2\lambda}(g_0 - \lambda(x_1 - x_0))$, or $x_1 - x_0 = \frac{1}{3\lambda} g_0$. Therefore, $g_1 = \frac{2}{3} g_0$ and $W_3 = M_0 + \frac{1}{3\lambda} g_0^2$. Going back to the date 0 optimization it becomes

$$\begin{aligned} \max_{x_0} W_3 &= M_0 + \frac{1}{3\lambda} g_0^2 = M + g_0 x_0 + \frac{1}{3\lambda} g_0^2 \\ &= M + (\overline{g} - \lambda x_0) x_0 + \frac{1}{3\lambda}(\overline{g} - \lambda x_0)^2, \end{aligned}$$

so the FOC yields $0 = \overline{g} - 2\lambda x_0 - \frac{2}{3}(\overline{g} - \lambda x_0)$, or $x_0 = \frac{1}{4\lambda}\overline{g}$. Therefore, $g_0 = \frac{3}{4}\overline{g}$, which implies that the monopoly gradually provides liquidity in the local markets:

$$x_{t,u} - x_{t-1,u} = \frac{1}{4\lambda}\overline{g} \text{ for } t = 0, 1 \text{ and } 2,$$

that is

$$x_{0,u} = \frac{1}{4\lambda}\overline{g},\ x_{1,u} = \frac{1}{2\lambda}\overline{g},\ \text{and}\ x_{2,u} = \frac{3}{4\lambda}\overline{g},$$

and the gap decreases linearly over time:

$$g_{0,u} = \frac{3}{4}\overline{g},\ g_{1,u} = \frac{1}{2}\overline{g},\ \text{and}\ g_{2,u} = \frac{1}{4}\overline{g},$$

where the subscript u refers to the arbitrageur being unconstrained. ■

Proof of Proposition 14. For the full consideration of the effect of the constraint on the optimal portfolio choice of the monopolistic arbitrageur, one should analyze the 2-period subgame in which the constraint affects the optimal trades, given the gap g_0 and the positions she has after trading at date 0, M_0 and x_0, and then consider the portfolio choice problem at date 0. However, on the unconstrained equilibrium path the gap converges to zero, i.e. $g_1 - g_0 < 0$, and (2.8) is always satisfied. Given that arbitrageur can never achieve higher utility in the constrained portfolio choice problem than in the unconstrained problem, but the unconstrained optimum is feasible when incorporating the constraint, it does not affect the equilibrium holdings and gap for a monopolistic arbitrageur. ■

2.7.2 Optimal trading of the duopoly

Following the same footsteps as with the monopoly, first I solve the problem without the wealth constraint.

Proof of Proposition 15. Arbitrageur i's optimization program, $i = 1, 2$, is given by

$$\max_{\{x_t^i\}_{t=0}^2} W_3^i = M^i + \sum_{t=0}^{2} g_t\left(x_t^i\right)\left(x_t^i - x_{t-1}^i\right), \tag{2.17}$$

where $g_t = g_{t-1} - \lambda\left(x_t^i - x_{t-1}^i\right) - \lambda\left(x_t^{-i} - x_{t-1}^{-i}\right)$ for $t = 0, 1, 2$ and $g_{-1} \equiv \overline{g}$. In period 2 it is given by

$$\max_{x_2^i} W_3^i = M_1^i + g_2\left(x_2^i\right)\left(x_2^i - x_1^i\right) = M_1^i + \left(g_1 - \lambda\left(x_2^i - x_1^i\right) - \lambda\left(x_2^{-i} - x_1^{-i}\right)\right)\left(x_2^i - x_1^i\right),$$

and the FOCs yield

$$x_2^i - x_1^i = x_2^{-i} - x_1^{-i} = \frac{1}{3\lambda} g_1 \text{ and } g_2 = \frac{1}{3} g_1. \tag{2.18}$$

Moreover, $W_3^i = M_1^i + g_2\left(x_2^i - x_1^i\right) = M_1^i + \frac{1}{9\lambda} g_1^2$. Going back one more period the optimization problem becomes

$$\begin{aligned}
\max_{x_1^i} W_3^i &= M_1^i + \frac{1}{9\lambda} g_1^2 \\
&= M_0^i + \left(g_0 - \lambda\left(x_1^i - x_0^i\right) - \lambda\left(x_1^{-i} - x_0^{-i}\right)\right)\left(x_1^i - x_0^i\right) \\
&\quad + \frac{1}{9\lambda}\left(g_0 - \lambda\left(x_1^i - x_0^i\right) - \lambda\left(x_1^{-i} - x_0^{-i}\right)\right)^2,
\end{aligned}$$

and the FOC yields $x_1^i - x_0^i = x_1^{-i} - x_0^{-i} = \frac{7}{23\lambda} g_0$. Therefore, $g_1 = \frac{9}{23} g_0$, and $W_3^i = M_0^i + \frac{72}{23^2\lambda} g_0^2$. Going back to the date 0 optimization, it becomes

$$\max_{x_0^i} W_3^i = M_0^i + \frac{72}{23^2\lambda} g_0^2 = M^i + \left(\overline{g} - \lambda x_0^i - \lambda x_0^{-i}\right) x_0^i + \frac{72}{23^2\lambda}\left(\overline{g} - \lambda x_0^i - \lambda x_0^{-i}\right)^2,$$

so the FOC yields $x_0^i = x_0^{-i} = \frac{385}{1299\lambda}\overline{g}$. Therefore, $g_0 = \frac{529}{1299}\overline{g}$, which implies that the duopoly gradually provides liquidity in the local markets: $x_0^i = \frac{385}{1299\lambda}\overline{g}$, $x_1^i = \frac{182}{433\lambda}\overline{g}$, and $x_2^i = \frac{205}{433\lambda}\overline{g}$, for $i = 1, 2$,and the gap's evolution is given by $g_0 = \frac{529}{1299}\overline{g}$, $g_1 = \frac{69}{433}\overline{g}$, and $g_2 = \frac{23}{433}\overline{g}$. ■

Next I consider the case with the wealth constraint. First I characterize the optimal trades in period 2 conditional on the status of the two agents, and derive the value functions given the states and positions.

Suppose that the positions before trade happens at date 2 are M_1^i and x_1^i, and M_1^{-i} and x_1^{-i}, for agents i and $-i$ in the riskless and the risky assets, respectively. As a reminder, subscript $\{t, jk\}$ refers to time period t and status of traders i and $-i$, respectively, where the date 2 state can take two values: $j, k \in \{s, l\}$, i.e. solvency or liquidation, and it corresponds to whether the agents satisfied or violated the wealth constraint. For example, $g_{2,sl}\left(x_1^i, x_1^{-i}\right)$ denotes the gap in period 2 as a function of the position x_1^i of the arbitrageur who remains solvent and the position x_1^{-i} of the arbitrageur who is liquidated.

The following propositions state the optimal trades, equilibrium gaps and value functions in three possible cases at date 2. First, I restate a previous result without proof for the ss state, then I solve for the equilibria of the sl and ll states.

Proposition 28 *In period 2, conditional on both agents being solvent, the first-best trade orders and equilibrium gap are given by*

$$x^i_{2,ss} - x^i_{1,ss} = \frac{1}{3\lambda} g_1 \text{ for } i = 1,2, \text{ and } g_{2,ss} = \frac{1}{3} g_1. \tag{2.19}$$

Proof. Straightforward from (2.18). It also results in a continuation value function of $V_{1,ss}\left(M^i_1, x^i_1, M^{-i}_1, x^{-i}_1\right) = M^i_1 + \frac{1}{9\lambda} g^2_1$. ■

Proposition 29 *In period 2, conditional on agent i being solvent and $-i$ being liquidated, the first-best trade order and the equilibrium gap are given by*

$$x^i_{2,sl}\left(x^i_1, x^{-i}_1\right) = x^i_1 + \frac{1}{2\lambda}\left(g_1 + \lambda x^{-i}_1\right),\ x^i_{2,ls} = 0 \text{ and } g_{2,sl}\left(x^i_1, x^{-i}_1\right) = \frac{1}{2}\left(g_1 + \lambda x^{-i}_1\right). \tag{2.20}$$

Proof. The optimization problem of agent i is the same as in the ss case as she remains solvent, which yields the same FOC

$$0 = g_1 - 2\lambda\left(x^i_2 - x^i_1\right) - \lambda\left(x^{-i}_2 - x^{-i}_1\right). \tag{2.21}$$

As agent $-i$ has to close her position, $x^{-i}_{2,ls} = 0$. Substituting into (2.21), it becomes

$$x^i_{2,sl} = x^i_1 + \frac{1}{2\lambda}\left(g_1 + \lambda x^{-i}_1\right),$$

and the gap is

$$g_{2,sl}\left(x^i_1, x^{-i}_1\right) = \frac{1}{2}\left(g_1 + \lambda x^{-i}_1\right).$$

The value functions for the two agents are $V_{1,sl}\left(M^i_1, x^i_1, M^{-i}_1, x^{-i}_1\right) = M^i_1 + \frac{1}{4\lambda}\left(g_1 + \lambda x^{-i}_1\right)^2$, and $V_{1,ls}\left(M^i_1, x^i_1, M^{-i}_1, x^{-i}_1\right) = M^i_1 - \frac{1}{2}\left(g_1 + \lambda x^i_1\right) x^i_1$. ■

Proposition 30 *In period 2, conditional on both agents being liquidated, the trade orders and the equilibrium gap are given by*

$$x^i_{2,ll} = 0 \text{ and } g_{2,ll} = g_1 + \lambda\left(x^i_1 + x^{-i}_1\right). \tag{2.22}$$

The value functions are $V_{1,ll}\left(M^i_1, x^i_1, M^{-i}_1, x^{-i}_1\right) = M^i_1 - \left(g_1 + \lambda\left(x^i_1 + x^{-i}_1\right)\right) x^i_1$, $i = 1,2$.

Suppose now that $M^a/x_0^a > M^c/x_0^c$, hence we can define

$$\overline{x} \equiv x_0^a + x_0^c - \frac{1}{\lambda}\frac{M^c}{x_0^c} > \underline{x} \equiv x_0^a + x_0^c - \frac{1}{\lambda}\frac{M^a}{x_0^a}$$

as the thresholds on trades at date 1 that change the state of the world for arbitrageurs. It implies that at date 1 arbitrageur i faces the following optimization problem:

$$\begin{aligned}\max_x M_0^i &+ g_1(x)\left(x - x_0^i\right)\\ &+ 1_{\frac{M^i}{x_0^i}, \frac{M^{-i}}{x_0^{-i}} \geq g_1(x) - g_0} \frac{1}{9\lambda} g_1(x)^2 + 1_{\frac{M^i}{x_0^i} \geq g_1(x) - g_0 > \frac{M^{-i}}{x_0^{-i}}} \frac{1}{4\lambda}\left(g_1(x) + \lambda x_1^{-i}\right)^2\\ &- 1_{\frac{M^{-i}}{x_0^{-i}} \geq g_1(x) - g_0 > \frac{M^i}{x_0^i}} \frac{1}{2}\left(g_1(x) + \lambda x\right)x - 1_{\frac{M^i}{x_0^i}, \frac{M^{-i}}{x_0^{-i}} < g_1(x) - g_0}\left(g_1(x) + \lambda\left(x + x_1^{-i}\right)\right)x,\end{aligned}$$

where I have combined the continuation values for the four states of the world, given above. From here it is easy to show that the FOCs become

$$\begin{aligned}0 &= g_1 - \lambda\left(x_1^a - x_0^a\right) - \frac{2}{9}g_1 \text{ if } x_1^a > \overline{x} - x_1^c,\\ 0 &\geq g_1 - \lambda\left(x_1^a - x_0^a\right) - \frac{2}{9}g_1 \text{ if } x_1^a = \overline{x} - x_1^c,\\ 0 &= g_1 - \lambda\left(x_1^a - x_0^a\right) - \frac{1}{2}\left(g_1 + \lambda x_1^c\right) \text{ if } \overline{x} - x_1^c > x_1^a > \underline{x} - x_1^c,\\ 0 &\geq g_1 - \lambda\left(x_1^a - x_0^a\right) - \frac{1}{2}\left(g_1 + \lambda x_1^c\right) \text{ if } x_1^a = \underline{x} - x_1^c,\\ 0 &= g_1 - \lambda\left(x_1^a - x_0^a\right) - \left(g_1 + \lambda\left(x_1^a + x_1^c\right)\right) \text{ if } x_1^a < \underline{x} - x_1^c\end{aligned}$$

for the aggressive trader and

$$\begin{aligned}0 &= g_1 - \lambda\left(x_1^c - x_0^c\right) - \frac{2}{9}g_1 \text{ if } x_1^c > \overline{x} - x_1^a,\\ 0 &\geq g_1 - \lambda\left(x_1^c - x_0^c\right) - \frac{2}{9}g_1 \text{ if } x_1^c = \overline{x} - x_1^a,\\ 0 &= g_1 - \lambda\left(x_1^c - x_0^c\right) - \frac{1}{2}\left(g_1 + \lambda x_1^c\right) \text{ if } \overline{x} - x_1^c > x_1^c > \underline{x} - x_1^a,\\ 0 &\geq g_1 - \lambda\left(x_1^c - x_0^c\right) - \frac{1}{2}\left(g_1 + \lambda x_1^c\right) \text{ if } x_1^c = \underline{x} - x_1^a,\\ 0 &= g_1 - \lambda\left(x_1^c - x_0^c\right) - \left(g_1 + \lambda\left(x_1^c + x_1^a\right)\right) \text{ if } x_1^c < \underline{x} - x_1^a\end{aligned}$$

for the cautious trader. Combining these gives the following cases:

- Suppose $x_1^a + x_1^c > \overline{x}$, then the FOCs yield

$$x_1^a - x_0^a = x_1^c - x_0^c = \frac{7}{23\lambda} g_0.$$

The condition $x_1^a + x_1^c \geq \overline{x}$ is equivalent to

$$M^c \geq -\frac{14}{23} g_0 x_0^c,$$

which always holds.

- Suppose$x_1^a + x_1^c = \overline{x}$, hence the FOCs simplify to

$$x_1^a - x_0^a \geq \frac{7}{9\lambda}\left(g_0 + \frac{M^c}{x_0^c}\right) \text{ and } x_1^c - x_0^c \geq \frac{7}{9\lambda}\left(g_0 + \frac{M^c}{x_0^c}\right).$$

It implies that

$$g_1 = g_0 - \lambda\left[x_1^a + x_1^c - (x_0^a + x_0^c)\right] \leq -\left(\frac{5}{9} g_0 + \frac{14}{9}\frac{M^c}{x_0^c}\right) < 0$$

which cannot happen if $M^c / x_0^c > 0$.

- Suppose we have $\overline{x} > x_1^a + x_1^c > \underline{x}$, which implies that

$$x_1^a - x_0^a = x_1^c - x_0^c = \frac{1}{5\lambda}\left(g_0 - \lambda x_0^c\right),$$

hence

$$g_1 = g_0 - \frac{2}{5}\left(g_0 - \lambda x_0^c\right),$$

and it must be that

$$0 < \frac{M^c}{x_0^c} < -\frac{2}{5}\left(g_0 - \lambda x_0^c\right) \leq \frac{M^a}{x_0^a}.$$

- Suppose $x_1^a + x_1^c = \underline{x}$, then the FOCs become

$$x_1^a + \frac{1}{2} x_1^c \geq \frac{1}{2\lambda}\left(g_0 + \frac{M^a}{x_0^a}\right) + x_0^a$$

and

$$x_1^c \geq \frac{1}{3\lambda}\left(g_0 + \frac{M^a}{x_0^a}\right) + \frac{2}{3}x_0^c.$$

Also it must be that $x_1^a + x_1^c = \underline{x} = x_0^a + x_0^c - \frac{1}{\lambda}\frac{M^a}{x_0^a}$.

- Finally, suppose $x_1^a + x_1^c < \underline{x}$, then we have

$$x_1^a - x_0^a = x_1^c - x_0^c = -\frac{1}{3}\left(x_0^a + x_0^c\right),$$

and it requires $x_1^a + x_1^c < \underline{x}$, i.e.

$$0 < \frac{M^c}{x_0^c} < \frac{M^a}{x_0^a} < \frac{2}{3}\lambda\left(x_0^a + x_0^c\right).$$

These conditions describe the locally optimal trades and the trivial constraints on the proportion of wealth invested, presented in Propositions 20-22. What is left is to check whether they are globally optimal too, i.e. whether any arbitrageur wants so deviate while changing the state too.

Optimal trading conditional on getting to state ss

There is no candidate equilibrium with the constraint binding for arbitrageur c. Besides the above requirements on arbitrageur capital, in an equilibrium with not binding constraints it must also be checked whether arbitrageur a would like to deviate and trigger the bankruptcy of arbitrageur c. This is because agent c's position in the first best solution is not equivalent to full liquidation, thus it might be profitable for agent a to trigger agent c's bankruptcy. For this, suppose that arbitrageur a deviates so that she remains solvent but arbitrageur c is liquidated. She is better off if and only if

$$\begin{aligned}&V_{1,sl}\left(M_0^a + g_1\left(x_1^a - x_0^a\right), x_1^a, M_0^c + g_1\left(x_{1,nn}^c - x_0^c\right), x_{1,nn}^c\right)\\&\qquad > V_{1,ss}\left(M_0^a + g_{1,nn}\left(x_{1,nn}^a - x_0^a\right), x_{1,nn}^a, M_0^c + g_{1,nn}\left(x_{1,nn}^c - x_0^c\right), x_{1,nn}^c\right),\end{aligned}$$

that is if her utility from getting into state sl with positions $M_0^a + g_1\left(x_1^a - x_0^a\right)$ and x_1^a in the riskless and the risky assets, respectively, is higher than the utility she derives in state ss from holding positions $M_0^a + g_{1,nn}\left(x_{1,nn}^a - x_0^a\right)$ and $x_{1,nn}^a$ (which are the locally optimal holdings in that state), while assuming that arbitrageur c stays with the equilibrium holding $x_{1,nn}^c$. After some algebra, she is better off deviating iff $x_1^a \in \left(\overline{x}_{nn,nv}^a - \delta_{nn,nv}^a, \overline{x}_{nn,nv}^a + \delta_{nn,nv}^a\right)$, where

$$\overline{x}_{nn,nv}^a = \frac{3}{23\lambda}g_0 + x_0^a + \frac{1}{3}x_0^c,$$

and

$$\delta_{nn,nv}^a = \frac{2}{3\lambda}\sqrt{\left(\frac{15-2\sqrt{6}}{23}g_0 + \lambda x_0^c\right)\left(\frac{15+2\sqrt{6}}{23}g_0 + \lambda x_0^c\right)}.$$

Given $g_0 \geq 0$ and the assumption $x_0^c > 0$, the discriminant is non-negative and $\delta_{nn,nv}^a$ exists. As arbitrageur 1 can only push arbitrageur 2 into liquidation by increasing the gap while making sure she does not get liquidated, i.e. by choosing a trade

$$-\frac{1}{\lambda}M^a/x_0^a - \left(x_{1,nn}^c - x_0^c\right) \leq x_1^a - x_0^a < -\frac{1}{\lambda}M^c/x_0^c - \left(x_{1,nn}^c - x_0^c\right),$$

her deviation can increase her utility if and only if both $-\frac{1}{\lambda}M^a/x_0^a - \left(x_{1,nn}^c - x_0^c\right) < \overline{x}_{nn,nv}^a - x_0^a + \delta_{nn,nv}^a$ and $\overline{x}_{nn,nv}^a - x_0^a - \delta_{nn,nv}^a < -\frac{1}{\lambda}M^c/x_0^c - \left(x_{1,nn}^c - x_0^c\right)$ hold. Hence a simple reorganization of these inequalities implies that a necessary condition for the existence of the equilibrium is that either $M^a/x_0^a \leq -\frac{10}{23}g_0 + \frac{1}{3}\lambda x_0^c - \lambda\delta_{nn,nv}^a$ or $M^c/x_0^c \geq -\frac{10}{23}g_0 + \frac{1}{3}\lambda x_0^c + \lambda\delta_{nn,nv}^a$. As $M^a/x_0^a > 0$, it must be that $M^c/x_0^c \geq \Delta_{nn,nv}^a \equiv -\frac{10}{23}g_0 + \frac{1}{3}\lambda x_0^c + \lambda\delta_{nn,nv}^a$.

Optimal trading conditional on getting to state sl

To check whether an equilibrium with arbitrageur a being solvent and arbitrageur c having to liquidate exists, it must be checked whether arbitrageur c would prefer to change her trading speed and remain solvent, whether she would prefer to force arbitrageur 1 to distress, or whether the constrained arbitrageur a would prefer to liquidate.

Equilibrium with non-binding constraint The possible deviations are when arbitrageur c forces arbitrageur a into distress, or when arbitrageur c rescues herself. As the constraint might not bind in equilibrium when arbitrageurs start with different positions, it must also be checked whether arbitrageur c wants to trigger the distress of arbitrageur a while rescuing herself.

Agent c forces agent a into liquidation. Arbitrageur c is better off forcing the liquidation or arbitrageur a iff

$$V_{1,ll}\left(M_0^c + g_1\left(x_1^c - x_0^c\right), x_1^c, M_0^a + g_1\left(x_{1,nv}^a - x_0^a\right), x_{1,nv}^a\right)$$
$$> V_{1,ls}\left(M_0^c + g_{1,nv}\left(x_{1,vn}^c - x_0^c\right), x_{1,vn}^c, M_0^a + g_1\left(x_{1,nv}^a - x_0^a\right), x_{1,nv}^a\right),$$

that is if her utility from getting into state ll with positions $M_0^c + g_1\left(x_1^c - x_0^c\right)$ and x_1^c in the riskless and the risky assets, respectively, is higher than the utility she derives in state ls from holding positions $M_0^c + g_{1,nv}\left(x_{1,vn}^c - x_0^c\right)$ and $x_{1,vn}^c$ (which are actually the optimal holdings in that state), while assuming that arbitrageur a stays with the equilibrium holding $x_{1,nv}^a$. After some algebra, she is better off deviating iff $x_1^c \in \left(\overline{x}_{nv,vv}^c - \delta_{nv,vv}^c, \overline{x}_{nv,vv}^c + \delta_{nv,vv}^c\right)$, where

$$\overline{x}_{nv,vv}^c = -\frac{1}{2}x_0^a + \frac{3}{5}x_0^c - \frac{1}{10\lambda}g_0,$$

and

$$\delta_{nv,vv}^c = \frac{1}{\sqrt{3}\lambda}\sqrt{\left(\frac{1}{10}g_0 - \frac{1}{2}\lambda x_0^a + \frac{7}{5}\lambda x_0^c\right)\left(-\frac{9}{10}g_0 - \frac{3}{2}\lambda x_0^a - \frac{3}{5}\lambda x_0^c\right)},$$

with the discriminant being negative (hence a deviation cannot increase her utility) iff

$$0 < x_0^a < \frac{1}{5\lambda}g_0 + \frac{14}{5}x_0^c.$$

Suppose that the discriminant is non-negative and hence $\delta_{nv,vv}^c$ exists.

Since $x_0^a > 0$, arbitrageur c can push arbitrageur a into liquidation by increasing the gap, i.e. by choosing a trade $x_1^c - x_0^c < -\frac{1}{\lambda}M^a/x_0^a - \left(x_{1,nv}^a - x_0^a\right)$. She can deviate while increasing her utility if and only if $-\frac{1}{\lambda}M^a/x_0^a - \left(x_{1,nv}^a - x_0^a\right) > \overline{x}_{nv,vv}^2 - x_0^c - \delta_{nv,vv}^c$. Hence a simple reorganization of this inequality implies that in equilibrium it must be that $M^a/x_0^a \geq -\frac{1}{10}g_0 + \frac{1}{2}\lambda x_0^a + \frac{3}{5}\lambda x_0^c + \lambda\delta_{nv,vv}^c$. Notice that as $x_0^c > 0$, arbitrageur c will remain distressed when pushing arbitrageur a into bankruptcy, and hence no other condition is needed.

Agent c rescues herself. Arbitrageur c is better off rescuing herself iff

$$V_{1,ss}\left(M_0^c + g_1\left(x_1^c - x_0^c\right), x_1^c, M_0^a + g_1\left(x_{1,nv}^a - x_0^a\right), x_{1,nv}^a\right)$$
$$> V_{1,ls}\left(M_0^c + g_{1,nv}\left(x_{1,vn}^c - x_0^c\right), x_{1,vn}^c, M_0^a + g_1\left(x_{1,nv}^a - x_0^a\right), x_{1,nv}^a\right),$$

that is if her utility from getting into state ss with positions $M_0^c + g_1\left(x_1^c - x_0^c\right)$ and x_1^c in the riskless and the risky assets, respectively, is higher than the utility she derives in state ls from holding positions $M_0^c + g_{1,nv}\left(x_{1,vn}^c - x_0^c\right)$ and $x_{1,vn}^c$ (which are actually the optimal holdings in that state), while assuming that arbitrageur a stays with the equilibrium holding $x_{1,nv}^a$. After some tedious algebra, she is better off deviating iff $x_1^c \in \left(\overline{x}_{nv,nn}^c - \delta_{nv,nn}^c, \overline{x}_{nv,nn}^c + \delta_{nv,nn}^c\right)$ with

$$\overline{x}_{nv,nn}^c = \frac{7}{20\lambda} g_0 + \frac{87}{80} x_0^c \text{ and}$$

$$\delta_{nv,nn}^c = \frac{3\sqrt{457}}{80\lambda}\sqrt{\left(\frac{228 + 20\sqrt{2}}{457} g_0 + \lambda x_0^c\right)\left(\frac{228 - 20\sqrt{2}}{457} g_0 + \lambda x_0^c\right)}.$$

Given $g_0 \geq 0$ and the assumption $x_0^c > 0$, the discriminant is non-negative and $\delta_{nv,nn}^2$ exists.

Since $x_0^a, x_0^c > 0$, arbitrageur c can rescue herself by shrinking the gap, i.e. by choosing a trade $x_1^c - x_0^c \geq -\frac{1}{\lambda} M^c / x_0^c - \left(x_{1,nv}^a - x_0^a\right)$. She can deviate while increasing her utility if and only if $-\frac{1}{\lambda} M^c / x_0^c - \left(x_{1,nv}^a - x_0^a\right) < \overline{x}_{nv,nn}^c - x_0^c + \delta_{nv,nn}^c$. Hence a simple reorganization of this inequality implies that in equilibrium it must be that $M^c / x_0^c \leq -\frac{11}{20} g_0 + \frac{9}{80} \lambda x_0^c - \lambda \delta_{nv,nn}^c$.

Summary for unconstrained sl equilibrium. Combining the initial constraints that are based on the equilibrium price and the above constraints regarding potential deviations yields that the unconstrained sl equilibrium exists if $0 < M^c / x_0^c \leq \Delta_{nv,nn}^c$ and $M^a / x_0^a \geq \Delta_{nv,vv}^c$, where $\Delta_{nv,nn}^c = -\frac{11}{20} g_0 + \frac{9}{80} \lambda x_0^c - \lambda \delta_{nv,nn}^c$ and

$$\Delta_{nv,vv}^c = \max\left\{-\frac{2}{5}\left(g_0 - \lambda x_0^2\right); -\frac{1}{10} g_0 + \frac{1}{2} \lambda x_0^a + \frac{3}{5} \lambda x_0^c + \lambda \delta_{nv,vv}^c\right\}.$$

Equilibrium with binding constraint The possible deviations from the equilibrium trades are when arbitrageur c either forces arbitrageur a into distress or rescues herself, and when arbitrageur a decides to liquidate.

Agent c forces agent a into liquidation. Arbitrageur c is better off forcing the liquidation or arbitrageur 1 iff

$$V_{1,ll}\left(M_0^c + g_1\left(x_1^c - x_0^c\right), x_1^c, M_0^a + g_1\left(x_{1,bv}^a - x_0^a\right), x_{1,bv}^a\right)$$
$$> V_{1,ls}\left(M_0^c + g_{1,bv}\left(x_{1,vb}^c - x_0^c\right), x_{1,vb}^c, M_0^a + g_1\left(x_{1,bv}^a - x_0^a\right), x_{1,bv}^a\right),$$

that is if her utility from getting into state ll with positions $M_0^c + g_1(x_1^c - x_0^c)$ and x_1^c in the riskless and the risky assets, respectively, is higher than the utility she derives in state ls from holding positions $M_0^c + g_{1,bv}\left(x_{1,vb}^c - x_0^c\right)$ and $x_{1,vb}^c$ (which are actually the optimal holdings in that state), while assuming that arbitrageur a stays with the equilibrium holding $x_{1,bv}^a$. After some algebra, she is better off deviating iff $x_1^c \in \left(\overline{x}_{bv,vv}^c - \delta_{bv,vv}^c, \overline{x}_{bv,vv}^c + \delta_{bv,vv}^c\right)$, where

$$\overline{x}_{bv,vv}^c = \frac{1}{6\lambda}g_0 - \frac{1}{2}x_0^a + \frac{1}{3}x_0^c + \frac{2}{3\lambda}\frac{M^a}{x_0^a}, \text{ and}$$

$$\delta_{bv,vv}^c = \frac{1}{\sqrt{3}\lambda}\sqrt{\left(\frac{M^a}{x_0^a} + \frac{1}{2}g_0 - \frac{1}{2}\lambda x_0^a + \lambda x_0^c\right)\left(\frac{M^a}{x_0^a} - \left(\frac{1}{2}g_0 + \frac{3}{2}\lambda x_0^a + \lambda x_0^c\right)\right)},$$

with the discriminant being negative (hence a deviation cannot increase her utility) iff

$$-\frac{1}{2}g_0 + \frac{1}{2}\lambda x_0^a - \lambda x_0^c < \frac{M^a}{x_0^a} < \frac{1}{2}g_0 + \frac{3}{2}\lambda x_0^a + \lambda x_0^c.$$

Suppose now that the discriminant is non-negative and hence $\delta_{bv,vv}^c$ exists. As $x_0^a, x_0^a > 0$, arbitrageur c can push arbitrageur a into liquidation by increasing the gap, i.e. by choosing a trade $x_1^c < x_{1,vb}^c$. She can deviate while increasing her utility if and only if $x_{1,vb}^c > \overline{x}_{bv,vv}^c - \delta_{bv,vv}^c$. A simple reorganization of this inequality implies that a necessary condition for the equilibrium is $M^a/x_0^a + g_0 + 2\lambda x_0^c \leq 0$, which cannot happen. Therefore the equilibrium can only exist if $-\frac{1}{2}g_0 + \frac{1}{2}\lambda x_0^a - \lambda x_0^c < M^a/x_0^a < \frac{1}{2}g_0 + \frac{3}{2}\lambda x_0^a + \lambda x_0^c$.

Agent a decides to liquidate. She is better off triggering her own liquidation iff

$$V_{1,ll}\left(M_0^a + g_1\left(x_1^a - x_0^a\right), x_1^a, M_0^c + g_{1,bv}\left(x_{1,vb}^c - x_0^c\right), x_{1,vb}^c\right)$$
$$> V_{1,sl}\left(M_0^a + g_1\left(x_{1,bv}^a - x_0^a\right), x_{1,bv}^a, M_0^c + g_{1,bv}\left(x_{1,vb}^c - x_0^c\right), x_{1,vb}^c\right),$$

that is if her utility from getting into state ll with positions $M_0^a + g_1\left(x_1^a - x_0^a\right)$ and x_1^a in the riskless and the risky assets, respectively, is higher than the utility she derives in state ls from holding positions $M_0^a + g_1\left(x_{1,bv}^a - x_0^a\right)$ and $x_{1,bv}^a$ (which are the optimal holdings in that state), while assuming that arbitrageur c stays with the equilibrium holding $x_{1,vb}^c$. After some algebra, she is better off deviating iff $x_1^a \in \left(\overline{x}_{bv,vv}^a - \delta_{bv,vv}^a, \overline{x}_{bv,vv}^a + \delta_{bv,vv}^a\right)$, where

$$\overline{x}_{bv,vv}^a = -\frac{1}{2}\left(\frac{1}{3\lambda}\frac{M^a}{x_0^a} + \frac{1}{3\lambda}g_0 - x_0^a + \frac{2}{3}x_0^c\right), \text{ and}$$

$$\delta_{bv,vv}^a = \frac{1}{2\lambda}\sqrt{\left(\frac{M^a}{x_0^a} + g_0 + \lambda x_0^a\right)\left(\frac{11}{3}\frac{M^a}{x_0^a} - \frac{1}{3}g_0 - 3\lambda x_0^a - \frac{8}{3}\lambda x_0^c\right)},$$

with the discriminant being negative (hence a deviation cannot increase her utility) iff

$$0 < \frac{M^a}{x_0^a} < \frac{1}{11}g_0 + \frac{9}{11}\lambda x_0^a + \frac{8}{11}\lambda x_0^c.$$

Suppose that the discriminant is non-negative and hence $\delta_{bv,vv}^a$ exists.

As $x_0^a, x_0^c > 0$, arbitrageur a can trigger her own liquidation by increasing the gap, i.e. by choosing a trade $x_1^a < x_{1,bv}^a$. She can deviate while increasing her utility if and only if $x_{1,bv}^a > \overline{x}_{bv,vv}^a - \delta_{bv,vv}^a$. A simple reorganization of this inequality implies that a necessary condition for the equilibrium is thus either $0 < M^a/x_0^a < \frac{1}{11}g_0 + \frac{9}{11}\lambda x_0^a + \frac{8}{11}\lambda x_0^c$ or $\frac{1}{11}g_0 + \frac{9}{11}\lambda x_0^a + \frac{8}{11}\lambda x_0^c \leq M^a/x_0^a \leq -\frac{1}{7}g_0 + \frac{3}{7}\lambda x_0^a + \frac{4}{7}\lambda x_0^c$. As this latter would also imply $3g_0 + 5\lambda x_0^a + 2\lambda x_0^c \leq 0$, which never holds, therefore the equilibrium only exists if $0 < M^a/x_0^a < \frac{1}{11}g_0 + \frac{9}{11}\lambda x_0^a + \frac{8}{11}\lambda x_0^c$.

Agent c rescues herself. Arbitrageur c is better off rescuing herself iff

$$V_{1,ss}\left(M_0^c + g_1\left(x_1^c - x_0^c\right), x_1^c, M_0^a + g_1\left(x_{1,bv}^a - x_0^a\right), x_{1,bv}^a\right)$$
$$> V_{1,ls}\left(M_0^c + g_{1,bv}\left(x_{1,vb}^c - x_0^c\right), x_{1,vb}^c, M_0^a + g_1\left(x_{1,bv}^a - x_0^a\right), x_{1,bv}^a\right),$$

that is if her utility from getting into state ss with positions $M_0^c + g_1(x_1^c - x_0^c)$ and x_1^c in the riskless and the risky assets, respectively, is higher than the utility she derives in state ls from holding positions $M_0^c + g_{1,bv}\left(x_{1,vb}^c - x_0^c\right)$ and $x_{1,vb}^c$ (which are actually the optimal holdings in that state), while assuming that arbitrageur a stays with the equilibrium holding $x_{1,bv}^a$. After some tedious algebra, she is better off deviating iff $x_1^c \in \left(\overline{x}_{bv,nn}^c - \delta_{bv,nn}^c, \overline{x}_{bv,nn}^c + \delta_{bv,nn}^c\right)$ with

$$\overline{x}_{bv,nn}^c = \frac{7}{12\lambda}\left(g_0 + \frac{M^a}{x_0^a}\right) + \frac{41}{48}x_0^c \text{ and}$$

$$\delta_{bv,nn}^c = \frac{\sqrt{7}}{4\lambda}\sqrt{\left(\frac{M^a}{x_0^a} + g_0 + \frac{23+3\sqrt{2}}{14}\lambda x_0^c\right)\left(\frac{M^a}{x_0^a} + g_0 + \frac{23-3\sqrt{2}}{14}\lambda x_0^c\right)},$$

where $x_0^a, x_0^c > 0$ implies that the discriminant is non-negative and hence $\delta_{bv,nn}^2$ exists.

As $x_0^c > 0$, arbitrageur c can rescue herself by shrinking the gap, i.e. by choosing a trade $x_1^c - x_0^c \geq -\frac{1}{\lambda}M^c/x_0^c - \left(x_{1,bv}^a - x_0^a\right)$. She can deviate while increasing her utility if and only if $-\frac{1}{\lambda}M^c/x_0^c - \left(x_{1,bv}^a - x_0^a\right) < \overline{x}_{bv,nn}^c - x_0^c + \delta_{bv,nn}^c$. A simple reorganization of this inequality implies that a necessary condition for the equilibrium is $M^c/x_0^c \leq \frac{3}{4}M^a/x_0^a - \frac{1}{4}g_0 - \frac{3}{16}\lambda x_0^c - \lambda\delta_{bv,nn}^c$.

However, given $M^c/x_0^c > 0$, it should be that $\lambda\delta_{bv,nn}^c < \frac{3}{4}M^a/x_0^a - \frac{1}{4}g_0 - \frac{3}{16}\lambda x_0^c$. After substituting in for $\delta_{bv,nn}^c$ and using that $M^a/x_0^a > 0$ as well, it can be shown that it cannot hold. Therefore agent c always rescues herself, and a constrained sl equilibrium thus never happens.

Optimal trading conditional on getting to state ll

As it is shown below, it is enough to consider the possible deviation when arbitrageur 1 rescues herself, as it already implies there is no equilibrium with both agents becoming distressed.

Arbitrageur a rescues herself. If, for example, it is shown that arbitrageur a deviates, there is no equilibrium with double liquidation at all. She is better off rescuing herself iff

$$\begin{aligned} &V_{1,sl}\left(M_0^a + g_1(x_1^a - x_0^a), x_1^a, M_0^c + g_1\left(x_{1,vv}^c - x_0^c\right), x_{1,vv}^c\right) \\ &\qquad > V_{1,ll}\left(M_0^a + g_{1,vv}\left(x_{1,v}^a - x_0^a\right), x_{1,vv}^a, M_0^c + g_{1,vv}\left(x_{1,vv}^c - x_0^c\right), x_{1,vv}^c\right), \end{aligned}$$

that is if her utility from getting into state sl with positions $M_0^a + g_1\left(x_1^a - x_0^a\right)$ and x_1^a in the riskless and the risky assets, respectively, is higher than the utility she derives in state ll from holding positions $M_0^a + g_{1,vv}\left(x_{1,v}^a - x_0^a\right)$ and $x_{1,vv}^a$ (which are actually the optimal holdings in that state), while assuming that arbitrageur 2 stays with the equilibrium holding $x_{1,vv}^c$. After some, she is better off deviating iff $x_1^a \in \left(\overline{x}_{vv,nv}^a - \delta_{vv,nv}^a, \overline{x}_{vv,nv}^a + \delta_{vv,nv}^a\right)$ with

$$\overline{x}_{vv,nv}^a = x_0^a + \frac{1}{3\lambda}\left(g_0 + \frac{2}{3}\lambda x_0^a - \frac{1}{3}\lambda x_0^c\right)$$

and

$$\delta_{vv,nv}^a = \frac{2}{3\lambda}\left|g_0 + \frac{5}{3}\lambda x_0^a + \frac{2}{3}\lambda x_0^c\right|.$$

As $x_0^a, x_0^c > 0$, agent a has to decrease the gap, i.e. buy more (or short less) to make sure $M^a/x_0^a \geq g_1 - g_0$, and for this she needs a trade of $x_1^a \geq -\frac{1}{\lambda}M^a/x_0^a - x_{1,vv}^c$. Combining with the other condition yields that she can deviate iff $-\frac{1}{\lambda}M^a/x_0^a - x_{1,vv}^c < \overline{x}_{vv,nv}^a + \delta_{vv,nv}^a$. However, as this constraint is equivalent to $0 < g_0 + \lambda x_0^a + M^a/x_0^a$, which always holds as all three components are non-negative, arbitrageur a always deviates and thus there is no equilibrium with both agents getting liquidated.

2.7.3 Trading under predatory threat

As mentioned in the main part of the chapter, the optimization programs of arbitrageurs with these constraints becomes difficult to solve in closed form (it includes solving 4th order equations). Thus, I provide some preliminary analysis in the following three Lemmas to decrease the potential set of equilibria, and then I solve the remaining problem numerically.

First, it is easy to see that:

Claim 31 *There exists no equilibrium without trading at date* 0*.*

This Lemma is rather intuitive. If, for example, arbitrageur 1 does not trade in period 0, arbitrageur 2 is better off investing a little into the arbitrage opportunity that staying out completely, as long as her trade satisfies (2.13). Given the assumption $M^2 > 0$, there exists a sufficiently small x_0^2 such that it is possible.

Similarly:

Claim 32 *There exists no equilibrium with only one arbitrageur trading at date* 0*.*

If, for example, arbitrageur 1 does trade in period 0, arbitrageur 2 can take an arbitrarily small position such that $M^2/x_0^2 > M^1/x_0^1$. It implies that she becomes the aggressive agent, and as there is no equilibrium in which both arbitrageurs are liquidated, she will never be liquidated. As there is no threat of predation on her, investing is strictly better than staying out, as it is a fundamentally riskless arbitrage opportunity.

Finally:

Claim 33 *There is no equilibrium with any agent having* $x_0^i < 0$*.*

First, an agent cannot have $x_0^i < 0$ and become liquidated later, because in this case not trading at date 0 would make her better off for two reasons: she can trade freely later; moreover, liquidation means closing positions that one has built up previously. Second, if and agent has $x_0^i < 0$ and she uses it to force the other trader to liquidation, arbitrageur $-i$ can decide to withdraw from trading in the first period, with which she stays solvent, moreover faces a better investment opportunity, because the effective gap for her is even larger than before.

Therefore, it must be that in equilibrium $x_0^1, x_0^2 > 0$. Suppose now that under their trades arbitrageurs end up in an unconstrained ss equilibrium. Going back to the date 0 optimization, arbitrageur i's optimization problem becomes

$$\max_{x_0^i} W_3^i = M_0^i + \frac{72}{23^2 \lambda} g_0^2 = M^i + \left(\overline{g} - \lambda x_0^i - \lambda x_0^{-i}\right) x_0^i + \frac{72}{23^2 \lambda} \left(\overline{g} - \lambda x_0^i - \lambda x_0^{-i}\right)^2, \quad (2.23)$$

so the FOC yields $x_0^1 = x_0^2 = (1/\lambda)\, \alpha \overline{g} > 0$ with $\alpha = \frac{385}{1299}$. Therefore, $g_0 = (1 - 2\alpha)\, \overline{g} = \frac{529}{1299} \overline{g}$, and the optimal holdings and the equilibrium gap are as in the model with no constrained, in Proposition 15. This yields a utility of $V\left(M^i\right) = M^i + g_0 x_0^i + \frac{72}{23^2 \lambda} g_0^2 = M^i + \frac{1}{\lambda} \Delta \overline{g}^2$ with $\Delta = (1 - 2\alpha)\, \alpha + \frac{72}{23^2} (1 - 2\alpha)^2$.

When is it an equilibrium in presence of the wealth constraint? First, it must satisfy the conditions derived in Appendix 2.7.2. The equilibrium candidate order and gap

$$x_0^2 = \frac{385}{1299 \lambda} \overline{g} \text{ and } g_0 = \frac{529}{1299} \overline{g}$$

satisfy

$$x_0^2 > \frac{2\left(\sqrt{233} - 15\right)}{23 \lambda} g_0,$$

hence it must be that

$$\frac{M^1}{x_0^1}, \frac{M^2}{x_0^2} \geq -\frac{10}{23} g_0 + \frac{1}{3} \lambda x_0^2 + \lambda \delta_{nn,nv}^1,$$

which is equivalent to

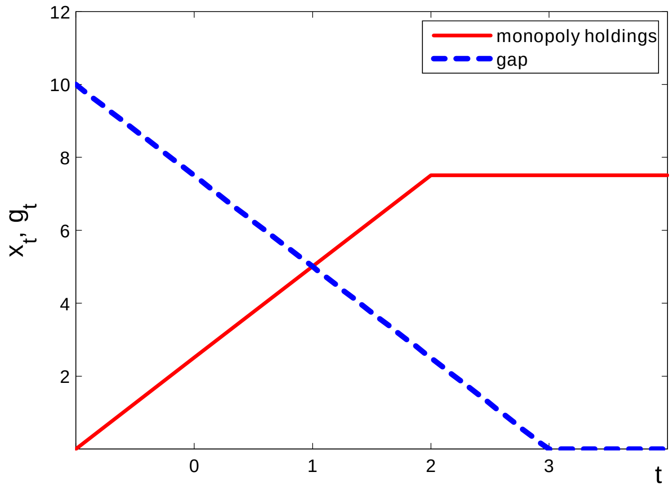
$$M^1, M^2 \geq \frac{1}{\lambda} \Omega \overline{g}^2$$

with

$$\begin{aligned} \Omega &\equiv \alpha \left[-\frac{10}{23} + \frac{83}{69} \alpha + \frac{2}{3} \sqrt{\left(\frac{15 - 2\sqrt{6}}{23} (1 - 2\alpha) + \alpha \right) \left(\frac{15 + 2\sqrt{6}}{23} (1 - 2\alpha) + \alpha \right)} \right] \\ &= \frac{385\left(4\sqrt{130051} - 305\right)}{3 * 1299^2} > 0. \end{aligned}$$

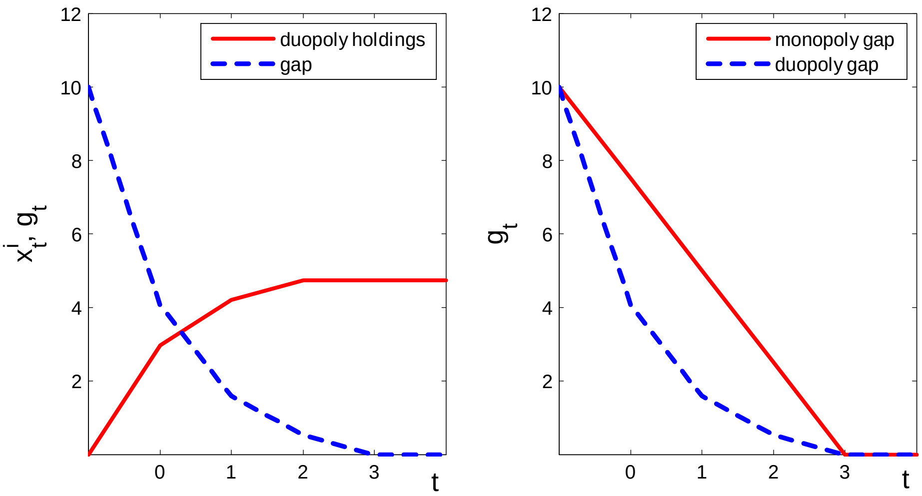
Second, it must be that none of the agents want to deviate from this profile and change the state. As having to liquidate puts a constraint on the strategy space of an arbitrageur, none of the arbitrageurs want to change the state in order to trigger her own distress. Therefore the only deviation one has to check is triggering the liquidation of the other arbitrageur. I confirm it numerically that as long as arbitrageurs trade such that they satisfy (2.13), it would be too costly for any trader to go long the gap and trigger the liquidation of the other arbitrageur. Therefore the equilibrium trades are those obtained from (2.23), subject to (2.13). This concludes the solution.

Figure 2-1: Equilibrium gap path and the optimal holdings of the monopoly in the arbitrage opportunity over time.



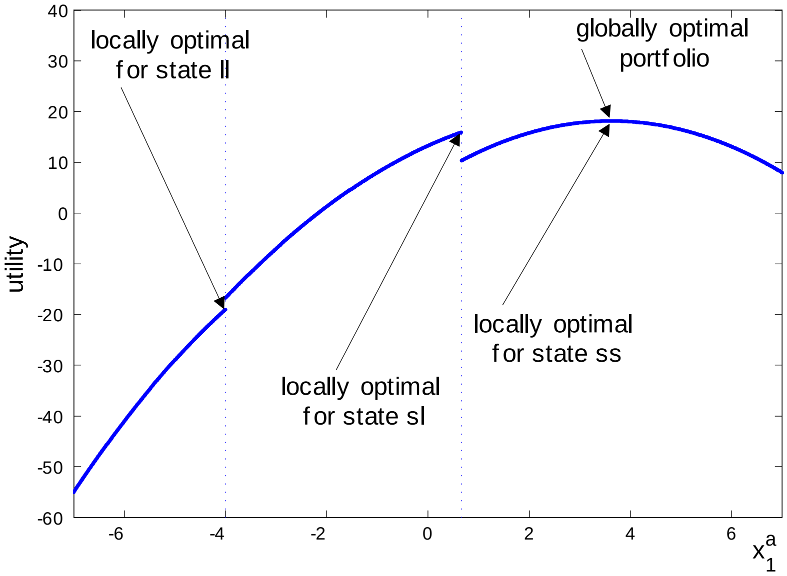
The dashed line shows the evolution of the gap and the solid line shows the evolution of the position of the monopolist arbitrageur as a function of time. The monopoly provides liquidity to local markets by trading at dates 0, 1 and 2, and the gap disappears at date 3 and remains closed thereafter. The model parameters are set to $\overline{g} = 10$ and $\lambda = 1$.

Figure 2-2: Equilibrium gap path and the optimal holdings of the duopoly over time.



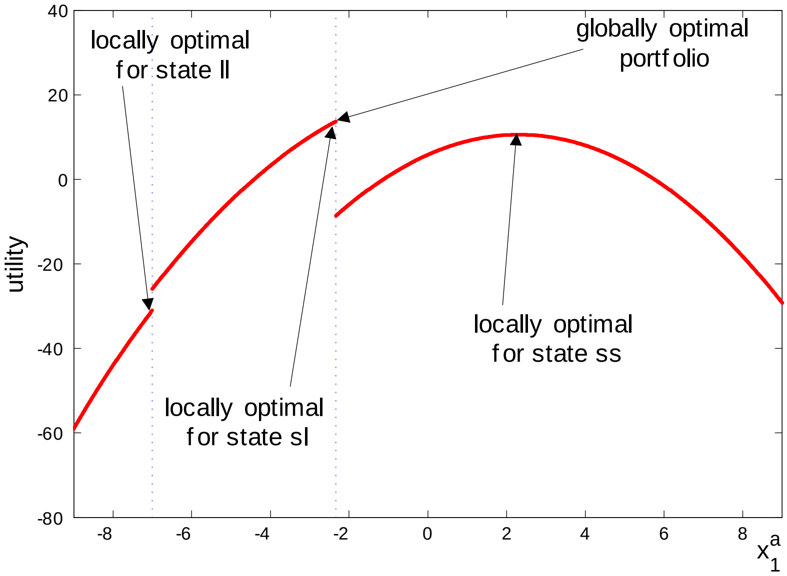
The left panel plots the evolution of the gap (dashed line) and the position of an unconstrained duopolist arbitrageur (solid line) as a function of time. The right panel compares the gap when a single arbitrageur (solid line) or two unconstrained arbitrageurs (dashed line) provide liquidity in local markets. The duopoly provides liquidity to local markets by trading at dates 0, 1 and 2, and the gap disappears at date 3 and remains closed thereafter. The model parameters are set to $\overline{g} = 10$ and $\lambda = 1$.

Figure 2-3: The utility of the aggressive arbitrageur as a function of her trade at date 1.



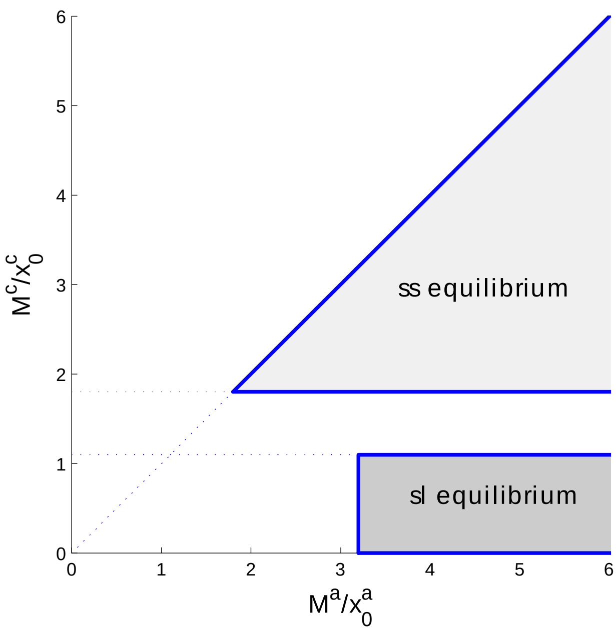
The utility of the aggressive arbitrageur as a function of her trade at date 1, x_1^a, while holding the cautious arbitrageur's date-1 trade constant at $x_1^c = 0$. The figure illustrates that the optimal strategy is to go short in the gap, and in this case both arbitrageurs remain solvent. The parameters are set to $g_0 = 3$, $\lambda = 1$, $M^a = 8$, $M^c = 10$, $x_0^a = 1$, $x_0^c = 3$.

Figure 2-4: The utility of the aggressive arbitrageur as a function of her trade at date 1.



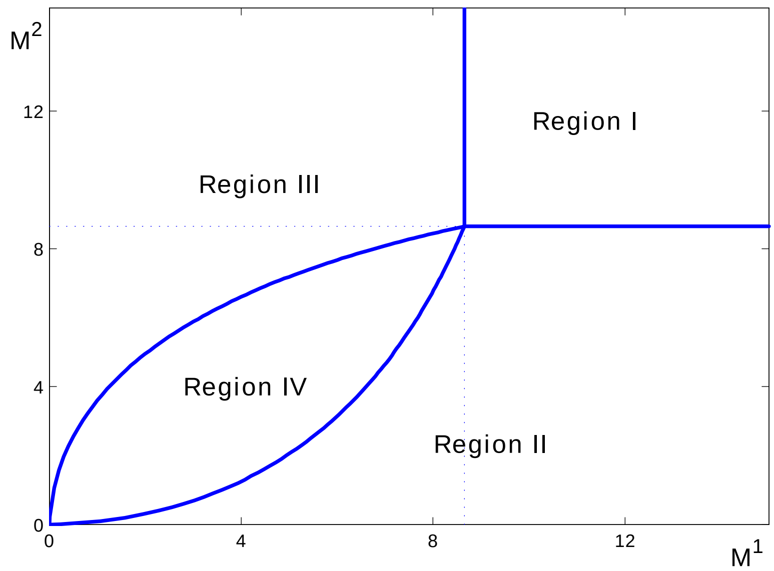
The utility of the aggressive arbitrageur as a function of her trade at date 1, x_1^a, while holding the cautious arbitrageur's date-1 trade constant at $x_1^c = 3$. The figure illustrates that the optimal strategy is to go long in the gap and force the cautious arbitrageur into distress. The parameters are set to $g_0 = 3$, $\lambda = 1$, $M^a = 8$, $M^c = 10$, $x_0^a = 1$, $x_0^c = 3$.

Figure 2-5: Capital thresholds for the different types of equilibria.



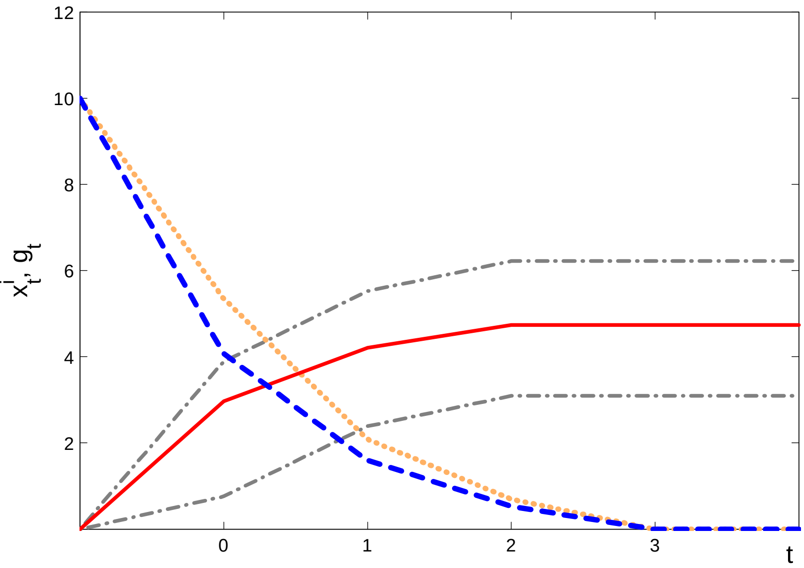
The horizontal axis plots the inverse of the proportion of capital invested in the arbitrage opportunity by the aggressive arbitrageur, M^a/x_0^a, and the vertical axis plots the same for the cautious arbitrageur, M^c/x_0^c. When both agents have low proportion of wealth invested in the risky assets, top right region, the wealth constraint does not affect arbitrage trading and the gap path, and an *sl* equilibrium exists. When arbitrageur a has a much lower proportion of wealth invested in the arbitrage opportunity that arbitrageur c, bottom left region, there exists an *sl* equilibrium. The aggressive arbitrageur predates on the cautious by widening the gap at date 1, and shorting it after the liquidation, at date 2. For other possible levels of proportion of capital invested in the arbitrage that satisfy $M^a/x_0^a \geq M^c/x_0^c$ there is no equilibrium.

Figure 2-6: Capital thresholds in the four different cases when arbitrageurs are subject to wealth constraints.



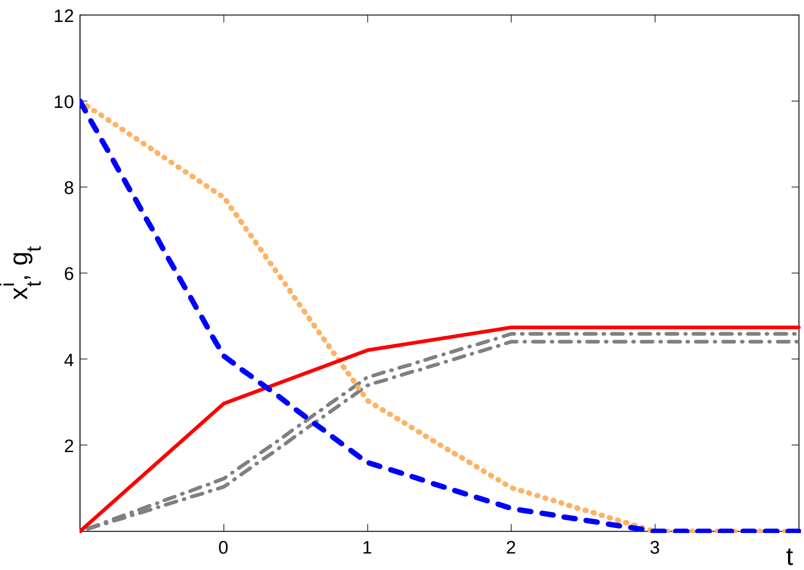
The horizontal axis plots the starting capital of arbitrageur 1, M^1, and the vertical axis plots the starting capital of arbitrageur 2, M^2. When both agents have large initial capital to invest, Region I, the wealth constraint does not affect arbitrage trading and the gap path. When at least one arbitrageur has a low capital level to start with, while the other has relatively more, i.e. Regions II and III, the wealth constraint affects the arbitrage positions but the gap path is close to the gap of the unconstrained case. Finally, when both arbitrageur have similarly low capital level to start with, Region IV, both arbitrageurs trade very little initially and the gap remains large. The model parameters are set to $\overline{g} = 10$ and $\lambda = 1$, which imply that the threshold for Region I is approximately $\frac{1}{\lambda}\Omega\overline{g}^2 \approx 8.651$.

Figure 2-7: Equilibrium gap path and the optimal holdings of the duopoly in the arbitrage opportunity over time.



The dotted line shows the evolution of the gap, and the dotted/dashed lines show the evolution of the positions of the constrained duopolist arbitrageurs as a function of time, when their initial capital levels M_1 and M_2 are significantly different from each other. To contrast, the dashed line shows the evolution of the gap and the solid line shows the evolution of the positions when the duopoly is unconstrained, as on Figure 2-2. Trading happens at dates 0, 1 and 2, and the gap disappears at date 3 and remains closed thereafter. The model parameters are set to $\overline{g} = 10$, $\lambda = 1$, $M_1 > 7.18$ and $M_2 = 5$.

Figure 2-8: Equilibrium gap path and the optimal holdings of the duopoly in the arbitrage opportunity over time.



The dotted line shows the evolution of the gap, and the dotted/dashed lines show the evolution of the positions of the constrained duopolist arbitrageurs as a function of time, when their initial capital levels M_1 and M_2 are similar. To contrast, the dashed line shows the evolution of the gap and the solid line shows the evolution of the positions when the duopoly is unconstrained, as on Figure 2-2. Trading happens at dates 0, 1 and 2, and the gap disappears at date 3 and remains closed thereafter. The model parameters are set to $\overline{g} = 10$, $\lambda = 1$, $M_1 = 1.17$ and $M_2 = 0.9$.

Chapter 3

On the uniqueness of equilibrium in the Grossman-Stiglitz noisy REE model

3.1 Introduction

In their seminal paper, Grossman and Stiglitz (1980) present a framework for a noisy rational expectations economy (REE), which since became a workhorse model studying asymmetric information in competitive financial markets. The basic purpose of the model is to resolve the paradox of fully revealing equilibria in models of asymmetric information: in such environments, prices would perfectly transmit the information of informed traders to uninformed ones, and therefore would imply that the value of information is zero. In their paper, they propose a model of asymmetric information and an equilibrium of the model, in which prices reflect the information of informed individuals, but only partially. Those who spend resources to obtain information do receive compensation from others without the information.

The standard Grossman and Stiglitz method, which has become widely used in models of asymmetric information, draws heavily on the fact that random variables are jointly normally distributed and investors have exponential (i.e. CARA) utilities, as it allows to work with linear demand functions and an equilibrium price function linear in state variables. However, the question whether there exist other equilibria of their model, which are potentially both less tractable and less appealing in their predictions, remains. In this chapter, which is joint work with Dömötör Pálvölgyi, we seek to explore this issue.

The main contribution of the chapter is to show that the model proposed by Grossman and Stiglitz indeed has a unique equilibrium when allowing for any continuous equilibrium price function, linear or not. Our solution method is different from the usual "conjecture and verify" approach, where the conjecture about a specific functional form of price naturally imposes limitations on the proof, and only allows to study existence and uniqueness in a particular class of functions. Here we propose a simple proof to show that no equilibrium besides the linear class can exist for this model, which hence complements the usual technique and leads to the conclusion that the unique equilibrium in the linear class of the Grossman and Stiglitz model is actually the sole equilibrium in the broader class of all continuous price functions as well.

Moreover, we show that relaxing the assumption of continuous prices leads to multiplicity in the possible equilibrium price functions. In particular, we provide an algorithm to create a (non-continuous) equilibrium price that is different from the Grossman-Stiglitz price function on a zero-measure set.

The noise in the REE proposed by Grossman and Stiglitz serves the purpose of revealing information only partially, therefore this chapter naturally belongs to the literature on partially-revealing rational expectations equilibria. Although the theory of fully-revealing REE is largely complete, with many studies on the generic existence and uniqueness, and some non-generic examples of non-existence (see, for example, Radner (1979), and Jordan (1982, 1983)), we know much less about partially-revealing REEs. Previous studies were mainly concerned about the existence of a rational expectations equilibrium in different settings, see for example Grossman (1976), Grossman and Stiglitz (1980), Hellwig (1980), Diamond and Verrecchia (1981), and Ausubel (1990). In contrast, we study uniqueness of the noisy REE in the model of Grossman and Stiglitz (1980). Our study is also related to DeMarzo and Skiadas (1998), who show uniqueness of a (non-noisy) perfectly revealing REE in the model of Grossman (1976), and give examples of partially revealing equilibria when payoffs are non-normal. Finally, Breon-Drish (2010) shows that in case of normality, the unique linear equilibrium is unique in the class of differentiable equilibrium price functions. Both of these studies use much more elaborate solution techniques than the present chapter, while restricting their attention in the possible equilibrium price functions.

The remainder of the chapter is organized as follows. Section 3.2 presents the model. Section 3.3 studies the linear equilibrium of the economy using the standard "conjecture and verify" method of Grossman and Stiglitz. Section 3.4 presents the argument for no equilibrium outside the linear class when allowing for any continuous price function. Section 3.5 provides an algorithm for a non-continuous price function. Finally, Section 3.6 concludes.

3.2 Model

This section introduces the baseline model. We consider a two-period economy with dates $t = 0$ and 1. Agents, specified later, trade at date 0. At date 1, assets pay off and agents consume.

3.2.1 Assets

There are two securities traded in a competitive market, a risk-free bond and a risky stock. The bond is in perfectly elastic supply and is used as numeraire, with the risk-free rate normalized to 0. The risky asset is assumed to be in random aggregate supply of $u \sim N\left(0, \sigma_u^2\right)$, and has random final payoff d at date 1, with ex ante distribution $d \sim N\left(\overline{d}, \sigma_d^2 = 1/\tau_d\right)$, which constitutes the common prior for all agents. The price of the stock at date 0 is denoted by p.

3.2.2 Agents

We assume that the asset market is populated by a continuum of agents (also called traders) with measure one. Agents do not hold endowments in the risky assets. Agent $k \in [0, 1]$ maximizes her negative exponential utility with CARA-coefficient α:

$$E\left[-\exp\left(-\alpha W_k\right)|\mathcal{I}_k\right],$$

where the final wealth $W_k = W_{k0} + x_k\left(d - p\right)$ is given by the starting wealth W_{k0}, plus the number of shares purchased, x_k, multiplied by the profit per share, $d - p$. $\mathcal{I}_k$ denotes the information set of a trader, and $E[.|\mathcal{I}_k]$ and $Var[.|\mathcal{I}_k]$ denote the expectation and variance conditional on the information set $\mathcal{I}_k$, respectively.

Traders can be either informed or uninformed. Informed traders, who form a mass of $0 < w \leq 1$, observe the signal $s = d + \eta$, with $\eta \sim N\left(0, \sigma_s^2 = 1/\tau_s\right)$.[1] The rest of the agents, with measure $1 - w$, are uninformed, and do not observe any signals about f. Instead, all agents of the model observe the market price p.[2] Formally, the information set of all informed traders is given by $\mathcal{I}^i = \{s, p\}$, but as the price does not contain more information about the final payoff than their private observation, $\{s\}$ is a sufficient statistics for $\{s, p\}$. The information set of uninformed traders is $\mathcal{I}^{ui} = \{p\}$. As agents inside the different investor classes are identical, we drop the subscript k from now on.

3.2.3 Equilibrium

We define an equilibrium of the above economy as follows.

Definition 34 *An equilibrium consists of a price function $P(s, u)$, and individual strategies for informed and uninformed traders, $x^i(s, p)$ and $x^{ui}(p)$, respectively, such that*

1. *demand is optimal for informed traders:*

$$x^i(s, p) \in \arg\max_{x \in \mathbb{R}} E\left[-\exp(-\alpha W^i)|s, p\right]; \tag{3.1}$$

2. *demand is optimal for uninformed traders:*

$$x^{ui}(p) \in \arg\max_{x \in \mathbb{R}} E\left[-\exp(-\alpha W^{ui})|p\right]; \tag{3.2}$$

3. *market clearing:*

$$wx^i(s, p) + (1 - w)\, x^{ui}(p) = u, \tag{3.3}$$

Conditions (3.1)-(3.3) define a competitive noisy rational expectations equilibrium. In particular, condition (3.1) states that individual asset demands are optimal for informed

[1] This assumption ensures that informed agents submit a downward-sloping demand curve.

[2] As it is standard in models with informational heterogeneity, the presence of random supply u makes sure that the price does not reveal d perfectly and hence the Grossman-Stiglitz paradox does not apply.

traders, conditioned on their observation and anything inferable from the price. Also, condition (3.2) states that individual asset demands are optimal for uninformed traders, conditioned on any information they infer from the price. Finally, (3.3) imposes that the market of the risky asset clears: aggregate demand equals supply.

3.3 The unique linear equilibrium

We first present the solution technique of Grossman and Stiglitz (1980). Solving for an equilibrium of the financial market requires three fairly standard steps. First, we postulate a REE price function. Given the price, we derive the optimal demand of uninformed traders. Finally, we check under what conditions the market clears at the conjectured price.

Following their seminal paper, we restrict our attention to the class of price functions that are linear in the state variables s and u. Formally, we conjecture a price function in the form

$$P(s,u) = A\overline{d} + B(s - Cu),$$

with constants A, B and C to be determined in equilibrium.

As it is well known, the assumption of the jointly normal distribution of d and s and the exponential utility implies that the informed optimization problem (3.1) simplifies to

$$x^i(s, P(s,u) = p) \in \arg\max_{x\in\mathbb{R}} E\left[W^i | s, P(s,u) = p\right] - \frac{\alpha}{2} Var\left[W^i | s, P(s,u) = p\right].$$

Moreover, the jointly normal distribution of d, s and u, and the conjectured linear form of the equilibrium price function implies that the price is normally distributed as well, which together with the exponential utility implies that the uninformed optimization problem (3.2) simplifies to

$$x^{ui}(P(s,u) = p) \in \arg\max_{x\in\mathbb{R}} E\left[W^{ui} | P(s,u) = p\right] - \frac{\alpha}{2} Var\left[W^{ui} | P(s,u) = p\right].$$

It is straightforward from here that optimal demands are the following: as informed traders cannot learn anything new from the price, they submit demand function

$$x^i(s,p) = \frac{E[d|s] - p}{\alpha Var[d|s]}, \tag{3.4}$$

while uninformed traders demand

$$x^{ui}(p) = \frac{E[d|p] - p}{\alpha Var[d|p]}. \tag{3.5}$$

Informed investors know the prior distribution of d and observe signal s, hence their posterior distribution is Gaussian with the mean and variance

$$E[d|s] = \frac{\tau_d \bar{d} + \tau_s s}{\tau_d + \tau_s} \text{ and } Var[d|s] = (\tau_d + \tau_s)^{-1}.$$

Given that uninformed agents know the prior distribution of d and observe the market price p, which is a normally distributed noisy signal about s and hence a normally distributed noisy signal about d itself, their posterior distribution is Gaussian with mean and variance:

$$E[d|p] = \frac{\tau_d \bar{d} + \tau_p \frac{p - A\bar{d}}{B}}{\tau_d + \tau_p} \text{ and } Var[d|p] = (\tau_d + \tau_p)^{-1},$$

where $\tau_p \equiv 1/\sigma_p^2$ denotes the precision of the price, with the price being a signal about payoff d with variance $\sigma_p^2 = \sigma_s^2 + C^2\sigma_u^2$.

The conjectured REE price function must equate demand and supply for each possible resolution of s and u. Substituting the optimal demands (3.4) and (3.5) into the market clearing condition (3.3) gives

$$w \frac{\frac{\tau_d \bar{d} + \tau_s s}{\tau_d + \tau_s} - p}{\alpha (\tau_d + \tau_s)^{-1}} + (1 - w) \frac{\frac{\tau_d \bar{d} + \tau_p \frac{p - A\bar{d}}{B}}{\tau_d + \tau_p} - p}{\alpha (\tau_d + \tau_p)^{-1}} = u.$$

Reorganizing the above equation, and using that in equilibrium the resulting coefficients must equal the conjectured A, B and C, we obtain the following result:

Theorem 35 (Grossman-Stiglitz) *In the family of linear equilibria there exists a unique*

REE of the model. The equilibrium price function is given in the form of

$$P(s,u) = A\overline{d} + B(s - Cu), \tag{3.6}$$

where

$$A = \frac{\tau_d}{\tau_d + w\tau_s + (1-w)\tau_p},\ B = 1 - A,\ \text{and } C = \frac{\alpha}{w\tau_s},$$

with $\tau_p = 1/\left(\sigma_s^2 + C^2\sigma_u^2\right)$.

3.4 The unique continuous price

In this section we show that the unique linear equilibrium of the model, solved by Grossman and Stiglitz (1980) and discussed in Section 3.3, is actually the unique equilibrium when allowing for any continuous price function.

The first observation we make is that the demand of informed traders is the same as before. Indeed, the only uncertainty they face is the unlearnable noise η, contained in their signal s, that is normally distributed. Therefore, the optimization problem of informed traders is unchanged and independent of the equilibrium price function. It implies that their demand is linear in signal s and price p:

$$x^i(s,p) = \frac{\frac{\tau_d\overline{d}+\tau_s s}{\tau_d+\tau_s} - p}{\alpha(\tau_d+\tau_s)^{-1}} = \frac{\tau_d}{\alpha}\overline{d} + \frac{\tau_s}{\alpha}s - \frac{\tau_d+\tau_s}{\alpha}p.$$

Second, whatever the equilibrium distribution of the market price is, uninformed traders' expectation and variance of d will be a function of the observed market price only, p, up to a constant. Therefore their demand is also a function of the equilibrium price p only; let us denote it by $x^{ui}(p)$. Market clearing implies that

$$w\left(\frac{\tau_d}{\alpha}\overline{d} + \frac{\tau_s}{\alpha}s - \frac{\tau_d+\tau_s}{\alpha}p\right) + (1-w)\,x^{ui}(p) = u,$$

that is

$$g(p) = s - Cu, \tag{3.7}$$

where

$$g(p) \equiv \frac{\tau_d + \tau_s}{\tau_s} p - (1-w) C x^{ui}(p) - \frac{\tau_d}{\tau_s} \bar{d},$$

and we already use the notation $C \equiv \alpha / (w\tau_s)$ from Theorem 35. Given the optimal demand function $x^{ui}(.)$ of uninformed investors (for which we want to solve for), one also knows the function g, and hence from a price realization $P(s,u) = p$ it is possible to compute $s - Cu$. Therefore, p always reveals $s - Cu$. The main question is whether p can tell more about s than just revealing $s - Cu$. In what follows, we make some simple observations based on (3.7) to argue that if $P(s,u)$ is continuous, it cannot. Hence p and $s - Cu$ are observationally equivalent.

Lemma 36 *For a given price realization p, the possible* (s,u) *pairs are all on a straight line.*

The proof of Lemma 36 is straightforward from (3.7). It implies that for a fixed $P(s,u) = p$ price the corresponding iso-price set on the (s,u) plane is a subset of a single straight line with tangency $1/C$ (see Figure 3-1). As such a line can be defined by its intercept with the horizontal axis, we can refer to it both as the line consisting of points that satisfy $s - Cu = l_p$ for a given constant l_p, or simply denote it l_p.

Next, we argue that given a p realization of the market price, uninformed agents cannot learn more about the signal s than its linear combination with the supply shock, $s - Cu$. Suppose that the converse is true, which implies that $P(s,u)$ must be a function of s not only through $s - Cu$. If this is the case, there must be two price realizations p_1 and p_2, such that $p_1 \neq p_2$ but $g(p_1) = g(p_2)$. It is equivalent to say that there are two pairs, $(s_1, u_1) \neq (s_2, u_2)$, such that they correspond to the two different prices, $P(s_1, u_1) = p_1$ and $P(s_2, u_2) = p_2$ while they are on the same line: $s_1 - Cu_1 = l_p = s_2 - Cu_2$ with $l_p \equiv l_{p_1} = l_{p_2}$.

As $P(s,u)$ is a continuous function of the random variables s and u, the Intermediate Value Theorem (see, for example, Bartle (1976), p. 153.) implies that if we connect the two points (s_1, u_1) and (s_2, u_2) with any simple curve γ, then there must be at least one point (s,u) on γ such that

$$P(s,u) = \frac{p_1 + p_2}{2}.$$

We will apply this theorem to two curves. The first curve will be simply the segment connecting (s_1, u_1) and (s_2, u_2), which is a part of line l_p. Take a point from this segment, denoted by (s^*, u^*), such that

$$P\left(s^*, u^*\right) = \frac{p_1 + p_2}{2},$$

as illustrated on Figure 3-2. (If there are at least two, take any one of them.) The second will be any curve whose intersection with the line is only (s_1, u_1) and (s_2, u_2). This will give a point outside l_p, denoted by (s_*, u_*), such that

$$P\left(s_*, u_*\right) = \frac{p_1 + p_2}{2}.$$

Given that $(s_*, u_*) \notin l_p$, it must be that $s_* - Cu_* \neq l_p$. Hence we found two points of the (s, u) plane such that they admit the same price, $P\left(s^*, u^*\right) = P\left(s_*, u_*\right) = \frac{p_1+p_2}{2}$, but $g\left(\frac{p_1+p_2}{2}\right) = s_* - Cu_* \neq s^* - Cu^* = g\left(\frac{p_1+p_2}{2}\right)$. This is clearly a contradiction.

Therefore, it must be that a price realization $P(s, u) = p$ is equivalent to observing the random variable $\widetilde{p} \equiv s - Cu$, i.e. the $p \longmapsto l_p$ mapping is a one-to-one mapping. If both s and u are normally distributed, $\widetilde{p}$ is normally distributed as well, and combining it with the normally distributed prior leads to Gaussian conditional distributions. Hence, uninformed agents' optimization program is necessarily a CARA-normal setting, with optimal demand

$$x^{ui}\left(\widetilde{p}\right) = \frac{E\left[d|\widetilde{p}\right] - p}{\alpha Var\left[d|\widetilde{p}\right]},$$

and the expectation and variance given by

$$E\left[d|\widetilde{p}\right] = \frac{\tau_d \overline{d} + \tau_{\widetilde{p}} \widetilde{p}}{\tau_d + \tau_{\widetilde{p}}} \text{ and } Var\left[d|\widetilde{p}\right] = \left(\tau_d + \tau_{\widetilde{p}}\right)^{-1}$$

for some $\tau_{\widetilde{p}}$. Market clearing then becomes

$$w\left(\frac{\tau_d}{\alpha}\overline{d} + \frac{\tau_s}{\alpha}s - \frac{\tau_d + \tau_s}{\alpha}p\right) + (1 - w)\frac{\tau_d \overline{d} + \tau_{\widetilde{p}}\widetilde{p} - \left(\tau_d + \tau_{\widetilde{p}}\right)p}{\alpha\left(\tau_d + \tau_{\widetilde{p}}\right)^{-1}} = u,$$

therefore p is linear in s, u and $\widetilde{p}$. Using $\widetilde{p} \equiv s - Cu$ we obtain that p must be linear in the state variables s and u. As we already showed that in the linear class there is a single equilibrium price function, we can conclude that the well-known Grossman-Stiglitz unique linear equilibrium is actually the unique equilibrium of the economy when allowing for any continuous price function:

Theorem 37 *In the family of continuous equilibrium price functions, there exists a unique REE of the model. The equilibrium price function is given in the linear form of*

$$P(s,u) = A\overline{d} + B[s - Cu],$$

where the coefficients A, B and C are those given in Theorem 35.

3.5 Non-continuous price functions

If the price function P does not have to be continuous but only measurable, then it is possible to have several equilibrium price functions. In this section we will show the existence of such price functions.

We construct a P that is non-continuous, and different from the Grossman-Stiglitz price function on a zero-measure subset of the whole (s,u) plane. We start with a definition.

Definition 38 *We say that a region R in the (s,u) plane is P-*homogenous *if $P(s,u)$ is constant. We also call region R p-*homogenous*, if $P(s,u) = p$ for all $(s,u) \in R$.*

From the definition of the equilibrium it is clear that the P-homogenous regions uniquely determine P. Moreover, let us denote the unique linear price function by P_0; we have seen that the P_0-homogenous regions are exactly the lines with tangency $1/C$.

Definition 39 *We further say that P_2 is a* refinement *of P_1 if every P_2-homogenous region is contained in a P_1-homogenous region.*

Note that Lemma 36 claims that every price function is a refinement of P_0.

To make a non-continuous valid price function, we will start from the unique linear equilibrium, P_0 and we will change it on some lines with tangency $1/C$. The following observation will be crucial:

Claim 40 *Suppose we take a refinement of P_0, P, such that each P-homogenous region is a line or halfline with tangency $1/C$. Suppose that on a line l with tangency $1/C$ the unique linear price took value p, and the two P-homogenous halflines are separated by the point $t \in l$. Denote by $p_-(t)$ the price on the left side of t (i.e. the set of points on l that have lower s-coordinate than t), and by $p_+(t)$ the price on the right side of t. Then we have $p_-(t) < p < p_+(t)$.*

Proof. The proof follows from the fact that the random variable describing the set of potential s values on the right half of the line first-order stochastically dominates the variable if s can be on the whole line, therefore any expected utility-maximizing uninformed agent prefers it. It increases their demand, and hence, by market clearing, it must fetch a higher price. A similar argument applies to the left side too. ■

We define a sequence of functions $\{P_i\}_{i=0}^{\infty}$ such that we start from the unique linear equilibrium P_0, and after countably many steps we obtain a non-continuous valid price function. In every step i we refine the previous function by splitting some P_{i-1}-homogenous lines into two P_i-homogenous halflines.

In the beginning we pick an arbitrary line with tangency $1/C$, denote it by l. We split it arbitrarily using Claim 40, obtaining two new price values, p_- and p_+, on this line. This gives us P_1. Notice that P_1 is not a valid price function, because it takes the value p_- on two different lines with tangency $1/C$: on (some part of) l and on l_{p_-}, which was the p_--homogenous line of P_0 before. Similarly, P_1 takes the value p_+ on some part of l and on l_{p_+}, which was the p_+-homogenous line of P_0 before. This would contradict Lemma 36, hence we have to modify P_1.

In the second step, we split l_{p_-} and l_{p_+} into two halflines, again using Claim 40. We take care that the four new prices that we obtain correspond to four lines that we have not modified yet. As we have only modified three lines so far, it is possible to find appropriate points for this splitting. This gives us P_2. Figures 3-3, 3-4 and 3-5 show the first two steps of this construction. As before, P_2 is not a valid price function, because there are some price realizations that correspond to more than one line, contradicting Lemma 36. Thus in the third step we modify these four lines, and so on.

After the first n steps we have already modified $2^n - 1$ lines. On $2^{n-1} - 1$ of them, P_n is in line with Lemma 36, but on the other 2^{n-1} the price function P_n takes 2^n values that are used on 2^n yet unchanged lines. Using Claim 40 we split these 2^n lines into 2^{n+1} halflines. Computing the demand of uninformed investors and applying the market clearing condition for all these halflines can yield 2^{n+1} new prices. This gives us P_{n+1}, which is still not a valid price function, and so on.

This algorithm gives a measurable price function that is continuous except for on countably many lines. Notice that this function is also the same as the original unique linear equilibrium except for a measure-zero set: the countable set of lines we have modified.

3.6 Conclusion

The standard method of conjecturing and then verifying a linear equilibrium price function has become widely used in models of asymmetric information. While papers following this technique show that the price function is unique in the linear class, they claim we do not know anything outside the linear class. Hence it is important to study whether there exist other equilibria of such a model, which are potentially both less tractable and less appealing in their predictions.

In this chapter we explore this question. Our solution method is different from the usual "conjecture and verify" approach, where the conjecture about a specific functional form of price naturally imposes limitations on the proof. Our contribution complements the usual techniques and leads to the conclusion that the unique linear equilibrium of the Grossman and Stiglitz model is unique when allowing for any continuous price function. We also provide an algorithm to create a (non-continuous) equilibrium price that is different from the Grossman-Stiglitz price function.

In this chapter we restrict our attention to the assumptions of the original Grossman and Stiglitz (1980) paper. There are some straightforward ways to generalize our results. First, our result on the non-continuous price function is probably extendable to a price function which is non-continuous on a positive measure set. Second, it would be interesting to see whether we could provide similar statements about the equilibrium in the same setting but with more general parameter distributions and utility functions. Third, an important question is whether our result would stay in a modification of the model with incorporating imperfect competition, as in Kyle (1989). Finally, it would be interesting to study whether other equilibria exist in certain settings of information aggregation in financial markets, when agents have differential information, and all of them have something to learn from the price, e.g. in Hellwig (1980) or Diamond and Verrecchia (1981). These problems are left for future research.

Figure 3-1: The set of (s, u) combinations that satisfy $s - Cu = l_p$.

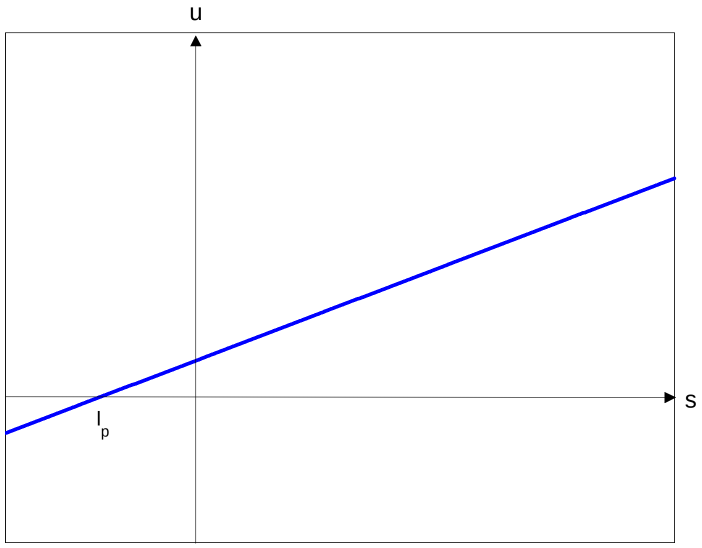


Figure 3-2: The relationship of l_p, (s^*, u^*) and (s_*, u_*).

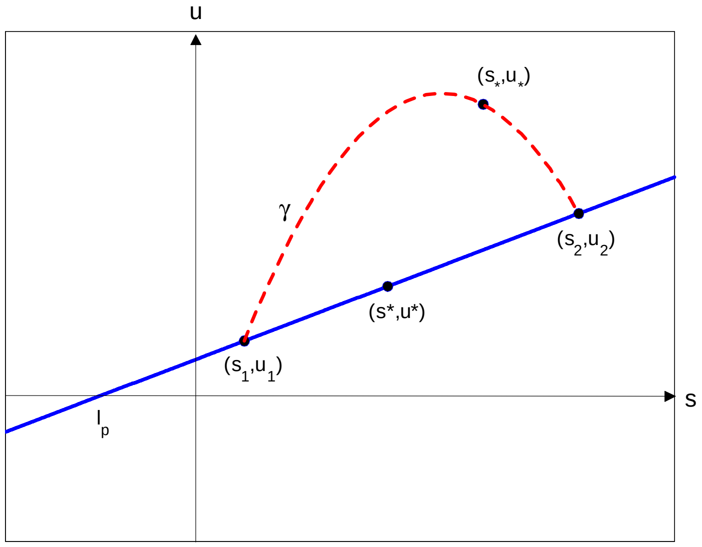
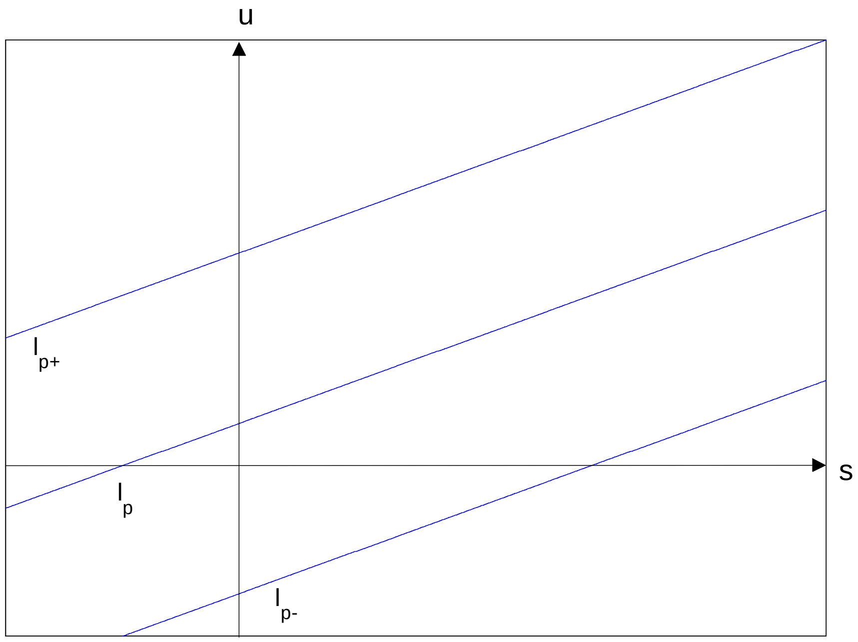
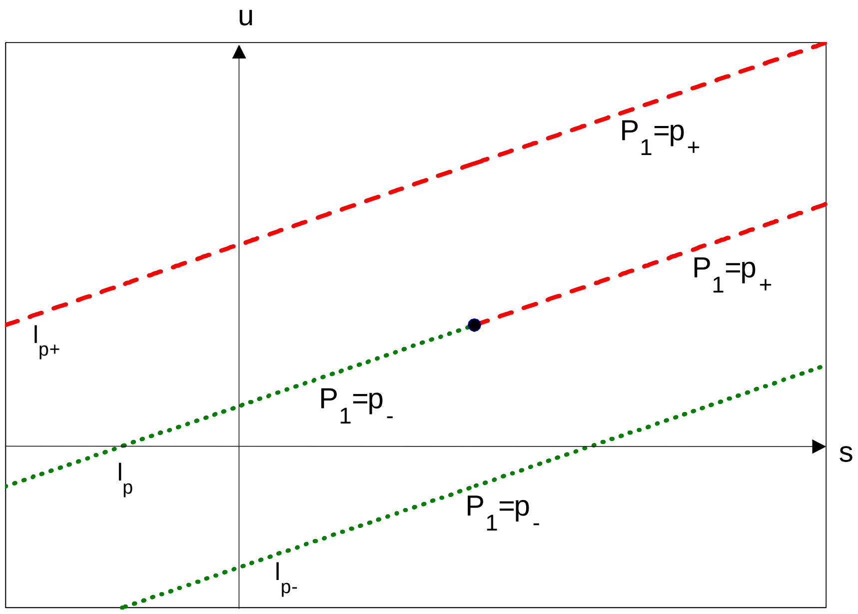


Figure 3-3: The unique continuous price function P_0.



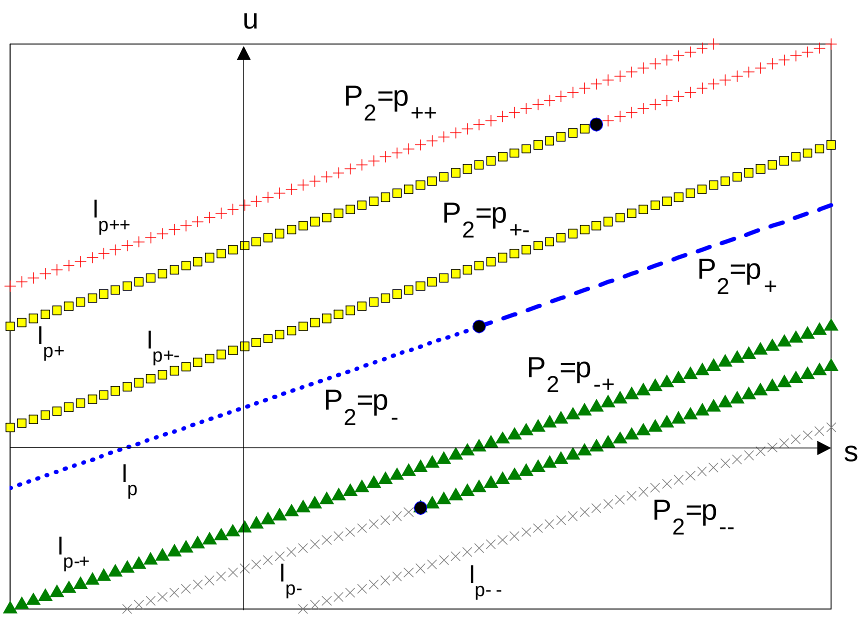
On line l_p it satisfies $P_0 = p$, on line l_{p_-} it takes the value $P_0 = p_-$, and on line l_{p_+} it takes the value $P_0 = p_+$.

Figure 3-4: The construction of P_1.



We take an arbitrary point on l_p and split it into two halflines. We compute uninformed traders's demand if they know s is on the halfline left to the splitting point, and plug it into the market-clearing condition, which leads to the price p_-. We proceed similarly with the halfline right to the splitting point and obtain p_+. Notice that it is not a valid price function yet, as for example $P_1 = p_+$ on some part of l_p and of course on the original l_{p_+} (illustrated by dashed lines), contradicting Lemma 36. Also, $P_1 = p_-$ on some part of l_p and of course on the original l_{p_-} (illustrated by dotted lines). Therefore, in the next step we have to modify the prices on l_{p_+} and l_{p_-}.

Figure 3-5: The construction of P_2.



We take one arbitrary point on l_{p_+} and one on l_{p_-}, and split each to two halflines. For all four halflines we compute the uninformed demands conditional on knowing that s is on the appropriate halfline, and then obtain the market clearing prices for them. In this process we make sure none of the prices are p, p_+ or p_-, otherwise we choose a different splitting. Notice that now there is no problem with prices p_+ and p_-, as they are fetched only if s is on a part of l_p. However, P_2 is still not a valid price function, because there are four price levels, p_{++}, p_{+-}, p_{-+}, and p_{--}, that are achieved on both a halfline and a line, contradicting Lemma 36 (illustrated by $+$, $\square$, $\triangle$ and $\times$, respectively). Hence, in the next step we must modify P_2 on the lines $l_{p_{++}}$, $l_{p_{+-}}$, $l_{p_{-+}}$, and $l_{p_{--}}$.

Bibliography

Abreu, D. and M. K. Brunnermeier (2002). Synchronization risk and delayed arbitrage. *Journal of Financial Economics 66*, 341–360.

Abreu, D. and M. K. Brunnermeier (2003). Bubbles and crashes. *Econometrica 71*, 173–204.

Almazan, A., K. C. Brown, M. Carlson, and D. A. Chapman (2004). Why constrain your mutual fund manager? *Journal of Financial Economics 73*, 289–321.

Angeletos, G.-M., C. Hellwig, and A. Pavan (2006). Signaling in a global game: coordination and policy traps. *Journal of Political Economy 114*, 452–484.

Angeletos, G.-M., C. Hellwig, and A. Pavan (2008). Robust predictions in global games with multiple equilibria: defense policies against currency attacks. Working paper, MIT.

Angeletos, G.-M. and I. Werning (2006). Crises and prices: information aggregation, multiplicity, and volatility. *American Economic Review 96*, 1720–1736.

Asquith, P., R. Gertner, and D. Scharfstein (1994). Anatomy of financial distress: an examination of junk-bond issuers. *Quarterly Journal of Economics 109*, 625–658.

Atkeson, A. (2001). Discussion on Morris and Shin. *NBER Macroeconomics Annual (2000)*, 162–171.

Attari, M. and A. S. Mello (2006). Financially constrained arbitrage in illiquid markets. *Journal of Economic Dynamics and Control 30*, 2793–2822.

Ausubel, L. M. (1990). Partially-revealing rational expectations equilibrium in a competitive economy. *Journal of Economic Theory 50*, 93–126.

Bai, Y., E. C. Chang, and J. Wang (2006). Asset prices under short-sale constraints. Working paper, MIT.

Barlevy, G. and P. Veronesi (2003). Rational panics and stock market crashes. *Journal of Economic Theory 110*, 234–263.

Bartle, R. G. (1976). *The elements of real analysis.* Wiley, New York, NY.

Beber, A. and M. Pagano (2011). Short-selling bans around the world: evidence from the 2007-09 crisis. *Journal of Finance*, forthcoming.

Berk, J. B. and R. C. Green (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy 112*, 1269–1295.

Bernardo, A. E. and I. Welch (2004). Liquidity and financial market runs. *Quarterly Journal of Economics 119*, 135–158.

Boehmer, E., C. M. Jones, and X. Zhang (2009). Shackling short sellers: the 2008 shorting ban. Working paper, Columbia University.

Breon-Drish, B. (2010). Asymmetric information in financial markets: anything goes. Working paper, University of California, Berkeley.

Bris, A., W. N. Goetzmann, and N. Zhu (2007). Efficiency and the bear: short sales and markets around the world. *Journal of Finance 62*, 1029–1079.

Brunner, A. and J. P. Krahnen (2008). Multiple lenders and corporate distress: evidence on debt restructuring. *Review of Economic Studies 75*, 415–442.

Brunnermeier, M. K. (2009). Deciphering the liquidity and credit crunch 2007-2008. *Journal of Economic Perspectives 23*, 77–100.

Brunnermeier, M. K. and L. H. Pedersen (2005). Predatory trading. *Journal of Finance 60*, 1825–1863.

Brunnermeier, M. K. and L. H. Pedersen (2009). Market liquidity and funding liquidity. *Review of Financial Studies 22*, 2201–2238.

Carlin, B. I., M. S. Lobo, and S. Vishwanathan (2007). Episodic liquidity crises: cooperative and predatory trading. *Journal of Finance 62*, 2235–2274.

Carlsson, E. and E. van Damme (1993). Global games and equilibrium selection. *Econometrica 61*, 989–1018.

Chamley, C. (1999). Coordinating regime switches. *Quarterly Journal of Economics 114*, 869–905.

Chen, Q., I. Goldstein, and W. Jiang (2010). Payoff complementarities and financial fragility - evidence from mutual fund outflows. *Journal of Financial Economics 97*, 239–262.

Chevalier, J. and G. Ellison (1997). Risk taking by mutual funds as a response to incentives. *Journal of Political Economy 105*, 1167–1200.

Coval, J. and E. Stafford (2007). Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics 86*, 479–512.

Danielsson, J., H. S. Shin, and J.-P. Zigrand (2004). The impact of risk regulation on price dynamics. *Journal of Banking and Finance 28*, 1069–1087.

Danielsson, J., H. S. Shin, and J.-P. Zigrand (2011). Balance sheet capacity and endogenous risk. Working paper, LSE.

Dasgupta, A. (2007). Coordination and delay in global games. *Journal of Economic Theory 134*, 195–225.

DeMarzo, P. and C. Skiadas (1998). Aggregation, determinacy, and informational efficiency for a class of economies with asymmetric information. *Journal of Economic Theory 80*, 123–152.

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

Diamond, D. W. and R. E. Verrecchia (1981). Information aggregation in a noisy rational expectations economy. *Journal of Financial Economics 9*, 221–235.

Diamond, D. W. and R. E. Verrecchia (1987). Constraints on short-selling and asset price adjustment to private information. *Journal of Financial Economics 18*, 277–311.

Duffie, D., N. Garleanu, and L. H. Pedersen (2002). Securities lending, shorting, and pricing. *Journal of Financial Economics 66*, 307–339.

Edwards, F. R. (1999). Hedge funds and the collapse of Long-Term Capital Management. *Journal of Economic Perspectives 13*, 189–210.

Financial Services Authority (2009). Short selling. Discussion paper.

Garleanu, N. and L. H. Pedersen (2007). Liquidity and risk management. *American Economic Review Papers and Proceedings 97*, 193–197.

Garleanu, N. and L. H. Pedersen (2011). Margin-based asset pricing and the law of one price. *Review of Financial Studies 24*, 1980–2022.

Gilchrist, S., C. P. Himmelberg, and G. Huberman (2005). Do stock price bubbles influence corporate investment? *Journal of Monetary Economics 52*, 805–827.

Goldstein, I. and A. Guembel (2008). Manipulation and the allocational role of prices. *Review of Economic Studies 75*, 133–164.

Goldstein, I., E. Ozdenoren, and K. Yuan (2009). Trading frenzies and their impact on real investment. Working paper, LSE.

Goldstein, I. and A. Pauzner (2005). Demand-deposit contracts and the probability of bank runs. *Journal of Finance 60*, 1293–1327.

Gromb, D. and D. Vayanos (2002). Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics 66*, 361–407.

Grossman, S. J. (1976). On the efficiency of competitive stock markets where traders have diverse information. *Journal of Finance 31*, 573–585.

Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient markets. *American Economic Review 70*, 393–408.

Harrison, J. M. and D. M. Kreps (1978). Speculative investor behavior in a stock market with heterogeneous expectations. *Quarterly Journal of Economics 92*, 323–336.

Hellwig, C. (2002). Public information, private information, and the multiplicity of equilibria in coordination games. *Journal of Economic Theory 107*, 191–222.

Hellwig, C., A. Mukherji, and A. Tsyvinski (2006). Self-fulfilling currency crises: the role of interest rates. *American Economic Review 96*, 1769–1787.

Hellwig, M. F. (1980). On the aggregation of information in competitive markets. *Journal of Economic Theory 22*, 477–498.

Hertzberg, A., J. M. Liberti, and D. Paravisini (2010). Public information and coordination: evidence from a credit registry expansion. *Journal of Finance 66*, 379–412.

Ho, K. W. (1996). Short-sales restrictions and volatility: the case of the stock exchange of singapore. *Pacific-Basin Finance Journal 4*, 377–391.

Hong, H. and J. Stein (2003). Differences of opinion, short-sales constraints, and market crashes. *Review of Financial Studies 16*, 377–391.

Ivashina, V. and D. S. Scharfstein (2010). Bank lending during the financial crisis of 2008. *Journal of Financial Economics 97*, 319–338.

Jones, C. M. and O. A. Lamont (2002). Short sale constraints and stock returns. *Journal of Financial Economics 97*, 319–338.

Jordan, J. (1982). The generic existence of rational expectations equilibrium in the higher dimensional case. *Journal of Economic Theory 26*, 224–243.

Jordan, J. (1983). On the efficient market hypothesis. *Econometrica 51*, 1325–1344.

Kondor, P. (2009). Risk in dynamic arbitrage: the price effects of convergence trading. *Journal of Finance 64*, 631–655.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica 53*, 1315–1336.

Kyle, A. (1989). Informed speculation with imperfect competition. *Review of Economic Studies 56*, 317–355.

Liu, J. and F. Longstaff (2004). Losing money on arbitrage: optimal dynamic portfolio choice in markets with arbitrage opportunities. *Review of Financial Studies 17*, 611–641.

Liu, X. (2010). Predatory short-selling and self-fulfilling crises. Working paper, HKUST.

Loewenstein, R. (2000). *When genius failed: the rise and Fall of Long-Term Capital Management*. Random House, New York.

Marin, J. M. and J. P. Olivier (2008). The dog that did not bark: insider trading and crashes. *Journal of Finance 63*, 2429–2476.

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion. *Journal of Finance 32*, 1151–1168.

Morris, S. and H. S. Shin (1998). Unique equilibrium in a model of self-fulfilling currency attacks. *American Economic Review 88*, 587–597.

Morris, S. and H. S. Shin (1999). Private versus public information in coordination problems. Working paper, Yale University and University of Oxford.

Morris, S. and H. S. Shin (2001). Rethinking multiple equilibria in macroeconomics. *NBER Macroeconomics Annual (2000)*, 139–161.

Morris, S. and H. S. Shin (2002). The social value of public information. *American Economic Review 92*, 1521–1534.

Morris, S. and H. S. Shin (2003). Global games: theory and applications. In *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*. Cambridge University Press, Cambridge, UK.

Morris, S. and H. S. Shin (2004). Coordination risk and the price of debt. *European Economic Review 48*, 133–153.

Nagel, S. (2005). Short sales, institutional investors and the cross-section of stock returns. *Journal of Financial Economics 75*, 277–309.

Obstfeld, M. (1986). Rational and self-fulfilling balance-of-payments crises. *American Economic Review 76*, 72–81.

Obstfeld, M. (2004). Models of currency crises with self-fulfilling features. *European Economic Review 40*, 1037–1047.

Ofek, E. and M. Richardson (2003). DotCom mania: the rise and fall of internet stock prices. *Journal of Finance 58*, 1113–1137.

Ozdenoren, E. and K. Yuan (2008). Feedback effects and asset prices. *Journal of Finance 63*, 1939–1975.

Panageas, S. (2003). Speculation, overpricing, and investment - theory and empirical evidence. Working paper, MIT.

Radner, R. (1979). Rational expectations equilibrium: generic existence and the information revealed by prices. *Econometrica 47*, 655–678.

Reed, A. V. (2007). Costly short selling and stock price adjustment to earnings announcements. Working paper, University of North Carolina.

Rochet, J.-C. and X. Vives (2004). Coordination failures and the lender of last resort: was bagehot right after all? *Journal of the European Economic Association 2*, 1116–1147.

Rubinstein, M. E. (2004). Great moments in financial economics: III. Short-sales and stock prices. *Journal of Investment Management 2*.

Scheinkman, J. A. and W. Xiong (2003). Overconfidence and speculative bubbles. *Journal of Political Economy 111*, 1183–1219.

Shin, H. S. (2009). Reflections on northern rock: the bank run that heralded the global financial crisis. *Journal of Economic Perspectives 23*, 101–119.

Shleifer, A. and R. W. Vishny (1997). The limits of arbitrage. *Journal of Finance 52*, 35–55.

Sirri, E. R. and P. Tufano (1998). Costly search and mutual fund flows. *Journal of Finance 53*, 1589–1622.

Tarashev, N. A. (2007). Speculative attacks and the information role of the interest rate. *Journal of the European Economic Association 5*, 1–36.

Van Nieuwerburgh, S. and L. Veldkamp (2009). Information immobility and the home bias puzzle. *Journal of Finance 64*, 1187–1215.

Van Nieuwerburgh, S. and L. Veldkamp (2010). Information acquisition and under-diversification. *Review of Economic Studies 77*, 779–805.

Xiong, W. (2001). Convergence trading with wealth effects. *Journal of Financial Economics 62*, 379–411.

Yuan, K. (2005). Asymmetric price movements and borrowing constraints: a REE model of crises, contagion, and confusion. *Journal of Finance 60*, 247–292.

Yuan, K. (2006). The price impact of borrowing and short-sale constraints. Working paper, University of Michigan.

Zigrand, J.-P. (2004). A general equilibrium analysis of strategic arbitrage. *Journal of Mathematical Economics 40*, 923–952.